CONSUMER PRICE INDEX THEORY

Chapter Titles

CHAPTER 6: ELEMENTARY INDEXES[1]

Draft: April 16, 2021

Erwin Diewert, University of British Columbia

Table of Contents

---

[1] This chapter draws heavily on Chapter 20 of the *Consumer Price Index Manual*; see ILO IMF OECD UNECE Eurostat World Bank (2004; 355-371). The author thanks Elizabeth Abraham, Carsten Boldsen, Yuri Dikhanov, Kevin Fox, Heidi Ertl, Robert Hill, Ronald Johnson, Claude Lamboray, Cindy MacDonald, Marshall Reinsdorf, Mark Ruddock, Chihiro Shimizu, Mick Silver, Jasmin Whelan and Grant Yake for helpful comments.

1. Introduction

In all countries, the calculation of a Consumer Price Index proceeds in two (or more) stages. In the first stage of calculation, *elementary price indexes* are calculated for the *elementary expenditure aggregates* of a CPI. In the second and higher stages of aggregation, these elementary price indexes are combined to obtain higher level indexes using information on the expenditures on each elementary aggregate as weights. An *elementary aggregate* consists of the expenditures by a specified group of consumers on a relatively homogeneous set of products defined within the consumption classification used in the CPI.

At the first stage of aggregation, one of two possible situations can occur:

- Detailed price and quantity (or price and value) information on all transacted products in the elementary aggregate is available for the time period under consideration.[2]
- Only price information is available for the products in the aggregate under consideration. Moreover, the price information may be collected only for a *sample* of the entire set of product prices that are in scope.

At higher levels of aggregation, typically price and quantity (or value) information is available. Thus for higher levels of aggregation and for situations where detailed price and quantity information is available at the first stage of aggregation, the materials in previous chapters can be applied; i.e., Lowe, Laspeyres, Paasche and Fisher indexes can be used at higher levels of aggregation and at the elementary level if detailed price and quantity information is available. However, for situations where quantity or value information is not available, most of the index number theory outlined in previous chapters is not directly applicable. In this case, an elementary price index is a more primitive concept that relies on price data only. The situation where only price information is available will be the focus of this chapter. However, some elementary indexes can be constructed using price and quantity (or expenditure) data and so some attention will be paid to this situation as well.[3]

The question of what is the most appropriate formula to use to construct an elementary price index is considered in this chapter.[4] The quality of a CPI depends heavily on the quality of the first stage of aggregation elementary indexes, which are the basic building blocks from which Consumer Price Indexes are constructed.

CPI compilers have to select *representative products* within an elementary aggregate and then collect a sample of prices for each of the representative products, usually from a sample of different outlets. The individual products whose prices are actually collected are described as the *sampled products*. Their prices are collected over successive time periods. An elementary price index is therefore typically calculated from two sets of *matched price observations*. In this chapter, we will assume that there are no missing observations and no changes in the quality of the products sampled so that the two sets of prices are perfectly matched. In the following

[2] With the increased availability of scanner data both for retail outlets as well as for individual consumers, the first situation is increasingly likely. Also it may be the case that the statistical office will have access to price and quantity data on deliveries to households from regulated electricity and telecom firms. In the Appendix to this chapter, we will use such a data set for the UK fixed line telecom sector in order to show how the various elementary indexes to be considered below perform in practice.

[3] Thus scanner data is increasingly being applied at the elementary level by national statistical agencies. The use of scanner data can lead to chain drift problems which will be addressed in the following chapter.

[4] The material in this chapter draws heavily on the contributions of Dalén (1992), Balk (1994) (2002) (2008) and Diewert (1995) (2002).

chapter, we will consider alternative strategies when there are multiple time periods and missing observations; i.e., in chapter 7, we will discuss *multilateral index number theory*. In chapter 8, the treatment of new and disappearing goods and services and the related problems associated with measuring *quality change* will be discussed.

Before we define the elementary indexes used in practice, we will first consider in section 2 what is a suitable definition for an *ideal elementary index*. An *ideal index* will make use of expenditure data (as well as price data) even though it cannot always be implemented in practice due to lack of expenditure and quantity data. The problems involved in aggregating transaction prices for the same product over *time* are also discussed in this section. In general, the discussion in section 2 provides a theoretical target index target index that uses both price and quantity information. "Practical" elementary price indexes that are constructed using only information on prices will be discussed in subsequent sections.

Section 3 provides some additional discussion about the problems involved in picking a suitable level of disaggregation for the elementary aggregates. Should the elementary aggregates have a regional dimension in addition to a product dimension? Should prices be collected from retail outlets or from households? These are the types of question discussed in this section.

Section 4 introduces the main elementary index formulae that are used in practice and section 5 develops some numerical relationships between the various "practical" indexes. These relationships will be illustrated for a particular data set in Appendix A to this chapter.

Section 6 develops the axiomatic or *test approach* to bilateral elementary indexes when only price information is available.

Section 7 contains some material on the importance of the time reversal test.

Section 8 concludes with an overview of the various results.

Appendix A looks at the problems that arise when households have to pay a fixed fee to gain access to various products or services that a firm sells. For the most part, these access fees are not very large so their treatment in a CPI does not make a material difference. However, in the case of telecommunication services, alternative treatments of access fees lead to very different price (and quantity) indexes as will be seen in the Appendix. Also, as mentioned above, the numerical relationships between the various elementary indexes which are developed in section 5 below will be illustrated in Appendix A with actual telecom data from the UK.

Appendix B lists the objections to the use of the Carli index that were made by Robert Hill in his testimony to the UK House of Lords on the use of the Carli index in the Retail Price Index that was once used in the UK.

2. Ideal Elementary Indexes

The aggregates covered by a CPI are usually arranged in the form of a tree like hierarchy, such as COICOP (Classification of Individual Consumption by Purpose). An aggregate is a set of economic transactions pertaining to a set of commodities and a set of economic agents over a specified time period. Every economic transaction relates to the change of ownership of a specific, well defined commodity (good or service) at a particular place and date, and comes with a quantity and a price. A price index for an aggregate is typically calculated as a weighted average of the price indexes for the subaggregates, the (expenditure or sales) weights and type of average being determined by the index formula. One can descend in such a hierarchy as far as available information allows the weights to be decomposed. The lowest level aggregates are called *elementary* aggregates. They are basically of two types:

- Those for which all detailed price and quantity information is available;
- Those for which the statistician, considering the operational cost and the response burden of getting detailed price and quantity information about all the transactions, decides to make use of a representative sample of commodities and respondents.

As indicated above, the practical relevance of studying this topic is large. Since the elementary aggregates form the building blocks of a CPI, the choice of an inappropriate formula at this level can have a tremendous impact on the overall index.

In this section, it will be assumed that detailed price and quantity information for all transactions pertaining to the elementary aggregate for the two time periods under consideration is available. This assumption allows us to define an *ideal elementary aggregate*. Subsequent sections will relax this strong assumption about the availability of detailed price and quantity data on transactions but in any case, it is useful to have a theoretically ideal target for the "practical" elementary index.

The detailed price and quantity data, although perhaps not available to the statistician, is, in principle, available in the outside world. It is frequently the case that at the respondent level (i.e., at the outlet or firm level), some aggregation of the individual transactions information has been executed, usually in a form that suits the respondent's financial or management information system. This respondent determined level of information could be called the *basic information level*. This is, however, not necessarily the finest level of information that could be made available to the price statistician. One could always ask the respondent to provide more disaggregated information. For instance, instead of monthly data one could ask for weekly data; or, whenever appropriate, one could ask for regional instead of global data; or, one could ask for data according to a finer commodity classification. The only natural barrier to further disaggregation is the individual transaction level.[5]

It is now necessary to discuss a problem[6] that arises when detailed data on *individual transactions* are available, either at the level of the individual household or at the level of an individual outlet. Recall that in previous chapters, the price and quantity indexes, $P(p^0,p^1,q^0,q^1)$ and $Q(p^0,p^1,q^0,q^1)$, were introduced. These (bilateral) price and quantity indexes decomposed the value ratio V^1/V^0 into a price change part $P(p^0,p^1,q^0,q^1)$ and a quantity change part $Q(p^0,p^1,q^0,q^1)$. In this framework, it was taken for granted that the period t price and quantity for commodity i, p_i^t and q_i^t respectively, were well defined.[7] However, these definitions are not straightforward since individual consumers may purchase the *same* item during period t *at different prices*. Similarly, if we look at the sales of a particular shop or outlet that sells to consumers*, the same item may sell at very different prices during the course of the period*. Hence before a traditional bilateral price index of the form $P(p^0,p^1,q^0,q^1)$ considered in previous chapters can be applied, there is a non-trivial *time aggregation problem* that must be solved in order to obtain the basic prices p_i^t and q_i^t that are the components of the price vectors p^0 and p^1 and the quantity vectors q^0 and q^1.

---

[5] The material in this section is based on Balk (1994).
[6] This time aggregation problem was discussed briefly in chapter 2.
[7] Note that the period of time t could represent any period of time: a quarter, a month, a week, a day or an hour.

Walsh[8] and Davies (1924) (1932), suggested a solution to this time aggregation problem: the appropriate quantity at this very first stage of aggregation is the *total quantity purchased* of the narrowly defined item and the corresponding price is the value of purchases of this item divided by the total amount purchased, which is a *narrowly defined unit value*. In more recent times, most researchers have adopted the Walsh and Davies solution to the time aggregation problem.[9] Note that this solution to the time aggregation problem has the following advantages:

- The quantity aggregate is intuitively plausible, being the total quantity of the narrowly defined item purchased by the household (or sold by the outlet) during the time period under consideration;
- The product of the price times quantity equals the total value purchased by the household (or sold by the outlet) during the time period under consideration.

We will adopt this solution to the time aggregation problem as our concept for the price and quantity at this preliminary stage of aggregation.

Having decided on an appropriate theoretical definition of price and quantity for an item at the very lowest level of aggregation (i.e., a narrowly defined unit value and the total quantity sold of that item at the individual outlet), we now consider how to aggregate these narrowly defined elementary prices and quantities into an overall elementary aggregate. Suppose that there are N lowest level items or specific commodities in this chosen elementary category. Denote the period t quantity of item n by q_n^t and the corresponding time aggregated unit value price by p_n^t for t = 0,1 and for items n = 1,2,...,N. Define the period t quantity and price vectors as $q^t \equiv [q_1^t,q_2^t,...,q_N^t]$ and $p^t \equiv [p_1^t,p_2^t,...,p_N^t]$ for t = 0,1. It is now necessary to choose a theoretically ideal index number formula $P(p^0,p^1,q^0,q^1)$ that will aggregate the individual item prices into an overall aggregate price relative for the N items in the chosen elementary aggregate. However, this problem of choosing a functional form for $P(p^0,p^1,q^0,q^1)$ is *identical to the overall index number problem* that was addressed in previous chapters. In these previous chapters, four different approaches to index number theory were studied that led to specific index number formulae as being "best" from each perspective. From the viewpoint of *fixed basket approaches*, it was found that the Fisher (1922) and Walsh (1901) price indexes, P_F and P_W, appeared to be "best". From the viewpoint of the *test approach*, the Fisher index appeared to be "best". From the viewpoint of the *stochastic approach* to index number theory, the Törnqvist Theil index number formula P_T emerged as being "best". Finally, from the viewpoint of the *economic approach* to index number theory, the Walsh price index P_W, the Fisher ideal index P_F and the Törnqvist Theil index number formula P_T were all regarded as being equally desirable. It was also shown that the same three index number formulae numerically approximate each other very closely under certain conditions and so it will not matter

[8] Walsh explained his reasoning as follows: "Of all the prices reported of the same kind of article, the average to be drawn is the arithmetic; and the prices should be weighted according to the relative mass quantities that were sold at them." Correa Moylan Walsh (1901; 96). "Some nice questions arise as to whether only what is consumed in the country, or only what is produced in it, or both together are to be counted; and also there are difficulties as to the single price quotation that is to be given at each period to each commodity, since this, too, must be an average. Throughout the country during the period a commodity is not sold at one price, nor even at one wholesale price in its principal market. Various quantities of it are sold at different prices, and the full value is obtained by adding all the sums spent (at the same stage in its advance towards the consumer), and the average price is found by dividing the total sum (or the full value) by the total quantities." Correa Moylan Walsh (1921a; 88).

[9] See for example Szulc (1987; 13), Dalén (1992; 135), Reinsdorf (1994), Diewert (1995; 20-21), Reinsdorf and Moulton (1997) and Balk (2002).

very much which of these alternative indexes is chosen.[10] Hence, the *theoretically ideal elementary index number formula* is taken to be one of the three formulae $P_F(p^0,p^1,q^0,q^1)$, $P_W(p^0,p^1,q^0,q^1)$ or $P_T(p^0,p^1,q^0,q^1)$ where the period t quantity of item n, q_n^t, is the total quantity of that narrowly defined item purchased by the household during period t (or sold by the outlet during period t) and the corresponding price for item n is p_n^t, the time aggregated unit value, for t = 0,1 and for items n = 1,2,...,N.

In the following sections, various "practical" elementary price indexes will be defined. These indexes do not have quantity weights and thus are functions only of the price vectors p^0 and p^1. Thus when a practical elementary index number formula, say $P_E(p^0,p^1)$, is compared to an ideal elementary price index, say the Fisher price index $P_F(p^0,p^1,q^0,q^1)$, then obviously P_E will differ from P_F because the *prices are not weighted according to their economic importance* in the practical elementary formula. It is useful to list the following possible sources of difference between a practical elementary price index $P_E(p^0,p^1)$ and an ideal target index:

- *Weighting bias* or more generally, *formula bias*; i.e., a price index of the form $P_E(p^0,p^1)$ is not able to weight prices according to the economic importance of the product in the consumer's total expenditures on the group of products under consideration.[11]
- *Sampling bias*; i.e., the statistical agency may not be able to collect information on all N products in the elementary aggregate; i.e., only a sample of the N prices may be collected.[12]
- *Time aggregation bias*; i.e., even if a price for a narrowly defined item is collected by the statistical agency, it may not be equal to the theoretically appropriate time aggregated unit value price.[13]
- *Item aggregation bias* or *unit value bias*. The statistical agency may classify certain distinct products as being essentially equivalent and thus the unit value aggregate for this group of aggregated products may not take into account possible significant quality differences in the group of aggregated products. For example, products that are thought to be very similar and are sold in the same units of measurement could be treated as a single product.[14]
- *Aggregation over agents* or *aggregation over entities bias* or *aggregation over outlets bias*. The unit value for a particular item may be constructed by aggregating over all households in a region or a certain demographic class or by aggregating over all outlets or shops that sell the item in a particular region.[15]

[10] Theorem 5 in Diewert (1978; 888) showed that P_F, P_T and P_W will approximate each other to the second order around an equal price and quantity point. However, if there are violent fluctuations in prices and quantities, a second order approximation to any one of these formulae may not be very accurate.
[11] For materials on how to measure formula bias, see Diewert (1998), White (1999) (2000) and chapter 7.
[12] This is a specialized topic with a long history. It will not be covered in this volume.
[13] Many statistical agencies send price collectors to various outlets on certain days of the month to collect list prices of individual items. Usually, price collectors do not work on weekends when many sales take place and thus the collected prices may not be fully representative of all transactions that occur. Thus these collected prices can be regarded only as approximations to the time aggregated unit values for those items.
[14] For materials on unit value bias, see Diewert and von der Lippe (2010) and Silver (2010) (2011) and the additional references in these papers.
[15] For materials on possible methods to measure outlet substitution bias, see Diewert (1998). The problems associated with measuring aggregation over consumers' bias were noted in the final sections of chapter 5.

- *New and disappearing products bias*; i.e., $P_E(p^0,p^1)$ measures price change only over matched products for the two periods being compared; new products and disappearing products are ignored in standard elementary indexes that depend only on prices.[16]

Approximations to the numerical differences between various elementary indexes of the form $P_E(p^0,p^1)$ and various superlative indexes will be developed in chapter 7.

In the following section, the problems of aggregation and classification will be discussed in more detail.

3. Aggregation and Classification Problems for Elementary Aggregates

Hawkes and Piotrowski (2003) noted that the definition of an elementary aggregate involves aggregation over *four* possible dimensions:[17]

- A *time* dimension; i.e., the item unit value could be calculated for all item transactions for a year, a month, a week, or a day.
- A *spatial* dimension; i.e., the item unit value could be calculated for all item transactions in the country, province, state, city, neighbourhood or individual location.
- A *product* dimension; i.e., the item unit value could be calculated for all item transactions in a broad general category (e.g., food), in a more specific category (e.g., margarine), for a particular brand (ignoring package size) or for a particular narrowly defined item (e.g., a particular AC Nielsen universal product code).
- A *sectoral* (or *entity* or *economic agent*) dimension; i.e., the item unit value could be calculated for a particular class of households or a particular class of outlets.

Each of the above dimensions for choosing the domain of definition for an elementary aggregate will be discussed in turn.

As the time period is compressed, several problems emerge:

- Purchases (by households) and sales (by outlets) become erratic and sporadic. Thus the frequency of unmatched purchases or sales from one period to the next increases and in the limit (choose the time period to be one minute), nothing will be matched and bilateral index number theory fails at the individual consumer level.[18]
- As the time period becomes shorter, chained indexes exhibit more "drift"; i.e., if the data at the end of a chain of periods reverts to the data in the initial period, the chained index does not revert back to unity. As was discussed in section 8 of chapter 2, it is only

[16] This problem was addressed in section 14 of chapter 5. It will be addressed in more detail in chapters 7 and 8.

[17] Hawkes and Piotrowski (2003; 31) combined the spatial and sectoral dimensions into the spatial dimension. They also acknowledged the pioneering work of Theil (1954), who identified three dimensions of aggregation: aggregation over individuals, aggregation over commodities and aggregation over time. It should be noted that William Hawkes was a pioneer in realizing the importance of scanner data for the construction of Consumer Price Indexes; see Hawkes (1997). Other important contributors include Reinsdorf (1996), Silver (1995), Silver and Heravi (2001) (2003) (2005), de Haan and van der Grient (2011), Ivancic, Diewert and Fox (2011) and de Haan and Krsinich (2014).

[18] This problem was noted in section 19 of chapter 5. David Richardson (2003; 51) also made this point: "Defining items with a finer granularity, as is the case if quotes in different weeks are treated as separate items, results in more missing data and more imputations." However, high frequency consumer price indexes could be successfully constructed if aggregation over households or outlets is permitted.

appropriate to use chained indexes when the underlying price and quantity data exhibit relatively smooth trends. When the time period is short, seasonal fluctuations[19] and periodic sales and advertising campaigns[20] can cause prices and quantities to oscillate (or "bounce" to use Szulc's (1983; 548) term) and hence it is not appropriate to use chained indexes under these circumstances. If fixed base indexes are used in this short time period situation, then the results will usually depend very strongly on the choice of the base period. In the seasonal context, not all commodities may even be in the marketplace during the chosen base period.[21] All of these problems can be mitigated by choosing a longer time period so that trends in the data will tend to dominate the short term fluctuations.

- As the time period contracts, virtually all goods become *durable* in the sense that they yield services not only for the period of purchase but for subsequent periods. Thus the period of purchase or acquisition becomes different from the periods of use, leading to many complications.[22]
- As the time period contracts, users will usually not be particularly interested in the short term fluctuations of the resulting index and there will be demands for smoothing the necessarily erratic results. Put another way, users will desire a way of summarizing the weekly or daily movements in the index into monthly or quarterly movements in prices. Hence from the viewpoint of meeting the needs of users, there may be relatively little demand for high frequency indexes.

In view of the above considerations, it is recommended that the index number time period be at least four consecutive weeks or a month.[23]

It is also necessary to choose the spatial dimension of the elementary aggregate. Should item prices in each city or region be considered as separate aggregates or should a national item aggregate be constructed? Obviously, if it is desired to have regional CPIs that aggregate up to a national CPI, then it will be necessary to collect item prices by region. However, it is not clear how fine the "regions" should be. It could be as fine as a grouping of households in a postal code or to individual outlets across the country.[24] There does not seem to be a clear consensus on what the optimal degree of spatial disaggregation should be.[25] Each statistical agency will have to make

---

[19] See chapter 9 for a monthly seasonal example where chained month to month indexes are a disaster.

[20] See Feenstra and Shapiro (2003) for an example of a weekly superlative index that exhibits massive chain drift. Substantial chain drift can also occur using monthly indexes; see Szulc (1983) (1987). See Richardson (2003; 50-51) and Ivancic, Diewert and Fox (2011) for additional discussions of the issues involved in choosing weekly unit values versus monthly unit values.

[21] See chapter 9 below for suggested solutions to these seasonality problems.

[22] See chapter 10 below for more material on the possible CPI treatment of durable goods.

[23] If there is very high inflation in the economy (or even hyperinflation), then it may be necessary to move to weekly or even daily indexes. Also, it should be noted that some index number theorists feel that new theories of consumer behavior should be developed that could utilize weekly or daily data: "Some studies have endorsed unit values to reduce high frequency price variation, but this implicitly assumes that the high frequency variation represents simply noise in the data and is not meaningful in the context of a COLI. That is debatable. We need to develop a theory that confronts the data, not truncate the data to fit the theory." Jack E. Triplett (2003; 153). However, until such new theories are adequately developed, it seems pragmatic to define the item unit values over months or quarters rather than days or weeks.

[24] Iceland no longer uses regional weights but uses individual outlets as the primary geographical unit; see Gudnason (2003; 18).

[25] Hawkes and Piotrowski note that it is quite acceptable to use national elementary aggregates when making international comparisons between countries: "When we try to compare egg prices across geography, however, we find that lacing across outlets won't work, because the eyelets on one side of the shoe (or outlets on one side of the river) don't match up with those on the other side. Thus, in making

its own judgements on this matter, taking into account the costs of data collection and the demands of users for a spatial dimension for the CPI.

How detailed should the product dimension be? The possibilities range from regarding all commodities in a general category as being equivalent to the other extreme, where only a commodity in a particular package size made by a particular manufacturer or service provider is regarded as being equivalent. All things being equal, Triplett (2004) stressed the advantages of matching products at the most detailed level possible, since this will prevent quality differences from clouding the period to period price comparisons. This is sensible advice but then what are the drawbacks to working with the finest possible commodity classification? The major drawback is that the finer the classification is, the more difficult it will be to *match* the item purchased or sold in the base period to the same item in the current period. Hence, the finer the product classification, the smaller will be the number of matched price comparisons that are possible.[26] This would not be a problem if the unmatched prices followed the same trend as the matched ones in a particular elementary aggregate, but in at least some circumstances, this will not be the case.[27] Thus the finer the classification system is, the more work (in principle) there will be for the statistical agency to quality adjust or impute the prices that do not match. Choosing a relatively coarse classification system can lead to a very cost efficient system of quality adjustment (i.e., essentially no explicit quality adjustment or imputation is done for the prices that do not exactly match) but it may not be very accurate. The statistical agency will have to balance the theoretical purity of a very fine classification system with the possible loss of product matches.

The final issue in choosing a classification scheme is the issue of choosing a sectoral dimension; i.e., should the unit value for a particular item be calculated for a particular outlet or a particular household or for a class of outlets or households? Before this question can be answered, it is necessary to ask whether the individual outlet or the individual household is the appropriate finest level of entity classification. If the economic approach to the consumer price index is taken, then the individual household is the appropriate finest level of entity classification.[28] However, if the time period is short, a single household will not work very well as the basic unit of entity observation due to the sporadic nature of many purchases by an individual household; i.e., there will be tremendous difficulties in matching prices across periods for individual households.

interspatial comparisons, we have no choice but to aggregate outlets all the way up to the regional (or, in the case of purchasing power parities, national) level. We have no hesitation about doing this for interspatial comparisons, but we are reluctant to do so for intertemporal ones. Why is this?" William J. Hawkes and Frank W. Piotrowski (2003; 31-32). An answer to their question is that it is preferable to match like with like as closely as possible which leads statisticians to prefer the finest possible level of aggregation, which, in the case of intertemporal comparisons, would be the individual household or the individual outlet. However, in making cross region comparisons, matching is not possible unless regional item aggregates are formed, as Hawkes and Piotrowski point out above.

[26] This is part of the *matching problem* discussed at the end of chapter 5.

[27] Silver and Heravi (2001) (2003; 286) (2005) and Koskimäki and Vartia (2001) stressed this point and presented empirical evidence to back up their point. Feenstra (1994) and Balk (2000) used the assumption of CES preferences to deal with the new products problem. Their approaches will be discussed in chapter 8.

[28] This point has been made emphatically by two authors in a book on scanner data and price indexes: "In any case, unit values across stores are not the prices actually faced by households and do not represent the per period price in the COLI, even if the unit values are grouped by type of retail outlet." Jack E. Triplett (2003; 153-154). "Furthermore, note that the relationship being estimated is not a proper consumer demand function but rather an 'establishment sales function'. Only after making further assumptions – for example, fixing the distribution of consumers across establishments – is it permissible to jump to demand functions". Eduardo Ley (2003; 380).

However, for a grouping of "similar" households that is sufficiently large, it does become feasible in theory to use the grouped household as the entity classification rather than the outlet as is usually done. This is not usually done because of the costs and difficulties involved in collecting individual household data on prices and expenditures.[29] Thus price information is usually collected from retail establishments or outlets that sell mainly to households. Matching problems are mitigated using this strategy (but not eliminated) because the retail outlet generally sells the same items on a continuing basis.

If expenditures by all households in a region are aggregated together, will they equal sales by the retail outlets in the region? Under certain conditions, the answer to this question is yes. The conditions are that the outlets do not sell any items to purchasers who are not local households (no regional exports or sales to local businesses or governments) and that the regional households do not make any purchases of consumption items other than from the local outlets (no household imports or transfers of commodities to local households by governments). Obviously, these restrictive conditions will not be met in practice but they may hold as a first approximation.

The effects of *regional aggregation* and *product aggregation* can be examined, thanks to a study by Koskimäki and Ylä-Jarkko (2003). This study utilized scanner data for the last week in September 1998 and September 2000 on butter, margarine and other vegetable fats, vegetable oils, soft drinks, fruit juices and detergents. This information was provided by the AC Nielsen company for Finland. At the finest level of item classification (the AC Nielsen Universal Product Code), the number of individual items in the sample was 1028. The total number of outlets in the sample was 338. Koskimäki and Ylä-Jarkko considered four levels of spatial disaggregation:

- The entire country (1 level);
- Provinces (4 levels);
- AC Nielsen regions (15 levels);
- Individual outlets (338 levels).

They also considered four levels of product disaggregation:

- The COICOP 5 digit classification (6 levels);
- The COICOP 7 digit classification (26 levels);
- The AC Nielsen brand classification (266 levels);
- The AC Nielsen individual Universal Product Code (1028 distinct products).

In order to illustrate the ability to match products over the two year period as a function of the degree of fineness of the classification, Koskimäki and Ylä-Jarkko (2003; 10) presented a table that shows that the proportion of transactions that could be matched across the two years fell steadily as the fineness of the classification scheme increased. At the highest level of aggregation (the national and COICOP 5 digit), all transactions could be matched over the two year period but at the finest level of aggregation (338 outlets times 1028 individual products or 347,464 classification cells in all), only 61.7 % of the value of transactions in 2000 could be matched back to their 1998 counterparts. Their Table 7 is reproduced as Table 1 below.

[29] However, it is possible to collect accurate household data in certain circumstances; see Gudnason (2003), who pioneered a receipts methodology for collecting household price and expenditure data in Iceland. Also, in the future, as monetary transactions are replaced by debit and credit card transactions, it will become possible to construct individual household estimates of real consumption, provided that product codes are included in the transaction records.

Table 1: Proportion of Transactions in 2000 that Could be Matched to 1998

	COICOP 5 digit	COICOP 7 digit	AC Nielsen Brand	AC Nielsen UPC
Country	1.000	1.000	.982	.801
Province	1.000	1.000	.975	.774
AC Nielsen Region	1.000	1.000	.969	.755
Individual Outlet	.904	.904	.846	.617

For each of the above 16 levels of product and regional disaggregation, for the products that were available in September of 1998 and 2000, Koskimäki and Ylä-Jarkko (2003; 9) calculated Laspeyres and Fisher price indexes. They found substantial differences in these indexes as the degree of disaggregation increased.

Another study on the effects of alternative methods of unit value aggregation over outlets (i.e., treat each unit value for each product in each store as a unique product versus aggregating products over stores and chains) was undertaken by Ivancic and Fox (2013). They used 65 weeks of scanner data on the sales of different types of instant coffee sold by four supermarket chains in Australia in 110 stores where the data were collected between February 1997 and April 1998. It contains information on 110 stores which belong to four supermarket chains located in the metropolitan area of one of the major capital cites in Australia. These stores accounted for over 80 percent of grocery sales in the various capital cities of Australia during this period. After data exclusions, 436,103 weekly observations on 157 coffee items were used in their study.[30] Their results on alternative methods of aggregation can be summarized as follows:

"The results show that when non-superlative index numbers are used to calculate price change, aggregation choices can have a huge impact. However, the issue of aggregation seems to become relatively trivial when the standard Fisher and Törnqvist superlative indexes are used, with an extremely close range of estimates of price change found across different aggregation methods. This result seems to provide further support for the use of these superlative indexes over the use of non-superlative indexes to estimate price change." Lorraine Ivancic and Kevin J. Fox (2013; 643).

The non-superlative index numbers[31] were chained Laspeyres and Paasche indexes and the superlative indexes were chained Fisher and Törnqvist Theil indexes.

Thus the problem of determining the "best" unit value to insert into an index number formula is far from settled. We will look at this problem again in Chapter 11.

Another issue that arises in the context of defining exactly what prices and quantities should be entered into an index number formula is the following one: should statistical agencies exclude sale prices? In general, this is not a recommended practice since very large amounts of a product can be sold at a sale price. Fox and Syed (2016; 404) found that the exclusion of sale prices can introduce a substantial bias. They also found that even when sale prices are included they are systematically under-weighted, but the under-weighting remains fairly stable over time so that inflation measurement is not significantly affected. They also found evidence that the typical practice of using data from an incomplete period in constructing unit values can lead to an upward bias in the resulting price index.[32]

[30] Their paper also lists some related studies.
[31] The weekly unit values by product were aggregated into monthly unit values.
[32] Diewert, Fox and de Haan (2016) also found this effect. The direction of this bias may be due to an increasing frequency of end of month or quarter sales.

4. Some Elementary Indexes that Have Been Suggested Over the Years

Suppose that there are N commodities in a chosen elementary category. Denote the period t price of item n by p_n^t for t = 0,1 and for items n = 1,2,...,N. As usual, define the period t price vector as $p^t \equiv [p_1^t,p_2^t,...,p_N^t]$ for t = 0,1.

The first simple elementary index number formula is due to the French economist Dutot (1738):

(1) $P_D(p^0,p^1) \equiv [\sum_{n=1}^N (1/N)\, p_n^1]/[\sum_{n=1}^N (1/N)\, p_n^0] = [\sum_{n=1}^N p_n^1]/[\sum_{n=1}^N p_n^0] = p^1\cdot 1_N/p^0\cdot 1_N.$

Thus the Dutot elementary price index is equal to the arithmetic average of the N period 1 prices divided by the arithmetic average of the N period 0 prices.

The second simple elementary index number formula is due to the Italian economist Carli (1764):

(2) $P_C(p^0,p^1) \equiv \sum_{n=1}^N (1/N)(p_n^1/p_n^0).$

Thus the Carli elementary price index is equal to the *arithmetic* average of the N item price ratios or price relatives, p_n^1/p_n^0. This formula was already encountered in our study of the unweighted stochastic approach to index numbers; recall definition (2) in chapter 4 above.

The third simple elementary index number formula is due to the English economist Jevons (1865):

(3) $P_J(p^0,p^1) \equiv \prod_{n=1}^N (p_n^1/p_n^0)^{1/N}.$

Thus the Jevons elementary price index is equal to the *geometric* average of the N item price ratios or price relatives, p_n^1/p_n^0. Again, this formula was introduced as formula (4) in our discussion of the unweighted stochastic approach to index number theory in chapter 4 above.

The fourth elementary index number formula P_H is the *harmonic* average of the N item price relatives and it was first suggested in passing as an index number formula by Jevons (1865; 121) and Coggeshall (1887):

(4) $P_H(p^0,p^1) \equiv [\sum_{n=1}^N (1/N)(p_n^1/p_n^0)^{-1}]^{-1}.$

Finally, the fifth elementary index number formula is the geometric average of the Carli and harmonic formulae; i.e., it is *the geometric mean of the arithmetic and harmonic means* of the N price relatives:

(5) $P_{CSWD}(p^0,p^1) \equiv [P_C(p^0,p^1)\, P_H(p^0,p^1)]^{1/2}.$

This index number formula was first suggested by Fisher (1922; 472) as his formula 101. Fisher also observed that, empirically for his data set, P_{CSWD} was very close to the Jevons index, P_J, and these two indexes were his "best' unweighted index number formulae. In more recent times, Carruthers, Sellwood and Ward (1980; 25) and Dalén (1992; 140) also proposed P_{CSWD} as an elementary index number formula.

It should be noted that the Jevons index is now the most commonly used elementary index (when only price information is available). The Dutot and Carli formulae are used by a few statistical agencies.

Having defined the most commonly used elementary formulae, the question now arises: which formula is "best"? Obviously, this question cannot be answered until desirable properties for elementary indexes are developed. This will be done in a systematic manner in section 6 below (using the test approach), but in the present section, one desirable property for an elementary index will be noted. This is the *time reversal test*, which was noted earlier in chapters 2 and 3. In the present context, this test for the elementary index $P(p^0,p^1)$ becomes:

(6) $P(p^0,p^1)P(p^1,p^0) = 1.$

This test says that if the prices in period 2 revert to the initial prices of period 0, then the product of the price change going from period 0 to 1, $P(p^0,p^1)$, times the price change going from period 1 to 2, $P(p^1,p^0)$, should equal unity; i.e., under the stated conditions, we should end up where we started.[33] It can be verified that the Dutot, Jevons and Carruthers, Sellwood, Ward and Dalén indexes, P_D, P_J and P_{CSWD}, all satisfy the time reversal test but that the Carli and Harmonic indexes, P_C and P_H, fail this test. In fact, these last two indexes fail the test in the following *biased* manner:

(7) $P_C(p^0,p^1)\ P_C(p^1,p^0) \geq 1$;
(8) $P_H(p^0,p^1)\ P_H(p^1,p^0) \leq 1$

with strict inequalities holding in (7) and (8) provided that the period 1 price vector p^1 is not proportional to the period 0 price vector p^0.[34] Thus the Carli index will generally have an *upward bias* while the Harmonic index will generally have a *downward bias*. Fisher (1922; 66 and 383) was quite definite in his condemnation of the Carli index due to its upward bias.[35] Because it fails the time reversal test, the Carli index is not used in compiling elementary price indexes for the Harmonized Index of Consumer Prices (HICP) which is the official Eurostat index used to compare consumer prices across European Union countries.

In the following section, some numerical relationships between the five elementary indexes defined in this section will be established. Then in the subsequent section, a more comprehensive list of desirable properties for elementary indexes will be developed and the five elementary formulae will be evaluated in the light of these properties or tests.

5. Numerical Relationships between Some Elementary Indexes

It can be shown[36] that the Carli, Jevons and Harmonic elementary price indexes satisfy the following inequalities:

---

[33] This test can also be viewed as a special case of Walsh's (1901) Multiperiod Identity Test, (63) in chapter 2.

[34] These inequalities follow from the fact that a harmonic mean of N positive numbers is always equal to or less than the corresponding arithmetic mean; see Walsh (1901;517) or Fisher (1922; 383-384). This inequality is a special case of Schlömilch's Inequality; see Hardy, Littlewood and Pólya (1934; 26).

[35] See also Szulc (1987; 12) and Dalén (1992; 139). Dalén (1994; 150-151) provided some nice intuitive explanations for the upward bias of the Carli index.

[36] Each of the three indexes P_H, P_J and P_C is a mean of order r where r equals −1, 0 and 1 respectively and so the inequalities follow from Schlömilch's inequality.

(9) $P_H(p^0,p^1) \le P_J(p^0,p^1) \le P_C(p^0,p^1)$;

i.e., the Harmonic index is always equal to or less than the Jevons index, which in turn, is always equal to or less than the Carli index. In fact, the strict inequalities in (9) will hold provided that the period 0 vector of prices, p^0, is not proportional to the period 1 vector of prices, p^1.

The inequalities (9) do not tell us by how much the Carli index will exceed the Jevons index and by how much the Jevons index will exceed the Harmonic index. Hence, in the remainder of this section, some approximate relationships between the five indexes defined in the previous section will be developed that will provide some practical guidance on the relative magnitudes of each of the indexes.

The first approximate relationship that will be derived is between the Jevons index P_J and the Dutot index P_D. For each period t, define the *arithmetic mean* of the N prices pertaining to that period as follows:

(10) $p^{t*} \equiv \sum_{n=1}^N (1/N)\, p_n^t$; $\qquad t = 0,1.$

Now define (implicitly) the *multiplicative deviation* of the nth price in period t relative to the mean price in that period, e_n^t, as follows:

(11) $p_n^t = p^{t*}(1+e_n^t)$; $\qquad n = 1,...,N ; t = 0,1.$

Note that (10) and (11) imply that the deviations e_n^t sum to zero in each period; i.e., we have:

(12) $\sum_{n=1}^N e_n^t = 0$; $\qquad t = 0,1.$

Note that the Dutot index can be written as the ratio of the mean prices, p^{1*}/p^{0*}; i.e., we have:

(13) $P_D(p^0,p^1) = p^{1*}/p^{0*}$.

Now substitute equations (11) into the definition of the Jevons index, (3):

(14) $P_J(p^0,p^1) = \prod_{n=1}^N [p^{1*}(1+e_n^1)/p^{0*}(1+e_n^0)]^{1/N}$
$= [p^{1*}/p^{0*}] \prod_{n=1}^N [(1+e_n^1)/(1+e_n^0)]^{1/N}$
$= P_D(p^0,p^1)\, f(e^0,e^1)$ $\qquad$ using definition (1)

where $e^t \equiv [e_1^t,...,e_N^t]$ for t = 0 and 1, and the function f is defined as follows:

(15) $f(e^0,e^1) \equiv \prod_{N=1}^N [(1+e_n^1)/(1+e_n^0)]^{1/N}$.

Expand $f(e^0,e^1)$ by a second order Taylor series approximation around $e^0 = 0_N$ and $e^1 = 0_N$. Using (12), it can be verified[37] that we obtain the following second order approximate relationship between P_J and P_D:

(16) $P_J(p^0,p^1) \approx P_D(p^0,p^1)[1 + (1/2N)e^0 \cdot e^0 - (1/2N)e^1 \cdot e^1]$
$= P_D(p^0,p^1)[1 + (1/2)\mathrm{var}(e^0) - (1/2)\mathrm{var}(e^1)]$

[37] This approximate relationship was first obtained by Carruthers, Sellwood and Ward (1980; 25).

where $\mathrm{var}(e^t)$ is the variance of the period t multiplicative deviations; i.e., for t = 0,1:

$$(17)\ \mathrm{var}(e^t) \equiv (1/N)\sum_{n=1}^{N} (e_n^t - e^{t*})^2$$
$$= (1/N)\sum_{n=1}^{N} (e_n^t)^2 \qquad \text{since } e^{t*} = 0 \text{ using (12)}$$
$$= (1/N)e^t \cdot e^t .$$

Under normal conditions[38], the variance of the deviations of the prices from their means in each period is likely to be approximately constant and so under these conditions, the Jevons price index will approximate the Dutot price index to the second order.

Note that with the exception of the Dutot formula, the remaining four elementary indexes defined in section 4 are functions of the relative prices of the N items being aggregated.[39] This fact is used in order to derive some approximate relationships between these four elementary indexes. Thus define the *nth price relative* as

$$(18)\ r_n \equiv p_n^1/p_n^0 ; \qquad n = 1,...,N.$$

Define the *arithmetic mean of the n price relatives* as

$$(19)\ r^* \equiv (1/N)\sum_{n=1}^{N} r_n = P_C(p^0,p^1)$$

where the last equality follows from the definition (2) for the Carli index. Finally, define (implicitly) the *deviation* e_n of the nth price relative r_n from the arithmetic average of the N price relatives r^* as follows:

$$(20)\ r_n = r^*(1+e_n) ; \qquad n = 1,...,N.$$

Note that (19) and (20) imply that the deviations e_n sum to zero; i.e., we have:

$$(21)\ \sum_{n=1}^{N} e_n = 0.$$

[38] If there are significant changes in the overall inflation rate, some studies indicate that the variance of deviations of prices from their means can also change. Also if N is small, then there will be sampling fluctuations in the variances of the prices from period to period, leading to random differences between the Dutot and Jevons indexes. If prices are normalized to equal 1 in period 0, this amounts to choosing particular units of measurement for the N commodities. In this case, $\mathrm{var}(e^0) = 0$, and the approximation (16) becomes the inequality $P_J(p^0,p^1) < P_D(p^0,p^1)$ if $\mathrm{var}(e^1) > 0$. In this case where normalized prices are used, the Dutot index becomes a Carli index which has an upward bias relative to the Jevons index. Appendix A shows that this bias can be substantial.

[39] The Dutot index can be rewritten as a function of relative prices and shares that depend only on period 0 prices as follows: $P_D(p^0,p^1) = \Sigma_{n=1}^{N} \sigma_n^0(p_n^1/p_n^0)$ where $\sigma_n^0 \equiv p_n^0/\Sigma_{i=1}^{N} p_i^0$ for n = 1,...,N; see the IMF, ILO, Eurostat, UNECE, OECD and World Bank (2020; 180). This publication also notes the following problem with the use of the Dutot formula: "Even when the varieties are fairly homogeneous and measured in the same units, the Dutot's implicit weights may still not be satisfactory. More weight is given to the price changes for the more expensive varieties, but in practice, they may well account for only small shares of the total expenditure within the aggregate. Consumers are unlikely to buy varieties at high prices if the same varieties are available at lower prices." IMF, ILO, Eurostat, UNECE, OECD and World Bank (2020; 180-181).

Now substitute equations (20) into the definitions of P_C, P_J, P_H and P_{CSWD}, (2)-(5) above, in order to obtain the following representations for these indexes in terms of the vector of deviations, $e \equiv [e_1,...,e_N]$:[40]

$$(22)\ P_C(p^0,p^1) = \sum_{n=1}^N (1/N)r_n = r^* \equiv r^* f_C(e) ;$$
$$(23)\ P_J(p^0,p^1) = \prod_{n=1}^N r_n^{1/N} = r^* \prod_{n=1}^N (1+e_n)^{1/N} \equiv r^* f_J(e) ;$$
$$(24)\ P_H(p^0,p^1) = [\sum_{n=1}^N (1/N)(r_n)^{-1}]^{-1} = r^*[\sum_{n=1}^N (1/N)(1+e_n)^{-1}]^{-1} \equiv r^* f_H(e) ;$$
$$(25)\ P_{CSWD}(p^0,p^1) = [P_C(p^0,p^1)P_H(p^0,p^1)]^{1/2} = r^*[f_C(e)f_H(e)]^{1/2} \equiv r^* f_{CSWD}(e)$$

where the last equation in (22)-(25) serves to define the deviation functions, $f_C(e)$, $f_J(e)$, $f_H(e)$ and $f_{CSWD}(e)$. The second order Taylor series approximations to each of these functions[41] around the point $e = 0_N$ are:

$$(26)\ f_C(e) \approx 1 ;$$
$$(27)\ f_J(e) \approx 1 - (1/2N)e\cdot e = 1 - (1/2)\mathrm{var}(e) ;$$
$$(28)\ f_H(e) \approx 1 - (1/N)e\cdot e = 1 - \mathrm{var}(e) ;$$
$$(29)\ f_{CSWD}(e) \approx 1 - (1/2N)e\cdot e = 1 - (1/2)\mathrm{var}(e)$$

where we have made repeated use of (21) in deriving the above approximations.[42] Thus to the second order, the Carli index P_C will *exceed* the Jevons and Carruthers Sellwood Ward Dalén indexes, P_J and P_{CSWD}, by $(1/2)r^*\mathrm{var}(e)$, which is r^* times one half the variance of the N price relatives p_n^1/p_n^0. Similarly, to the second order, the Harmonic index P_H will *lie below* the Jevons and Carruthers Sellwood Ward Dalén indexes, P_J and P_{CSWD}, by r^* times one half the variance of the N price relatives p_n^1/p_n^0.

Thus empirically, it is expected that the Jevons and Carruthers Sellwood Ward and Dalén indexes will be very close to each other.[43] Using the previous approximation result (16), it is expected that the Dutot index P_D will also be fairly close to P_J and P_{CSWD}, with some fluctuations over time due to changing variances of the period 0 and 1 deviation vectors, e^0 and e^1. Thus it is expected that these three elementary indexes will give much the same numerical answers in empirical applications. On the other hand, the Carli index can be expected to be substantially *above* these three indexes, with the degree of divergence growing as the variance of the N price relatives grows. Similarly, the Harmonic index can be expected to be substantially *below* the three middle indexes, with the degree of divergence growing as the variance of the N price relatives grows.

6. The Test Approach to Elementary Indexes

Recall that in chapter 3, the axiomatic approach to bilateral price indexes $P(p^0,p^1,q^0,q^1)$ was developed. In the present section, the elementary price index $P(p^0,p^1)$ depends only on the period 0 and 1 price vectors, p^0 and p^1 respectively, so that the elementary price index does not depend on the period 0 and 1 quantity vectors, q^0 and q^1. One approach to obtaining new tests or axioms

[40] Note that the vector of deviations e defined by equations (20) is different from the deviation vectors e^0 and e^1 defined by equations (11).

[41] From (22), it can be seen that $f_C(e)$ is identically equal to 1 so that (26) will be an exact equality rather than an approximation.

[42] These second order approximations are due to Dalén (1992; 143) for the case $r^* = 1$ and to Diewert (1995; 29) for the case of a general r^*.

[43] Reinsdorf and Triplett (2009; 63) noted that for the case N = 2, $P_{CSWD}(p^0,p^1) = P_J(p^0,p^1)$. This paper and Diewert (1993) provide a review of early approaches to index number theory and the construction of a Consumer Price index.

for an elementary index is to look at the twenty or so axioms that were listed in Chapter 3 for bilateral price indexes $P(p^0,p^1,q^0,q^1)$ and adapt those axioms to the present context; i.e., use the old bilateral tests for $P(p^0,p^1,q^0,q^1)$ that do not depend on the quantity vectors q^0 and q^1 as tests for an elementary index $P(p^0,p^1)$.[44] This approach will be utilized in the present section.

The first eight tests or axioms are reasonably straightforward and uncontroversial:

T1: *Continuity*: $P(p^0,p^1)$ is a continuous function of the N positive period 0 prices $p^0 \equiv [p_1{}^0,...,p_N{}^0]$ and the N positive period 1 prices $p^1 \equiv [p_1{}^1,...,p_N{}^1]$.

T2: *Identity*: $P(p,p) = 1$; i.e., the period 0 price vector equals the period 1 price vector, then the index is equal to unity.

T3: *Monotonicity in Current Period Prices*: $P(p^0,p^1) < P(p^0,p)$ if $p^1 < p$; i.e., if any period 1 price increases, then the price index increases.

T4: *Monotonicity in Base Period Prices*: $P(p^0,p^1) > P(p,p^1)$ if $p^0 < p$; i.e., if any period 0 price increases, then the price index decreases.

T5: *Proportionality in Current Period Prices*: $P(p^0,\lambda p^1) = \lambda P(p^0,p^1)$ if $\lambda > 0$; i.e., if all period 1 prices are multiplied by the positive number λ, then the initial price index is also multiplied by λ.

T6: *Inverse Proportionality in Base Period Prices*: $P(\lambda p^0,p^1) = \lambda^{-1} P(p^0,p^1)$ if $\lambda > 0$; i.e., if all period 0 prices are multiplied by the positive number λ, then the initial price index is multiplied by $1/\lambda$.

T7: *Mean Value Test*: $\min_n\{p_n{}^1/p_n{}^0 : n = 1,...,N\} \leq P(p^0,p^1) \leq \max_n\{p_n{}^1/p_n{}^0 : n = 1,...,N\}$; i.e., the price index lies between the smallest and largest price relative.

T8: *Symmetric Treatment of Outlets*: $P(p^0,p^1) = P(p^{0*},p^{1*})$ where p^{0*} and p^{1*} denote the *same* permutation of the components of p^0 and p^1; i.e., if we change the ordering of the outlets (or households) from which we obtain the price quotations for the two periods, then the elementary index remains unchanged.

Eichhorn (1978; 155) showed that Tests 1, 2, 3 and 5 imply Test 7, so that not all of the above tests are logically independent.

The following tests are more controversial and are not necessarily accepted by all price statisticians.

T9: *The Price Permutation Test*: $P(p^0,p^1) = P(p^{0*},p^{1**})$ where p^{0*} and p^{1**} denote possibly *different* permutations of the components of p^0 and p^1; i.e., if the ordering of the price quotes for both periods is changed in possibly different ways, then the elementary index remains unchanged.

Obviously, T8 is a special case of T9 where the two permutations of the initial ordering of the prices are restricted to be the same. Thus T9 implies T8. Test T9 is due to Dalén (1992; 138). He justified this test by suggesting that the price index should remain unchanged if outlet prices

[44] This was the approach used by Diewert (1995; 5-17), who drew on the earlier work of Eichhorn (1978; 152-160) and Dalén (1992).

"bounce" in such a manner that the outlets are just exchanging prices with each other over the two periods. While this test has some intuitive appeal, it is not consistent with the idea that the price of a specific product in a specific outlet (which may have some special characteristics which are not present in other outlets) should be matched with the same product price in the same outlet in a one to one manner across the two periods.[45]

The following test was also proposed by Dalén (1992) in the elementary index context:

T10: *Time Reversal*: $P(p^1,p^0) = 1/P(p^0,p^1)$; i.e., if the data for periods 0 and 1 are interchanged, then the resulting price index should equal the reciprocal of the original price index.

It is difficult to accept an index that gives a different answer if the ordering of time is reversed.

T11: *Circularity*: $P(p^0,p^1)P(p^1,p^2) = P(p^0,p^2)$; i.e., the price index going from period 0 to 1 times the price index going from period 1 to 2 equals the price index going from period 0 to 2 directly.

The circularity and identity tests imply the time reversal test; (just set $p^2 = p^0$). The circularity property would seem to be a very desirable property: it is a generalization of a property that holds for a single price relative.

Elementary price indexes may be calculated as direct price indexes by comparing the prices of the current period with those of a fixed price reference period or as chained short-term indexes obtained by multiplying the monthly (or quarterly) price indexes into a long-term price index. Many statistical offices chose to calculate the elementary price indexes by chaining the short term (monthly or quarterly) indexes because this has some practical advantages when dealing with replacements in the sample. For elementary indexes calculated as chained short-term price indexes it is crucial that the index meets the circularity test.

T12: *Commensurability*: $P(\lambda_1 p_1^0,..., \lambda_N p_N^0; \lambda_1 p_1^1,..., \lambda_N p_N^1) = P(p_1^0,...,p_N^0; p_1^1,...,p_N^1) = P(p^0,p^1)$ for all $\lambda_1 > 0, ..., \lambda_N > 0$; i.e., if we change the units of measurement for each commodity in each outlet, then the elementary index remains unchanged.

In the bilateral index context, virtually every price statistician accepts the validity of this test. However, in the elementary context, this test is more controversial. If the N items in the elementary aggregate are all very homogeneous, then it makes sense to measure all of the items in the same units. Hence, if we change the unit of measurement in this homogeneous case, then test T12 should restrict each of the λ_n to be the same number (say λ) and test T12 becomes the following test:

(30) $P(\lambda p^0,\lambda p^1) = P(p^0,p^1)$ for all $p^0 >> 0_N$, $p^1 >> 0_N$ and $\lambda > 0$.

Note that (30) will be satisfied if tests T5 and T6 are satisfied.

[45] Since a typical official Consumer Price Index consists of approximately 600 to 1000 separate strata where an elementary index needs to be constructed for each stratum, it can be seen that many strata will consist of quite heterogeneous items. Thus for a fruit category, some of the N items whose prices are used in the elementary index will correspond to quite different types of fruit with quite different prices. Randomly permuting these prices in periods 0 and 1 will lead to very odd price relatives in many cases, which may cause the overall index to behave badly unless the Jevons or Dutot formula is used.

However, in actual practice, elementary strata may not be very homogeneous: there may be thousands of individual items in each elementary aggregate and the hypothesis of item homogeneity may not be warranted. Under these circumstances, it is important that the elementary index satisfy the commensurability test, since the units of measurement of the heterogeneous items in the elementary aggregate are arbitrary and hence *the price statistician can change the index simply by changing the units of measurement for some of the items*.[46]

This completes the listing of the tests for an elementary index. There remains the task of evaluating how many tests are passed by each of the five elementary indexes defined in section 2 above.

The following results hold:

- The Jevons elementary index P_J satisfies *all* of the above tests.
- The Dutot index P_D satisfies all of the tests with the exception of the important Commensurability Test T12, which it fails.
- The Carli and Harmonic elementary indexes, P_C and P_H, fail the price permutation test T9, the time reversal test T10 and the circularity test T11 but pass the other tests.
- The geometric mean of the Carli and Harmonic elementary indexes, P_{CSWD}, fails only the (suspect) price permutation test T9 and the circularity test T11.[47]

Since the Jevons elementary index P_J satisfies *all* of the tests, it emerges as being "best" from the viewpoint of the axiomatic approach to elementary indexes.

The Dutot index P_D satisfies all of the tests with the important exception of the Commensurability Test T12, which it fails. If there are heterogeneous items in the elementary aggregate, this is a rather serious failure and hence price statisticians should be careful in using this index under these conditions. If the N items under consideration are all measured in the same units and the products are close substitutes so that the product prices vary in a proportional manner over time, then the Dutot index could be used.[48] But if prices vary almost proportionally over time, then almost any reasonable index number formula will pick up the common factor of proportionality. The empirical example in Appendix A shows that if there are systematic divergent trends in prices, then the Dutot index can change dramatically as the units of measurement are changed.

The use of the Dutot, Carli and Harmonic indexes should be avoided.

The geometric mean of the Carli and Harmonic elementary indexes fail only the (suspect) price permutation test T9 and the circularity test T11. The failure of test T9 is probably not a fatal failure and P_{CSWD} will usually be numerically close to P_J so it will be close to satisfying the circularity test.[49]

The Carli and Harmonic elementary indexes, P_C and P_H, fail the (suspect) price permutation test T9, the time reversal test T10 and the circularity test T11 and pass the other tests. The failure of

[46] The empirical example in Appendix A shows that changing the units of measurement for the Dutot index makes a huge difference.
[47] But using the approximations given by (27) and (29), P_{CSWD} will satisfy circularity approximately.
[48] Evans (2012; 4) compared the Slovenian CPI with its corresponding Harmonized Index of Consumer Prices (HICP) and found very little difference over the period 1998-2011. The Slovenian national CPI used Dutot indexes at the elementary level and the Slovenian HICP used Jevons indexes at the elementary level.
[49] This is the case for the numerical example in the Appendix.

the time reversal test T10 (with an upward bias for the Carli and a downward bias for the Harmonic) is a rather serious failure and so price statisticians should not use these indexes.

In the following section, we present an argument due originally to Irving Fisher on why it is desirable for an index number formula to satisfy the time reversal test.

7. Fisher's Rectification Procedure and the Time Reversal Test

There is a problem with the Carli and Harmonic indexes that was first pointed out by Irving Fisher:[50] the rate of price change measured by the index number formula between two periods is dependent on which period is regarded as the base period. Thus the Carli index, $P_C(p^0,p^1)$ as defined by (2), takes period 0 as the base period and calculates (one plus) the rate of price change between periods 0 and 1.[51] Instead of choosing period 0 to be the base period, we could equally choose period 1 to be the base period and measure a reciprocal inflation rate going *backwards* from period 1 to period 0 and this *backwards measured inflation rate* would be $\sum_{n=1}^{N} (1/N)(p_n^0/p_n^1)$. In order to make this backwards inflation rate comparable to the forward inflation rate, we then take the reciprocal of $\sum_{n=1}^{N} (1/N)(p_n^0/p_n^1)$ and thus the overall inflation rate going from period 0 to 1 using period 1 as the base period is the following *Backwards Carli index* P_{BC}:[52]

$$(31)\ P_{BC}(p^0,p^1) \equiv [\sum_{n=1}^{N} (1/N)(p_n^1/p_n^0)^{-1}]^{-1} = P_H(p^0,p^1) ;$$

i.e., the Backwards Carli index turns out to equal the Harmonic index $P_H(p^0,p^1)$ defined earlier by (4).

If the forward and backwards methods of computing price change between periods 0 and 1 using the Carli formula were equal, then we would have the following equality:[53]

$$(32)\ P_C(p^0,p^1) = P_H(p^0,p^1).$$

Fisher argued that a good index number formula should satisfy (32) since the end result of using the formula should not depend on which period was chosen as the base period.[54] This seems to be a persuasive argument: if for whatever reason, a particular formula is favoured, where the base period 0 is chosen to be the period that appears before the comparison period 1, then the same

---

[50] "Just as the very idea of an index number implies a set of commodities, so it implies two (and only two) times (or places). Either one of the two times may be taken as the 'base'. Will it make a difference which is chosen? Certainly it *ought* not and our Test 1 demands that it shall not. More fully expressed, the test is that the formula for calculating an index number should be such that it will give the same ratio between one point of comparison and the other point, *no matter which of the two is taken as the base*." Irving Fisher (1922; 64).

[51] Instead of calculating price inflation between periods 0 and 1, period 1 can be replaced by any period t that follows period 1; i.e., p^1 in the Carli formula $P_C(p^0,p^1)$ can be replaced by p^t and then the index $P_C(p^0,p^t)$ measures price change between periods 0 and t. The arguments concerning $P_C(p^0,p^1)$ that follow apply equally well to $P_C(p^0,p^t)$.

[52] Fisher (1922; 118) termed the backward looking counterpart to the usual forward looking index the *time antithesis* of the original index number formula. Thus P_H is the time antithesis to P_C. The Harmonic index defined by (4) is also known as the Coggeshall (1887) index.

[53] Of course, equation (32) is *not* satisfied.

[54] "The justification for making this rule is twofold: (1) no reason can be assigned for choosing to reckon in one direction which does not also apply to the opposite, and (2) such reversibility does apply to any *individual* commodity. If sugar costs twice as much in 1918 as in 1913, then necessarily it costs half as much in 1913 as in 1918." Irving Fisher (1922; 64).

arguments that justify the forward looking version of the index number formula can be used to justify the backward looking version. If the forward and backward versions of the index agree with one another, then it does not matter which version is used and this equality provides a powerful argument in favour of using the formula. If the two versions do not agree, then rather than picking the forward version over the backward version, a more symmetric procedure would be to take an average of the forward and backward looking versions of the index formula.

Fisher provided an alternative way for justifying the equality of the two indexes in equation (32). He argued that the forward looking inflation rate using the Carli formula is $P_C(p^0,p^1) = \sum_{n=1}^N (1/N)(p_n^1/p_n^0)$. As noted above, the backwards looking inflation rate using the Carli formula is $\sum_{n=1}^N (1/N)(p_n^0/p_n^1) = P_C(p^1,p^0)$. Fisher[55] argued that the product of the forward looking and backward looking indexes should equal unity; i.e., a good formula *should* satisfy the following equality (which is equivalent to (32)):

$$(33)\ P_C(p^0,p^1)P_C(p^1,p^0) = 1.$$

But (33) is the usual *time reversal test* that was listed in the previous section. Thus Fisher provided a reasonably compelling case for the satisfaction of this test.

As we have seen in section 4 above,[56] the problem with the Carli formula is that it not only does not satisfy the equalities (32) or (33) but it *fails* (33) with the following definite inequality:

$$(34)\ P_C(p^0,p^1)P_C(p^1,p^0) > 1$$

unless the price vector p^1 is proportional to p^0 (so that $p^1 = \lambda p^0$ for some scalar $\lambda > 0$), in which case, (33) will hold. The main implication of the inequality (34) is that the use of *the Carli index will tend to give higher measured rates of inflation* than a formula that satisfies the time reversal test (using the same data set and the same weighting).

Fisher showed how the downward bias in the backwards looking Carli index P_H and the upward bias in the forward looking Carli index P_C could be cured. The Fisher *time rectification procedure*[57] as a general procedure for obtaining a bilateral index number formula that satisfies the time reversal test works as follows. Given a bilateral price index P, Fisher (1922; 119) defined the *time antithesis* P° for P as follows:

$$(35)\ P^{\circ}(p^0,p^1,q^0,q^1) \equiv 1/P(p^1,p^0,q^1,q^0).$$

Thus P° is equal to the reciprocal of the price index that has reversed the role of time, $P(p^1,p^0,q^1,q^0)$. Fisher (1922; 140) then showed that the geometric mean of P and P°, say $P^* \equiv [P\times P^{\circ}]^{1/2}$, satisfies the time reversal test, $P^*(p^0,p^1,q^0,q^1)P^*(p^1,p^0,q^1,q^0) = 1$.

In the present context, P_C is only a function of p^0 and p^1, but the same rectification procedure works and the time antithesis of P_C is the harmonic index P_H. Applying the Fisher rectification

[55] "Putting it in still another way, more useful for practical purposes, the forward and backward index number multiplied together should give unity." Irving Fisher (1922; 64).

[56] Recall the inequalities (7) and (8) above.

[57] Actually, Walsh (1921b; 542) showed Fisher (1921) how to rectify a formula so it would satisfy the factor reversal test and Fisher (1922) simply adapted the methodology of Walsh to the problem of rectifying a formula so that it would satisfy the time reversal test.

procedure to the Carli index, the resulting rectified Carli formula, P_{RC}, turns out to equal the Carruthers, Sellwood and Ward (1980) and Dalén elementary index P_{CSWD} defined earlier by (5):

(36) $P_{RC}(p^0,p^1) \equiv [P_C(p^0,p^1)P_{BC}(p^0,p^1)]^{1/2} = [P_C(p^0,p^1)P_H(p^0,p^1)]^{1/2} = P_{CSWD}(p^0,p^1)$.

Thus P_{CSWD} is the geometric mean of the forward looking Carli index P_C and its backward looking counterpart $P_{BC} = P_H$, and, of course, P_{CSWD} will satisfy the time reversal test.[58]

8. Conclusion

The main results in this chapter can be summarized as follows:

- In order to define a "best" elementary index number formula, it is necessary to have a target index number concept. In section 2, it is suggested that normal bilateral index number theory applies at the elementary level as well as at higher levels and hence the target concept should be one of the Fisher, Törnqvist or Walsh formulae.
- When aggregating the prices of the same narrowly defined item within a period, the narrowly defined unit value is a reasonable target price concept.
- The axiomatic approach to traditional elementary indexes (i.e., no quantity or value weights are available) supports the use of the Jevons formula under most circumstances.[59] The Carruthers, Sellwood and Ward formula can be used as an alternative to the Jevons formula but both will give much the same numerical answers.
- The Carli index has an upward bias (with respect to satisfying the time reversal test) and the Harmonic index has a downward bias.
- All five unweighted elementary indexes are not really satisfactory. A much more satisfactory approach would be to collect quantity or value information along with price information and form sample superlative indexes as the preferred elementary indexes. However, if a chained superlative index is calculated, it should be examined for chain drift; i.e., a chained index should only be used if the data are relatively smooth and subject to long term trends rather than short term fluctuations.[60]

Appendix A: Alternative Approaches to the Treatment of Access Charges

An interesting Consumer Price Index problem arises when there is a fixed access charge for the right of consumers to purchase products or services from a supplier. Examples of such charges are annual club memberships, annual fees for the use of a credit card and fixed charges for access to telecommunication services. In this Appendix, we will outline three different approaches that could be used by consumer price statisticians to deal with these charges which are independent of the actual consumption of the goods and services that the payment of the a fixed charge allows consumers to purchase.

---

[58] See Chart 2 in the Appendix where it will be seen that for our empirical example, the Jevons index cannot be distinguished from the Carruthers, Sellwood, Ward and Dalen index.

[59] One exception to this advice is when a price can be zero in one period and positive in another comparison period. In this situation, the Jevons index will fail and the corresponding item will have to be ignored in the elementary index. The problems raised by missing prices will be considered at greater length in the subsequent chapters on multilateral methods and strongly seasonal commodities.

[60] If the price and quantity data are subject to large fluctuations, then multilateral methods should be used instead of a bilateral index number formula. Multilateral methods will be studied in chapter 7.

The notation that is used in this Appendix is similar to that used in section 2 above with some new notation for the fixed charge. Thus let $p^t \equiv [p_1^t,...,p_N^t]$ and $q^t \equiv [q_1^t,...,q_N^t]$ be the period t price and quantity vectors for the purchases of the goods or services that the payment of the access charge $P^t > 0$ allows the consumer or group of consumers to purchase for t = 0,1.

Define e^t as the *period t expenditure* on the actual goods and services purchased and v^t as the *value of period t total expenditures* on the group of commodities which is equal to e^t plus the period t access fixed charge P^t. It is also useful to define the period t fixed cost *margin* m^t as the ratio of P^t to e^t. Thus we have the following definitions:

(A.1) $e^t \equiv p^t \cdot q^t \equiv \Sigma_{n=1}^N p_n^t q_n^t$; $v^t \equiv p^t \cdot q^t + P^t = e^t + P^t$; $m^t \equiv P^t/e^t$; t = 0,1.

In the analysis which follows, we will look at some "practical" price indexes and compare their magnitudes. However, before we define these indexes, it is useful to look at three alternative utility maximization models which will help to motivate the alternative practical indexes.

Suppose the "consumer" has the utility function f(q). The first utility maximization model that we will consider is a "traditional" model which treats the period t fixed charge as a charge on the "income" that the consumer allocates to the N commodities in the group of commodities under consideration. Thus the *Model 1* period t utility maximization problem for the subgroup of commodities under consideration is the following one:

(A.2) $\max_q \{f(q): p^t \cdot q \le v^t - P^t = e^t ; q \ge 0_N\}$.

If the consumer price index were constructed in only a single stage, then Model 1 is a "practical" model that price statisticians could use to guide the construction of the national CPI. However, a typical CPI is constructed by aggregating over both commodity groupings and outlets or households. In order to implement the Model 1 approach, price statisticians would have to keep track of the various fixed charges that occur for various outlets and commodity groups as well as collecting the basic price and quantity information. The CPI subindexes which would be computed using this approach would also have to include (separately) information on the fixed charges by commodity group. The national accounts division of the national statistical agency would not be able to take a CPI subindex and use it for deflation purposes if that subgroup of commodities included substantial fixed charges; i.e., the period t CPI subindex would be appropriate for deflating the actual commodity expenditures e^t but *the subindex would not be appropriate for deflating actual group expenditures* (including the fixed charges) P^t.

The second utility maximization problem treats the access charge as a separate commodity that gives utility to consumers even if they do not consume any products or services that the access charge enables. The new utility function is $f^*(q,Q)$ where Q = 1 represents the contribution of access to overall utility for the subgroup of commodities under consideration. Thus the *Model 2* period t utility maximization problem for the subgroup of commodities under consideration is the following one:

(A.3) $\max_q \{f^*(q,1): p^t \cdot q + P^t \le v^t ; q \ge 0_N\}$.

This way of thinking about fixed charges in the telecommunications context is used by national regulators. The approach taken to the treatment of access charges is of some importance in measuring the productivity of telecommunications firms as will be seen in the example which follows. The advantage of this approach is that the CPI index that is constructed using this framework will be suitable for national accounts deflation purposes; i.e., the period t subindex that is a result of using this approach can be used to deflate total period t expenditures v^t on the commodity class.

The third utility maximization problem allocates the period t fixed charge P^t in a proportional to expenditure manner across the "usage" prices p^t. Recall that (A.1) defined the period t margin m^t as P^t/e^t. The margin is treated in much the same way that a general sales tax is treated; i.e., it is added on to the period t usage prices p^t. Thus the *Model 3* period t utility maximization problem for the subgroup of commodities under consideration is the following one:[61]

(A.4) $\max_q \{f(q): (1 + m^t)p^t \cdot q \leq v^t \,;\, q \geq 0_N\}$.

When price statisticians apply the economic approach to index number theory, it is assumed that the observed period t quantity vector q^t solves the corresponding period t utility maximization problem. It is also assumed that the first inequality constraint in problems (A.2)-(A.4) holds with equality. Thus if q^t solves problem (A.2) for period t, then $p^t \cdot q^t = v^t - P^t = e^t$ for t = 0,1; if q^t solves problem (A.3) for period t, then $p^t \cdot q^t = v^t - P^t = e^t$ for t = 0,1 and if q^t solves problem (A.4) for period t, then $(1 + m^t)p^t \cdot q^t = v^t$ for t = 0,1. Using the definitions for m^t, e^t and v^t in (A.1), it can be seen that $(1 + m^t)p^t \cdot q^t = [1 + (P^t/e^t)]p^t \cdot q^t = [1 + (P^t/p^t \cdot q^t)]p^t \cdot q^t = p^t \cdot q^t + P^t = v^t$ for t = 0,1. Thus for all three utility maximization problems, it is assumed that the various equalities in definitions (A.1) are satisfied.

We use the above alternative models of economic behavior to motivate the definitions of the alternative Laspeyres and Paasche indexes. Below, we will define the Laspeyres and Paasche indexes that correspond to the three models and compare their magnitudes.

The Laspeyres and Paasche indexes comparing the prices of period 1 to the corresponding prices of period 0 *using the Model 1 framework*, P_{L1} and P_{P1} respectively, are defined as follows:

(A.5) $P_{L1} \equiv p^1 \cdot q^0 / p^0 \cdot q^0$;
(A.6) $P_{P1} \equiv p^1 \cdot q^1 / p^0 \cdot q^1$.

The Laspeyres and Paasche indexes comparing the prices of period 1 to the corresponding prices of period 0 *using the Model 2 framework*, P_{L2} and P_{P2} respectively, are defined as follows:

(A.7) $P_{L2} \equiv [p^1 \cdot q^0 + P^1]/[p^0 \cdot q^0 + P^0]$
$= [P_{L1} + (P^1/e^0)]/[1 + (P^0/e^0)]$ dividing numerator and denominator by e^0;

(A.8) $P_{P2} \equiv [p^1 \cdot q^1 + P^1]/[p^0 \cdot q^1 + P^0]$
$= [1 + (P^1/e^1)]/[P_{P1}^{-1} + (P^0/e^1)]$ dividing numerator and denominator by e^1.

We also used definitions (A.5) and (A.6) in deriving the second lines of (A.7) and (A.8).

Using definitions (A.5) and (A.7), it is possible to compare P_{L1} to P_{L2}:

(A.9) $P_{L1} - P_{L2} = P_{L1} - \{[P_{L1} + (P^1/e^0)]/[1 + (P^0/e^0)]\}$
$= [1 + (P^0/e^0)]^{-1}[P_{L1}\{1 + (P^0/e^0)\} - P_{L1} - (P^1/e^0)]$
$= [1 + m^0]^{-1}[P_{L1}(P^0/e^0) - (P^1/e^0)]$

[61] Models 1 and 3 will not work if $q^t = 0_N$ for some period t. If this case occurs empirically, then Model 2 or some other model will have to be used.

$$= [1 + m^0]^{-1}[P_{L1} (P^0/e^0) - (P^1/P^0)(P^0/e^0)]$$
$$= [m^0/(1 + m^0)][P_{L1} - (P^1/P^0)].$$

Thus if the Laspeyres price index P_{L1} for the N products that are made available by paying the access charge in each period is equal to one plus the growth rate in the access charges, P^1/P^0, then P_{L1} will be equal to P_{L2} (which is the Laspeyres price index that treats the access charge as a normal commodity). If P_{L1} is greater than P^1/P^0, then P_{L1} will be greater than P_{L2}; if P_{L1} is less than P^1/P^0, then P_{L1} will be less than P_{L2}. If m^0 is large and the difference between P_{L1} and P^1/P^0 is also large, then the difference between P_{L1} and P_{L2} can be substantial. This case can occur in the case of a telecommunications subindex.[62]

Using definitions (A.6) and (A.8), it is possible to compare P_{P1} to P_{P2} but the resulting formula is a bit more complicated:

$$\text{(A.10)}\ P_{P1}^{-1} - P_{P2}^{-1} = P_{P1}^{-1} - \{[P_{P1}^{-1} + (P^0/e^1)]/[1 + (P^1/e^1)]\}$$
$$= [1 + (P^1/e^1)]^{-1}\,[P_{P1}^{-1}\{1 + (P^1/e^1)\} - P_{P1}^{-1} - (P^0/e^1)]$$
$$= [1 + m^1]^{-1}[P_{P1}^{-1}(P^1/e^1) - (P^0/e^1)]$$
$$= [1 + m^1]^{-1}[P_{P1}^{-1}(P^1/e^1) - (P^0/P^1)(P^1/e^1)]$$
$$= [m^1/(1 + m^1)][P_{P1}^{-1} - (P^1/P^0)^{-1}].$$

Multiply both sides of (A.10) by $P_{P1}P_{P2}$ and the following expression is obtained:

$$\text{(A.11)}\ P_{P2} - P_{P1} = [m^1/(1 + m^1)]\, P_{P2}\, [1 - P_{P1}\, (P^1/P^0)^{-1}]$$
$$= [m^1/(1 + m^1)]\, P_{P2}\, [P^1/P^0]^{-1}\, [(P^1/P^0) - P_{P1}].$$

Finally, multiply both sides of (A.11) through by −1 in order to obtain the following counterpart to (A.9):

$$\text{(A.12)}\ P_{P1} - P_{P2} = [m^1/(1 + m^1)]\, P_{P2}\, [P^1/P^0]^{-1}\, [P_{P1} - (P^1/P^0)].$$

Thus if the Paasche price index P_{P1} for the N products that are made available by paying the access charge in each period is equal to one plus the growth rate in the access charges, P^1/P^0, then P_{P1} will be equal to P_{P2} (which is the Paasche price index that treats the access charge as a normal commodity). If P_{P1} is greater than P^1/P^0, then P_{P1} will be greater than P_{P2}; if P_{P1} is less than P^1/P^0, then P_{P1} will be less than P_{P2}. If m^1 is large and the difference between P_{P1} and P^1/P^0 is also large, then the difference between P_{P1} and P_{P2} can be substantial.[63]

We turn now to the Model 3 framework. The Laspeyres and Paasche indexes comparing the prices of period 1 to the corresponding prices of period 0 *using the Model 3 framework*, P_{L3} and P_{P3} respectively, are defined as follows:

$$\text{(A.13)}\ P_{L3} \equiv (1 + m^1)p^1\cdot q^0/(1 + m^0)p^0\cdot q^0$$

[62] Our analysis for the case of Laspeyres price indexes also applies to other fixed basket indexes; i.e., simply replace the base period quantity vector q^0 by the fixed basket quantity vector q^* and apply our analysis pertaining to the differences between the various Laspeyres indexes. The definitions for e^0, v^0 and m^0 become $e^0 \equiv p^0\cdot q^*$, $v^0 \equiv e^0 + P^0$ and $m^0 \equiv P^0/e^0$. P_{L1} becomes $p^1\cdot q^*/p^0\cdot q^*$, P_{L2} becomes $[p^1\cdot q^* + P^1]/[p^0\cdot q^* + P^0]$ and P_{L3} (which will be defined below) becomes $(1 + m^1)p^1\cdot q^*/(1 + m^0)p^0\cdot q^*$ where $e^1 \equiv p^1\cdot q^*$, $v^1 \equiv e^1 + P^1$ and $m^1 \equiv P^1/e^1$.

[63] Note that the conditions for "bias" between P_{L1} and P_{L2} and for "bias" between P_{P1} and P_{P2} are very similar in structure.

$= [(1 + m^1)/(1 + m^0)]\, P_{L1}$ dividing numerator and denominator by e^0;

(A.14) $P_{P3} \equiv (1 + m^1)p^1 \cdot q^1/(1 + m^0)p^0 \cdot q^1$
$= (1 + m^1)/[(1 + m^0)P_{P1}{}^{-1}]$ dividing numerator and denominator by e^1;
$= [(1 + m^1)/(1 + m^0)]P_{P1}$.

It is very easy to compare P_{L3} to P_{L1} and to compare P_{P3} to P_{P1}. Using definitions (A.13) and (A.14), we have:

(A.15) $P_{L3}/P_{L1} = P_{P3}/P_{P1} = (1 + m^1)/(1 + m^0)$.

Thus P_{L3} will equal P_{L1} and P_{P3} will equal P_{P1} if $m^1 \equiv P^1/e^1$ is equal to $m^0 \equiv P^0/e^0$ or if $P^1/P^0 = e^1/e^0$. P_{L3} will be greater than P_{L1} and P_{P3} will be greater than P_{P1} if $m^1 > m^0$ or if $P^1/P^0 > e^1/e^0$. These results are very straightforward and easy to understand.

The more interesting comparisons are between P_{L3} and P_{L2} and between P_{P3} and P_{P2}. For the Laspeyres comparisons, using (A.7) and (A.13), we have:

(A.16) $P_{L2} - P_{L3} = \{[P_{L1} + (P^1/e^0)]/[1 + (P^0/e^0)]\} - \{(1 + m^1)P_{L1}/(1 + m^0)\}$
$= [1 + m^0]^{-1}[P_{L1} + (P^1/e^0) - (1 + \{P^1/e^1\})P_{L1}]$
$= [1 + m^0]^{-1}[(P^1/e^0) - (P^1/e^1)P_{L1}]$
$= m^1[1 + m^0]^{-1}[(e^1/e^0) - P_{L1}]$.

Thus if the usage expenditure ratio, e^1/e^0, is equal to the Laspeyres price index for the available products or services, P_{L1}, then P_{L2} will equal P_{L3}. In the telecommunications context, typically usage expenditures will grow more rapidly than the usage Laspeyres price index so that e^1/e^0 will be much greater than P_{L1} which will imply that P_{L2} will be greater than P_{L3} using (A.16). If m^1 is also large, then P_{L2} will be substantially greater than P_{L3}.[64] In the telecommunications context, the choice of index number method will matter as will be shown in the empirical example below.

Using definitions (A.8) and (A.14), we have the following equality:

(A.17) $P_{P2}{}^{-1} - P_{P3}{}^{-1} = \{[P_{P1}{}^{-1} + (P^0/e^1)]/[1 + (P^1/e^1)]\} - \{(1 + m^0)P_{P1}{}^{-1}/(1 + m^1)\}$
$= [1 + m^1]^{-1}\{P_{P1}{}^{-1} + (P^0/e^1) - P_{P1}{}^{-1} - P_{P1}{}^{-1}(P^0/e^0)\}$
$= [1 + m^1]^{-1}[(P^0/e^1) - P_{P1}{}^{-1}(P^0/e^0)]$
$= m^0[1 + m^1]^{-1}[(e^1/e^0)^{-1} - P_{P1}{}^{-1}]$.

Divide both sides of (A.17) by $P_{P3}{}^{-1}$ in order to obtain the following equalities:

(A.18) $P_{P3}/P_{P2} - 1 = m^0[1 + m^1]^{-1}[(e^1/e^0)^{-1} - P_{P1}{}^{-1}]P_{P3}$
$= m^0[1 + m^1]^{-1}[(e^1/e^0)^{-1} - P_{P1}{}^{-1}](1 + m^1)(1 + m^0)^{-1}P_{P1}$ using (A.14)
$= m^0\,(1 + m^0)^{-1}[P_{P1}(e^1/e^0)^{-1} - 1]$
$= m^0\,(1 + m^0)^{-1}(e^1/e^0)^{-1}[P_{P1} - (e^1/e^0)]$.

Multiply both sides of (A.18) by $-P_{P2}$ in order to obtain the following equality:

(A.19) $P_{P2} - P_{P3} = m^0\,(1 + m^0)^{-1}(e^1/e^0)^{-1}P_{P2}[(e^1/e^0) - P_{P1}]$.

[64] Note that $e^1/e^0 = P_{L1}Q_{P1}$ where $Q_{P1} \equiv p^1 \cdot q^1/p^1 \cdot q^0$ is the Paasche quantity index for usage expenditures. Thus (A.16) can be rewritten as $P_{L2} - P_{L3} = m^1[1 + m^0]^{-1}P_{L1}[Q_{P1} - 1]$. Thus if $Q_{P1} > 1$, then $P_{L2} > P_{L3}$.

Thus if the usage expenditure ratio, e^1/e^0, is equal to the Paasche price index for the available products or services, P_{P1}, then P_{P2} will equal P_{P3}. As noted above, in the telecommunications context, typically usage expenditures will grow more rapidly than the usage Paasche price index so that e^1/e^0 will be much greater than P_{P1} which will imply that P_{P2} will be greater than P_{P3} using (A.19). If m^0 is also large, then P_{P2} will be substantially greater than P_{P3}.[65] Thus again, in the telecommunications context, the choice of index number method will matter.

For empirical evidence on the huge differences in actual national indexes that the alternative treatment of access charges can make in the telecommunications context, we draw on the UK data that is listed in the recent study by Abdirahman, Coyle, Heys and Stewart (2020).[66] The UK retail telecom revenues for fixed lines $v_n^t \equiv p_n^t q_n^t$ and the corresponding quantities q_n^t for the years 2010-2017 are listed below in Table A.1. These data are not "pure" CPI data in that they do not refer to the purchases by households but instead refer to all retail purchases. However, these data will serve as an example that will show that the above three alternative treatment of access charges can lead to significantly different price (and quantity) indexes.

Table A.1: Fixed Line UK Retail Telecommunications Revenues and Quantities

Year t	v_1^t	v_2^t	v_3^t	v_4^t	v_5^t	q_1^t	q_2^t	q_3^t	q_4^t	q_5^t	e^t	v^t
2010	935	293	849	824	3259	65134	4850	5642	14736	23752	2901	6160
2011	787	237	675	742	3375	56083	4570	4471	13066	23872	2441	5816
2012	723	198	566	659	3706	51985	4111	3902	11506	24462	2146	5852
2013	673	155	488	620	3964	46191	3455	3351	10681	24970	1936	5900
2014	577	132	430	620	4148	40766	3015	2940	9028	25549	1759	5907
2015	498	123	369	604	4462	35586	2749	2735	8855	26075	1594	6056
2016	428	111	270	596	4776	30471	2169	2811	7826	26482	1405	6181
2017	362	89	228	543	4969	24705	1550	2587	6126	26661	1222	6191

The revenues in Table A1 are expressed in millions of United Kingdom Pounds. The five "products" and their units of measurement for the corresponding quantities are as follows:

- 1 = UK geographic calls in millions of minutes;
- 2 = International calls in millions of minutes;
- 3 = Calls to mobile phones in millions of minutes;
- 4 = Other calls in millions of minutes;
- 5 = Fixed line access charges; units are the number of lines in thousands.

Note that $e^t \equiv v_1^t + v_2^t + v_3^t + v_4^t$ is the total revenue or expenditure for year t on the various types of calls made from fixed lines in the UK and $v^t \equiv e^t + v_5^t$ is total expenditure including access charges v_5^t. The ratio of access charges in year t to the corresponding total call revenues is the *margin* $m^t \equiv v_5^t/e^t$ which is listed in Table A2 below. From Table A2, it can be seen that m^t increases steadily from 1.12 in 2010 to 4.07 in 2017. Thus the treatment of access charges is likely to make a substantial difference to any telecom price index based on the above data.

65 Note that $e^1/e^0 = P_{P1}Q_{L1}$ where $Q_{L1} \equiv p^0 \cdot q^1/p^0 \cdot q^0$ is the Laspeyres quantity index for usage expenditures. Thus (A.19) can be rewritten as $P_{P2} - P_{P3} = m^0 (1 + m^0)^{-1}(e^1/e^0)^{-1}P_{P2}P_{P1}[Q_{L1} - 1]$. Thus if $Q_{L1} > 1$, then $P_{P2} > P_{P3}$.

66 Their recent study extends their earlier important study; see Abdirahman, Coyle, Heys and Stewart (2017). These papers make clear that the alternative treatment of access charges makes a big difference not only to price indexes but also to the measurement of national output, consumption and productivity.

The *unit value prices* for each product can be constructed using the information in Table A1; i.e, we have $p_n^t \equiv v_n^t/q_n^{t*}$ for n = 1,...,5 and t = 2010,...,2017. In order to see more clearly how the prices of the various telecom products have changed over the sample period, normalize the unit value prices to equal 1 in the base year, 2010; i.e., define the *normalized prices and quantities*, p_n^{t*} and q_n^{t*}, as follows:[67]

(A.20) $p_n^{t*} \equiv p_n^t/p_n^{2010}$; $q_n^{t*} \equiv q_n^t p_n^{2010}$; $\quad$ n = 1,...,5; t = 2010,...,2017.

Table A.2 below lists the normalized prices and quantities for the five products along with the margin series, $m^t \equiv v_5^t/e^t$.

Table A.2: Normalized Prices and Quantities for the UK Fixed Line Retail Sector

Year t	p_1^{t*}	p_2^{t*}	p_3^{t*}	p_4^{t*}	p_5^{t*}	q_1^{t*}	q_2^{t*}	q_3^{t*}	q_4^{t*}	q_5^{t*}	m^t
2010	1.0000	1.0000	1.0000	1.0000	1.0000	935.00	293.00	849.00	824.00	3259.00	1.1234
2011	0.9776	0.8584	1.0033	1.0156	1.0304	805.07	276.08	672.79	730.62	3275.47	1.3826
2012	0.9689	0.7973	0.9640	1.0243	1.1042	746.25	248.36	587.17	643.39	3356.42	1.7269
2013	1.0150	0.7426	0.9678	1.0381	1.1570	663.07	208.72	504.25	597.25	3426.12	2.0475
2014	0.9860	0.7247	0.9720	1.2282	1.1833	585.20	182.14	442.41	504.82	3505.57	2.3582
2015	0.9749	0.7406	0.8966	1.2198	1.2472	510.84	166.07	411.56	495.15	3577.74	2.7993
2016	0.9785	0.8471	0.6383	1.3619	1.3144	437.41	131.03	423.00	437.61	3633.58	3.3993
2017	1.0208	0.9505	0.5857	1.5852	1.3583	354.64	93.64	389.29	342.55	3658.14	4.0663

It can be seen that relative prices and relative quantities vary considerably over the sample period. This will lead to dispersion among alternative index number formulae. We utilize the data in the above Tables to compute alternative indexes for each of the three approaches outlined above for the treatment of access charges.[68]

For the Approach 1 indexes, we ignore the access charges and simply compute the alternative indexes using only the prices and quantities for the first 4 products. In Table A.3 below, the "unweighted" price indexes[69] that were defined in section 4 above are listed. The fixed base Harmonic, Caruthers-Sellwood-Ward-Dalén, and Carli indexes, P_H^t, P_{CWSD}^t, P_C^t and their chained counterparts, P_{HCH}^t, P_{CWSDCH}^t, P_{CCH}^t, are listed in this table. The fixed base and chained Dutot and Jevons indexes coincide and so these indexes are simply listed as P_D^t and P_J^t in Table A.3. These indexes were calculated using the $p_n^t = v_n^t/q_n^t$ where the v_n^t and q_n^t are listed in Table A.1. All of these indexes with the exception of the Dutot index are independent of the units of measurement. Instead of using the original units of measurement to calculate the Dutot index, we could "standardize" the unit value prices by using the normalized prices $p_n^{t*} \equiv p_n^t/p_n^{2010}$ listed in Table A.2 and calculate a new Dutot index using the normalized prices.[70] It turns out that this new Dutot index P_{DN}^t using normalized prices in place of the original prices is equal to the fixed base Carli index P_C^t so we did not list P_{DN}^t in Table A.3. For all of the index number formulae that appear in

[67] If we change the units of measurement of prices, then we have to change the corresponding units of measurement for quantities in the opposite direction in order to preserve values.

[68] We will also consider a fourth approach which is relevant for producer price indexes.

[69] The term "unweighted" really means "equally weighted". These indexes do not make any use of quantity or value information. Thus they do not take into account the economic importance of each product. This is not a problem if expenditure shares are roughly equal but typically this is not the case.

[70] Thus define $P_{DN}^t \equiv [p_1^{t*}+p_2^{t*}+p_3^{t*}+p_4^{t*}]/[p_1^{2010*}+p_2^{2010*}+p_3^{2010*}+p_4^{2010*}] = [p_1^{t*}+p_2^{t*}+p_3^{t*}+p_4^{t*}]/[4] = (1/4)\Sigma_{n=1}^4 (p_n^t/p_n^{2010}) \equiv P_C^t$ where the second equality follows using $p_n^{2010*} = 1$ for n = 1,2,3,4. Thus the Dutot index using normalized prices in place of the initial prices is equal to the fixed base Carli index, P_C^t, for t = 2010,...,2017.

Table A.3 with the exception of the Dutot index P_D^t, it does not matter whether we use the prices and quantities listed in Table A.1 or their normalized counterparts listed in Table A.2.

Table A.3: Approach 1 Unweighted Price Indexes

Year t	P_D^t	P_J^t	P_H^t	P_{CSWD}^t	P_C^t	P_{HCH}^t	P_{CSWDCH}^t	P_{CCH}^t
2010	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
2011	**0.9733**	**0.9616**	**0.9594**	**0.9616**	**0.9637**	**0.9594**	**0.9616**	**0.9637**
2012	**0.9404**	**0.9345**	**0.9302**	**0.9344**	**0.9386**	**0.9319**	**0.9345**	**0.9370**
2013	**0.9358**	**0.9328**	**0.9241**	**0.9324**	**0.9409**	**0.9294**	**0.9328**	**0.9362**
2014	**0.9705**	**0.9610**	**0.9440**	**0.9607**	**0.9777**	**0.9544**	**0.9611**	**0.9677**
2015	**0.9314**	**0.9427**	**0.9278**	**0.9428**	**0.9580**	**0.9355**	**0.9427**	**0.9499**
2016	**0.8445**	**0.9213**	**0.8882**	**0.9217**	**0.9565**	**0.8972**	**0.9204**	**0.9442**
2017	**0.8851**	**0.9742**	**0.9153**	**0.9736**	**1.0355**	**0.9447**	**0.9731**	**1.0024**

It can be seen that the Dutot index using normalized prices, $P_{DN}^t = P_C^t$, ends up well above its Jevons index counterpart, P_J^t, when t = 2017. In section 5 above, we indicated that under certain circumstances, the Jevons and Dutot indexes should be approximately equal; see the approximate equality (16). Using our current notation, the approximate equality (16) becomes the following one:

$$\text{(A.21)}\quad P_J^t \approx P_{DN}^t[1 + (1/2)\text{var}(\varepsilon^{2010}) - (1/2)\text{var}(\varepsilon^t)] ; \qquad t = 2011, 2012, ..., 2017$$

where $p_A^t \equiv (1/4)(p_1^{t*} + p_2^{t*} + p_3^{t*} + p_4^{t*})$, $\varepsilon_n^t \equiv (p_n^{t*}/p_A^t) - 1$ for n = 1,2,3,4, $\varepsilon^t \equiv [\varepsilon_1^t, \varepsilon_2^t, \varepsilon_3^t, \varepsilon_4^t]$ and $\text{var}(\varepsilon^t) \equiv (1/4)\Sigma_{n=1}^4 (\varepsilon_n^t)^2$ for t = 2010,...,2017. Since the normalized prices p_n^{t*} all equal 1 when t = 2010, we see that $\text{var}(\varepsilon^{2010}) = 0$. Moreover, because p_3^{t*} trends down and p_4^{t*} trends up as t increases, $\text{var}(\varepsilon^t)$ is increasing over time and hence, using the above approximate equality, it can be seen that P_J^t will tend to be less than P_{DN}^t and the gap will grow over time as the variance $\text{var}(\varepsilon^t)$ increases. Thus we have an explanation for why the gap between P_J^t and $P_{DN}^t = P_C^t$ increases over time.[71]

The large differences between the Dutot index using the original units of measurement, P_D^t, and the version of the Dutot index that uses normalized prices, P_{DN}^t (which turns out to be equal to the fixed base Carli index P_C^t), indicates that *the Dutot formula should be used with extreme caution* even if there are common units of measurement for the individual commodities in scope for the index.

From Table A.3, it can be seen that the Jevons index is approximately equal to both the fixed base and chained Carruthers, Ward, Sellwood and Dalén indexes; i.e., we have the following approximate equalities which are consistent with the analysis in section 5 above:

$$\text{(A.22)}\quad P_J^t \approx P_{CSWD}^t \approx P_{CSWDCH}^t ; \qquad t = 2011, 2012, ..., 2017.$$

Looking at Table A.3, it can be seen that the following inequalities hold:

[71] As we have seen above, using normalized prices in the Dutot formula converts the fixed base Dutot index into a fixed base Carli index. Hence the divergence is explained by the fact that a geometric mean of numbers that are not all equal (the Jevons index) will always be less than the corresponding arithmetic mean (the Dutot index using normalized prices which is the fixed base Carli index). Recall that the indexes other than the Dutot index are invariant to the units of measurement.

(A.23) $P_H^t < P_J^t < P_C^t$; $P_{HCH}^t < P_J^t < P_{CCH}^t$; t = 2011, 2012, ..., 2017.

These inequalities are consistent with the inequalities (9) in section 5 above.

Note that in 2017, the Dutot index P_D^{2017} was equal to 0.8851 while the fixed base Carli index P_C^{2017} was equal to 1.0355. Thus P_C^{2017}/P_D^{2017} = 1.0355/0.8851 = 1.170. Thus there is a 17.0 % spread between these indexes listed in Table A.3, which is substantial. The choice of an unweighted index number formula matters.

The fact that the Jevons indexes P_J^t approximate the Carruthers, Sellwood, Ward and Dalén indexes P_{CSWD}^t can be demonstrated in another way. From Diewert (1978; 893), it is known that the Fisher index number formula, $P_F(p^1,p^t,q^1,q^t)$, approximates the Törnqvist Theil index, $P_T(p^1,p^t,q^1,q^t)$, to the second order around a point where $p^1 = p^t$ and $q^1 = q^t$. It is obvious that the Törnqvist Theil index collapses down to the Jevons index $P_J^t = P_J(p^1,p^t) \equiv \Pi_{n=1}^N (p_{tn}/p_{1n})^{1/N}$ if each expenditure share in periods 1 and t is equal to 1/N. Reinsdorf and Triplett (2009; 63) and Diewert (2013; 6) showed that if all expenditure shares in periods 1 and t are equal to 1/N, then the Fisher index collapses down to the Carruthers, Sellwood, Ward and Dalén index $P_{CSWD}(p^1,p^t) = P_{CSWD}^t$. Thus using the Diewert (1978; 893) second order approximation result, it can be seen that $P_J(p^1,p^t)$ will approximate $P_{CSWD}(p^1,p^t)$ to the second order around any point where $p^1 = p^t$. [72] In chapter 5, the Walsh (1901; 398) (1921a; 97) index was defined as follows:[73]

$$(A.24)\ P_W(p^1,p^t,q^1,q^t) \equiv \Sigma_{n=1}^N p_{tn}(q_{tn}q_{1n})^{1/2}/\Sigma_{n=1}^N p_{1n}(q_{tn}q_{1n})^{1/2}$$
$$= \Sigma_{n=1}^N (p_{tn}/p_{1n})^{1/2} (p_{tn}q_{tn}p_{1n}q_{1n})^{1/2}/\Sigma_{n=1}^N (p_{1n}/p_{tn})^{1/2} (p_{tn}q_{tn}p_{1n}q_{1n})^{1/2}$$
$$= \Sigma_{n=1}^N (p_{tn}/p_{1n})^{1/2} (s_{tn}s_{1n})^{1/2}/\Sigma_{n=1}^N (p_{1n}/p_{tn})^{1/2} (s_{tn}s_{1n})^{1/2}$$

where $s_{1n} \equiv p_{1n}q_{1n}/p^1\cdot q^1$ and $s_{tn} \equiv p_{tn}q_{tn}/p^t\cdot q^t$ for n = 1,...,N are the period 1 and t expenditure shares. If we again assume that all expenditure shares in periods 1 and t are equal to 1/N, then the Walsh index collapses down to the following *Dikhanov elementary index* $P_{DI}(p^1,p^t)$: [74]

$$(A.25)\ P_{DI}(p^1,p^t) \equiv \Sigma_{n=1}^N (p_{tn}/p_{1n})^{1/2} /\Sigma_{n=1}^N (p_{1n}/p_{tn})^{1/2}.$$

Diewert's 1978 second order approximation result also applies to Walsh and Fisher indexes so it carries over in the present special case where expenditure shares are assumed to be equal and constant across periods. Thus $P_{DI}(p^1,p^t)$ will approximate $P_J(p^1,p^t)$ and $P_{CSWD}(p^1,p^t)$ to the second order around any point where $p^t = \lambda p^1$. [75]

We turn to the *Approach 1 weighted indexes* for our UK telecom data set. Denote the year t fixed base Laspeyres, Paasche and Fisher indexes by P_{L1}^t, P_{P1}^t and P_{F1}^t and their chained counterparts by P_{LCH1}^t, P_{PCH1}^t and P_{FCH1}^t. These indexes are listed in Table A.4.

Table A.4: Approach 1 Laspeyres, Paasche and Fisher Indexes

Year t	P_{L1}^t	P_{P1}^t	P_{LCH1}^t	P_{PCH1}^t	P_{F1}^t	P_{FCH1}^t
2010	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**

[72] This second order approximation result also holds if $p^t = \lambda p^1$ for any scalar $\lambda > 0$.
[73] All prices and quantities are assumed to be positive.
[74] Yuri Dikhanov in a private communication suggested this approximation to the Walsh index.
[75] Using the data for our telecom example, the Dikhanov indexes P_{DI}^t were as follows; 1.0000, 0.9616, 0.9345, 0.9327, 0.9609, 0.9427, 0.9214, 0.9740. These numbers are very close to their P_J^t and P_{CWSD}^t counterparts.

2011	**0.9839**	**0.9825**	**0.9839**	**0.9825**	**0.9832**	**0.9832**
2012	**0.9658**	**0.9644**	**0.9661**	**0.9648**	**0.9651**	**0.9655**
2013	**0.9802**	**0.9811**	**0.9805**	**0.9797**	**0.9807**	**0.9801**
2014	**1.0243**	**1.0259**	**1.0274**	**1.0249**	**1.0251**	**1.0262**
2015	**0.9979**	**1.0066**	**1.0034**	**1.0009**	**1.0022**	**1.0022**
2016	**0.9746**	**0.9832**	**0.9931**	**0.9790**	**0.9789**	**0.9860**
2017	**1.0466**	**1.0355**	**1.0690**	**1.0481**	**1.0411**	**1.0585**

Note that the weighted indexes listed in Table A.4 are generally higher than their unweighted counterparts listed in Table A.3. The chained Laspeyres indexes are always above their chained Paasche counterparts but this is not always the case for the fixed base Laspeyres and Paasche indexes. Note also that the spread between the six weighted indexes listed in Table A.4 for 2017 is much smaller than the corresponding spread between the unweighted indexes in Table A.3: the highest index value was 1.0690 for the chained Laspeyres index and the lowest index value was 1.0355 for the fixed base Paasche index. Thus the index spread in 2017 was 1.0690/1.0355 = 1.032 or a 3.2 % spread which is far smaller than the unweighted index spread in 2017 which was 17.0 %.

Since the Paasche and Laspeyres indexes have equal justifications, we prefer the Fisher index which is an average of these two indexes which satisfies the time reversal test. To choose between the fixed base Fisher and its chained counterpart, we look at the spread between the Laspeyres and Paasche indexes in 2017. For the fixed base versions of these indexes, the spread is equal to $P_{LFB}^{2017}/P_{PFB}^{2017}$ = 1.0466/1.0355 = 1.011 or 1.1%. For the chained versions of these indexes the spread is equal to $P_{LCH}^{2017}/P_{PCH}^{2017}$ = 1.0690/1.0481 = 1.020 or 2.0%. Since the spread is smaller for the fixed base indexes, we prefer the fixed base indexes over the chained indexes and hence our preferred index for the present data set is the Fisher fixed base index, P_{FFB}^t.

For the *Approach 2 weighted indexes*, we treat the total access charges $v_5^t \equiv P^t$ as the aggregate price of access in year t [76] and we set the corresponding year t quantity, Q^t, equal to 1. The prices and quantities for products 1-4 are the p_n^t and q_n^t that are listed in Table A.1. The price of access, $P^t = v_5^t$, is listed in Table A.1. Denote the resulting year t fixed base Laspeyres, Paasche and Fisher indexes by P_{L2}^t, P_{P2}^t and P_{F2}^t and their chained counterparts by P_{LCH2}^t, P_{PCH2}^t and P_{FCH2}^t. These indexes are listed in Table A.5. We also list (one plus) the rate of growth in the access charges, P^t/P^{2010}, and (one plus) the rate of growth in expenditures on products 1-4, e^t/e^{2010}. Note that P^t/P^{2010} increases rapidly over time while e^t/e^{2010} decreases rapidly.

Table A.5: Approach 2 Laspeyres, Paasche, Fisher Indexes, P^t/P^{2010} and e^t/e^{2010}

Year t	P_{L2}^t	P_{P2}^t	P_{LCH2}^t	P_{PCH2}^t	P_{F2}^t	P_{FCH2}^t	P^t/P^{2010}	e^t/e^{2010}
2010	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
2011	**1.0112**	**1.0126**	**1.0112**	**1.0126**	**1.0119**	**1.0119**	**1.0356**	**0.8414**
2012	**1.0565**	**1.0671**	**1.0611**	**1.0658**	**1.0618**	**1.0634**	**1.1372**	**0.7397**
2013	**1.1051**	**1.1276**	**1.1137**	**1.1203**	**1.1163**	**1.1170**	**1.2163**	**0.6674**
2014	**1.1558**	**1.1877**	**1.1659**	**1.1722**	**1.1716**	**1.1691**	**1.2728**	**0.6063**
2015	**1.1943**	**1.2506**	**1.2198**	**1.2282**	**1.2221**	**1.2240**	**1.3691**	**0.5495**
2016	**1.2343**	**1.3185**	**1.2797**	**1.2870**	**1.2757**	**1.2834**	**1.4655**	**0.4843**
2017	**1.2996**	**1.3946**	**1.3419**	**1.3465**	**1.3463**	**1.3442**	**1.5247**	**0.4212**

[76] This is only an approximation to Model 2 defined by (A.3) since the UK data is aggregate retail sales data rather than individual household consumption data. Also Model 2 defined by (A.3) is a model that applies to a single household; we have neglected the complications that arise when aggregating over households.

Bringing access charges into the scope of the index has led to a general increase in the weighted index numbers. The fixed base Fisher index for Approach 1 ended up at 1.0481 in 2017 whereas the fixed base Fisher index for Approach 2 ended up at 1.3442. This is a very large difference. The fixed base Laspeyres index ended up at 1.2996 while the counterpart fixed base Paasche index ended up at 1.3946. The corresponding chained indexes ended up at 1.3419 and 1.3465. Thus for Approach 2, we prefer the chained Fisher index over its fixed base counterpart since the spread between the Laspeyres and Paasche indexes is much smaller for the chained indexes. However, the two Fisher indexes were very close to each other and they ended up at 1.3463 and 1.3442, so in this case, it does not matter which Fisher index is chosen.

Recall equations (A.9) which established the following relationship between the year t Approach 1 Laspeyres index, P_{L1}^t, and the Approach 2 Laspeyres index, P_{L2}^t: $P_{L1}^t - P_{L2}^t = [m^{2010}/(1 + m^{2010})][P_{L1}^t - (P^t/P^{2010})]$. From Tables A.4 and A.5, it can be seen that $P_{L1}^t < P^t/P^{2010}$ for all t > 2010 and thus $P_{L1}^t < P_{L2}^t$ for t = 2011,...,2017. Similarly, (A.12) established the following relationship between the year t Approach 1 Paasche index, P_{P1}^t, and the Approach 2 Paasche index, P_{P2}^t: $P_{P1}^t - P_{P2}^t = [m^t/(1 + m^t)]P_{P2}^t\,[P^t/P^{2010}]^{-1}\,[P_{P1}^t - (P^t/P^{2010})]$. From Tables A.4 and A.5, it can be seen that $P_{P1}^t < P^t/P^{2010}$ for all t > 2010 and thus $P_{P1}^t < P_{P2}^t$ for t = 2011,...,2017. These inequalities also imply that $P_{F1}^t < P_{F2}^t$ for t = 2011,...,2017. Thus due to the very rapid growth in access charges over the sample period, the Approach 2 Laspeyres, Paasche and Fisher indexes will be much larger than their Approach 1 counterparts.

For the *Approach 3 weighted indexes*, the access charges are spread across products 1-4 in a proportional manner. Thus define $1 + m^t \equiv v^t/e^t$ and $p_n^{t**} \equiv (1 + m^t)p_n^{t*}$ for n = 1,2,3,4 and t = 2010,...,2017. The corresponding quantities are the q_n^{t*} listed in Table A.2.[77] . Denote the Approach 3 year t fixed base Laspeyres, Paasche and Fisher indexes by P_{L3}^t, P_{P3}^t and P_{F3}^t and their chained counterparts by P_{LCH3}^t, P_{PCH3}^t and P_{FCH3}^t. These indexes are listed in Table A.6.

Table A.6: Approach 3 Laspeyres, Paasche and Fisher Indexes

Year t	P_{L3}^t	P_{P3}^t	P_{LCH3}^t	P_{PCH3}^t	P_{F3}^t	P_{FCH3}^t
2010	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2011	1.1040	1.1024	1.1040	1.1024	1.1032	1.1032
2012	1.2403	1.2385	1.2407	1.2391	1.2394	1.2399
2013	1.4068	1.4081	1.4072	1.4061	1.4074	1.4066
2014	1.6199	1.6225	1.6249	1.6209	1.6212	1.6229
2015	1.7854	1.8010	1.7953	1.7909	1.7932	1.7931
2016	2.0191	2.0369	2.0575	2.0282	2.0280	2.0428
2017	2.4972	2.4706	2.5506	2.5008	2.4839	2.5256

Allocating the access charges across the first four type of call products leads to a very large increase in the weighted index numbers. The fixed base Fisher indexes for Approach 1 and 2 ends up at 1.0481 and 1.3442 respectively in 2017 whereas the fixed base Fisher index for Approach 3 ends up at 2.4839. These differences are very large. The Approach 3 fixed base Laspeyres and Paasche spread in 2017 was smaller than the corresponding spread in their chained counterparts so the fixed base Fisher index P_{F3}^t is our preferred weighted index for this approach.

[77] Instead of using the $p_n^{t**} \equiv (1+m^t)p_n^{t*}$ and q_n^{t*} for n = 1,..,4 from Table A.2 as the primary data that is used in the various Laspeyres, Paasche and Fisher indexes, we could use $(1+m^t)p_n^t$ and q_n^t for n = 1,..,4 from Table A.1 as the primary data. The indexes remain the same since the Laspeyres, Paasche and Fisher indexes are invariant to changes in the units of measurement.

Using our current notation, the equalities in (A.15) translate into the following equalities:

(A.26) $P_{L3}{}^t/P_{L1}{}^t = P_{P3}{}^t/P_{P1}{}^t = (1 + m^t)/(1 + m^{2010})$; t = 2010,...,2017.

From Table A.2, we see that m^t is monotonically increasing. Thus using (A.26), it can be seen that the inequalities $P_{L3}{}^t > P_{L1}{}^t$ and $P_{P3}{}^t > P_{P1}{}^t$ for t > 2010 must hold.

Using our current notation, (A.16) can be rewritten as follows:

(A.27) $P_{L3}{}^t - P_{L2}{}^t = m^t[1 + m^{2010}]^{-1}[P_{L1}{}^t - (e^t/e^{2010})]$; t = 2010,...,2017.

Tables A.4 and A.5 list the usage expenditure ratios (e^t/e^{2010}) and the Approach 1 Laspeyres indexes $P_{L1}{}^t$. Using these series, it can be seen that $P_{L1}{}^t > e^t/e^{2010}$ for t > 2010. Thus using (A.27), we must have $P_{L3}{}^t > P_{L2}{}^t$ for t > 2010.

Using our current notation, (A.19) can be rewritten as follows:

(A.28) $P_{P3}{}^t - P_{P2}{}^t = m^{2010}[1 + m^{2010}]^{-1} P_{P2}{}^t[P_{P1}{}^t - (e^t/e^{2010})]$; t = 2010,...,2017.

Tables A.4 and A.5 list the usage expenditure ratios (e^t/e^{2010}) and the Approach 1 Paasche indexes $P_{P1}{}^t$ and it can be seen that $P_{P1}{}^t > e^t/e^{2010}$ for t > 2010. Thus using (A.28), we must have $P_{P3}{}^t > P_{P2}{}^t$ for t > 2010. It follows that it is also the case that $P_{F3}{}^t > P_{F2}{}^t$ for t > 2010.

Finally, we consider *Approach 4*. This approach is an approach that is used when constructing producer price indexes for the telecom sector in the regulation literature that attempts to measure the Total Factor Productivity of the sector. In this approach, the number of line connections is used as the output measure for access charges.[78] Thus this approach simply uses the $v_n{}^t$ and $q_n{}^t$ that are listed in Table A.1 (and the implied prices $p_n{}^t \equiv v_n{}^t/q_n{}^t$ for n = 1,...,5) in the usual index number formulae that are considered in this Appendix. Denote the Approach 4 year t fixed base Laspeyres, Paasche and Fisher indexes by $P_{L4}{}^t$, $P_{P4}{}^t$ and $P_{F4}{}^t$ and their chained counterparts by $P_{LCH4}{}^t$, $P_{PCH4}{}^t$ and $P_{FCH4}{}^t$. These indexes are listed in Table A.7.

Table A.7: Approach 4 Laspeyres, Paasche and Fisher Indexes

Year t	$P_{L4}{}^t$	$P_{P4}{}^t$	$P_{LCH4}{}^t$	$P_{PCH4}{}^t$	$P_{F4}{}^t$	$P_{FCH4}{}^t$
2010	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2011	1.0085	1.0097	1.0085	1.0097	1.0091	1.0091
2012	1.0390	1.0484	1.0427	1.0470	1.0437	1.0449
2013	1.0737	1.0927	1.0800	1.0857	1.0832	1.0829
2014	1.1084	1.1316	1.1135	1.1178	1.1199	1.1156
2015	1.1298	1.1733	1.1479	1.1541	1.1513	1.1510
2016	1.1544	1.2209	1.1904	1.1953	1.1872	1.1929
2017	1.2115	1.2796	1.2419	1.2438	1.2451	1.2428

[78] See for example Lawrence and Diewert (2006; 218) where the distributor's number of line connections is regarded as an output of the firm. Their paper is concerned with electricity distribution but the same methodology is used for telecommunication firms.

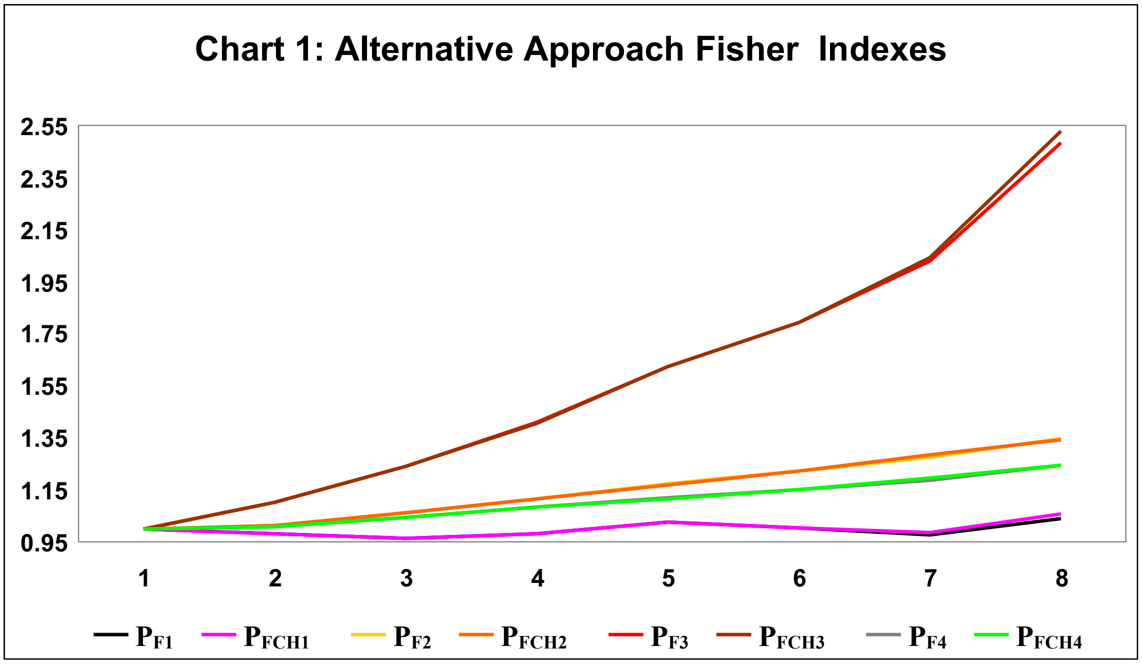
Using the Approach 4 methodology, it can be seen that the fixed base Paasche index grows more rapidly than the corresponding fixed base Laspeyres index. The addition of product 5 to the first four products has caused this somewhat unusual phenomenon. The price of product 5 increases 1.36 fold over the sample period which is much higher than a weighted average of the prices of the first 4 products; i.e., P_{L1}^t and P_{P1}^t increased 1.047 fold and 1.036 fold respectively over the sample period. At the same time, q_5^t increased while q_1^t-q_4^t decreased substantially over the sample period. Under these conditions, P_{P4}^t will increase more rapidly than P_{L4}^t. Table A.7 also indicates that the spread between P_{L4}^{2017} and P_{P4}^{2017} is larger than the spread between the chained indexes, P_{L4CH}^{2017} and P_{P4CH}^{2017}. Under these conditions, we prefer the chained Fisher index P_{FCH4}^t over its fixed base counterpart P_{F4}^t. However, Table A.7 indicates that the difference between the fixed base and chained Fisher indexes is negligible using Approach 4.

The following table lists the fixed base and chained Fisher indexes for all four approaches.

Table A.8: Fixed Base and Chained Fisher Indexes for All Four Approaches

Year t	P_{F1}^t	P_{FCH1}^t	P_{F2}^t	P_{FCH2}^t	P_{F3}^t	P_{FCH3}^t	P_{F4}^t	P_{FCH4}^t
2010	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2011	0.9832	0.9832	1.0119	1.0119	1.1032	1.1032	1.0091	1.0091
2012	0.9651	0.9655	1.0618	1.0634	1.2394	1.2399	1.0437	1.0449
2013	0.9807	0.9801	1.1163	1.1170	1.4074	1.4066	1.0832	1.0829
2014	1.0251	1.0262	1.1716	1.1691	1.6212	1.6229	1.1199	1.1156
2015	1.0022	1.0022	1.2221	1.2240	1.7932	1.7931	1.1513	1.1510
2016	0.9789	0.9860	1.2757	1.2834	2.0280	2.0428	1.1872	1.1929
2017	1.0411	1.0585	1.3463	1.3442	2.4839	2.5256	1.2451	1.2428

From Table A.8, it can be seen that the Approach 1 Fisher indexes (which ignored the access charges) generate the lowest increase in prices, followed by the Approach 4 indexes (include access charges as a regular commodity with the quantity set equal to the number of lines), followed by the Approach 2 indexes (include access charges but hold the corresponding quantity fixed at unity) and finally followed by the Approach 3 Fisher indexes (which spread the access charges across the other products). These alternative approach Fisher indexes are plotted in Chart 1.



It can be seen that the differences between fixed base and chained Fisher indexes for each approach are small but the differences between the four approaches is very large indeed. Thus in the case of fixed line telecommunications services, the choice of an Approach to the treatment of access charges is important.

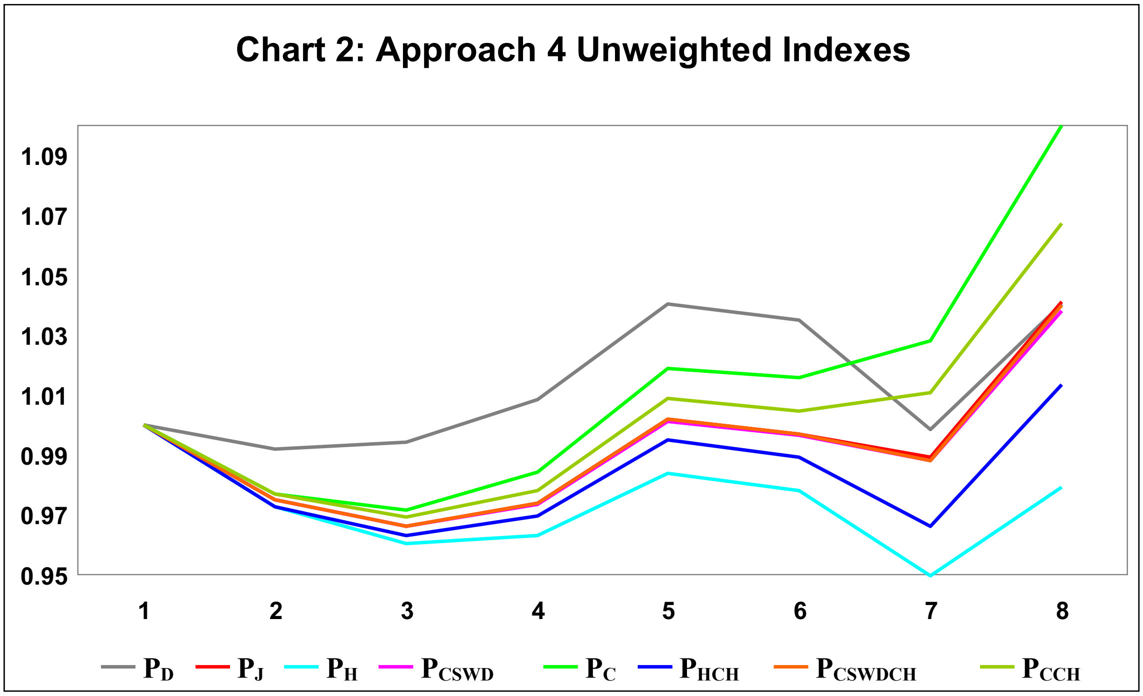
In the case where quantity or expenditure weights are not available, the choice of an elementary index number formula is also important for the telecommunications sector; recall Table A.3 above which listed the unweighted indexes using the prices of products 1-4. To conclude this appendix, we list the same unweighted indexes as were listed in Table A.3 but using the prices of products 1-5 in Table A.9 below.

Table A.9: Approach 4 Unweighted Price Indexes

Year t	P_D^t	P_J^t	P_H^t	P_{CSWD}^t	P_C^t	P_{HCH}^t	P_{CSWDCH}^t	P_{CCH}^t
2010	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2011	0.9920	0.9750	0.9728	0.9749	0.9770	0.9728	0.9749	0.9770
2012	0.9941	0.9662	0.9605	0.9661	0.9717	0.9629	0.9662	0.9694
2013	1.0083	0.9739	0.9629	0.9734	0.9841	0.9697	0.9738	0.9780
2014	1.0403	1.0019	0.9838	1.0012	1.0188	0.9950	1.0019	1.0088
2015	1.0349	0.9970	0.9779	0.9967	1.0158	0.9891	0.9970	1.0048
2016	0.9986	0.9892	0.9498	0.9882	1.0280	0.9660	0.9882	1.0108
2017	1.0403	1.0412	0.9792	1.0379	1.1001	1.0133	1.0400	1.0674

The 2017 spread in the above unweighted indexes is 1.1001/0.9792 = 1.123 or 12.3%. Recall that the corresponding index spread for the Approach 1 unweighted price indexes was 17.0% so the addition of product 5 has lowered the spread significantly. The above indexes used the prices that correspond to the values and quantities listed in Table A.1. Recall that the Dutot index using normalized prices, P_{DN}^t, was equal to the chained Carli index, P_{CCH}^t, listed in Table A.3. A similar result holds here: P_{DN}^t is equal to P_{CCH}^t listed in Table A.8. The indexes listed in Table A.8 are plotted on Chart 2. It can be seen that P_J^t, P_{CSWD}^t and P_{CSWDCH}^t cannot be distinguished on Chart

2.[79] These series are in the middle of the listed indexes, with the chained Carli and Carli indexes, $P_{CCH}{}^t$ and $P_C{}^t$, well above the middle series and the chained Harmonic and Harmonic indexes, $P_{HCH}{}^t$ and $P_H{}^t$, well below the middle series. The Dutot series $P_D{}^t$ is initially well above the other series but it joins up with the middle series at the end of the sample period. The Dutot index $P_{DN}{}^t$ using the normalized prices listed in Table A.2 coincides with the fixed base Carli index $P_C{}^t$. Thus there is a substantial difference in the Dutot indexes as the units of measurement change. The remaining indexes are invariant to changes in the units of measurement.



Some of the conclusions that can be drawn from this appendix are as follows:

- Unweighted elementary indexes can differ substantially depending on which formula is used.
- The Carli fixed base and chained indexes are not recommended due to their failure of the time reversal test with a built in upward bias.
- The Dutot index is also not recommended due to its lack of invariance to changes in the units of measurement. Even if the units of measurement are the same, the empirical example shows that changing the units of measurement can make a huge difference.
- The approach used to allocate access charges can make a substantial difference to the CPI in the case of regulated network industries where access charges can be substantial.[80]
- In the case of regulated industries, often price and quantity data will often be available to the price statistician.[81] In this case, weighted indexes are preferred over unweighted

[79] Recall the approximate equalities (27) and (29) in section 5 above.
[80] This point is due to Abdirahman, Coyle, Heys and Stewart (2017) (2020).
[81] Unfortunately, data submitted to regulators is usually quarterly data which presents challenges in the context of producing a monthly CPI. However, national income accountants have to produce quarterly consumer price indexes and perhaps more importantly, national accounts price indexes can be revised. Hence as better information becomes available to the price statistician, better (revised) indexes can be produced.

indexes because they take into account the economic importance of the various outputs of the regulated industry. The example in this appendix shows that there can be significant differences between weighted and unweighted indexes.

Appendix B: Additional Problems Associated with the use of the Carli Index

Robert Hill (2018) submitted some testimony to the United Kingdom's House of Lords Economic Affairs Committee on the use of the Carli Index in the UK's Retail Price Index. His points 3-5 listed below deal with problems associated with the use of the Carli index. Since his testimony is not easily accessible and some of his points were not made in this chapter, the first five points in his testimony are quoted below.

1. I am responding to the latest call for evidence from the House of Lords Economic Affairs Committee in my capacity as a researcher in the field of price indices. I am a British citizen based at University of Graz in Austria, where I am Professor of Macroeconomics. I served on the Expert Advisory Group for Paul Johnson's report on UK Consumer Prices Statistics. I have also served as an advisor to Eurostat on the treatment of owner-occupied housing (OOH) in the harmonized index of consumer prices (HICP).
2. In this statement I will focus on what I think are the two most serious problems with the RPI. These are its use of the Carli formula at the elementary level, and its treatment of OOH.
3. Irving Fisher warned against using the Carli formula in his 1922 book on index numbers. Carli fails the time reversal test, and suffers from a systematic upward bias. For example, if prices change from periods 1 to 2, but then in period 3 return to their original period 1 levels, a chained Carli index will always find that the price level is higher in period 3 than in period 1 (except in the special case where all prices change by exactly the same proportion from one period to the next).
4. Levell (2015) provides a detailed comparison of the Carli and Jevons price index formulas. Carli takes an arithmetic mean of the price relatives while Jevons takes a geometric mean. While Levell ends up rightly favouring Jevons, he is at times too kind to Carli, which could cause some confusion among users. Indeed there seems to be a perception in some circles that there are trade-offs between Carli and Jevons. For example, Leyland (2011) states that: "*The RSS does not have a view on whether the arithmetic or geometric mean is the better approach but it does consider the issue a major concern*."
5. In my opinion the use of the Carli index is indefensible. To see why, I will revisit some of the points made by Levell. Levell assesses the Carli index from three perspectives, referred to in the literature as the test, statistical and economic approaches. From the test perspective, Jevons is unambiguously better than Carli. Jevons is the only elementary price index formula that satisfies all the 14 tests considered by Levell. Up to this point I am in complete agreement with Levell. Turning to the statistical approach on page 316, Levell states that: "*Ultimately our object of interest here is* $E(p_1^i/p_0^i)$." He then goes on to show that Jevons is a downward biased measure of $E(p_1^i/p_0^i)$. My problem here is that I disagree that $E(p_1^i/p_0^i)$ should be our object of interest since it treats price rises and falls asymmetrically. A better approach is to focus on the natural logarithm of the price indices with the following object of interest: $E[\ln(p_1^i/p_0^i)]$. In this setting Jevons unambiguously outperforms Carli under the statistical approach. Turning finally to the economic approach, Levell notes that in the case of Leontief preferences – where there is no substitution effect – a case can be made for Carli. This argument dates back at least to the ILO CPI Manual of 2004, which on page 16 contains the following statement: "*With Leontief preferences, a Laspeyres index provides an exact measure of the cost of living index. In this case, the Carli calculated for a random sample would provide an estimate of the cost of living provided that the items were selected with probabilities proportional to the population expenditure shares.*" This statement has caused a lot of confusion in the literature. I agree that in this case Laspeyres is an exact measure. But what follows regarding Carli is misleading. First, the whole point with elementary indices is that there are no expenditure shares. Second, if we assume the items are sampled proportionally to expenditure shares, then what we have is not Carli but a weighted arithmetic mean of the price relatives. If we assume further that the reference expenditure shares are those of the earlier of the two periods being compared, then instead of Carli we have Laspeyres. So what this statement is really saying is that if we have Leontief preferences and we replace Carli with Laspeyres, then we will get the right answer. This is not very helpful. It is not true that Carli performs well when preferences

are Leontief or close to Leontief. The only situation when Carli is free of upward bias is when all prices change at the same rate (which is the Hicks, not Leontief, aggregation case). In conclusion, whichever way you look at it the Carli index is flawed and should not be used. Jevons has much better properties. Robert Hill (2018).

Hill makes two important points in his point 5 above:

- The econometric or statistical approach to index number theory frequently assumes that the goal of the exercise is to measure the average relative price increase; i.e., to measure some average over n of the price ratios, p_{tn}/p_{1n}. Using this perspective, econometricians may assert that for example, the Törnqvist Theil index is a biased estimator for the target index. But this "bias" vanishes if we make the goal the measurement of the average $\log(p_{tn}/p_{1n})$. As Hill notes above, the first approach treats price rises and falls more asymmetrically than the second approach. In any case, the more important economic and basket approaches to consumer index number theory do not take the statistical approach to index number theory.
- Hill's second main point has to do with justifications for the use of the Carli index under special assumptions about the nature of consumer preferences. His dismissal of this type of argument seems to be on target.

References

Abdirahman, M., D. Coyle, R. Heys and W. Stewart (2017), "A Comparison of Approaches to Deflating Telecoms Services Output", Economic Statistics Centre of Excellence (ESCoE) Discussion Paper 2017-04, London.

Abdirahman, M., D. Coyle, R. Heys and W. Stewart (2020), "Telecoms Deflators: A Story of Volume and Revenue Weights", Economic Statistics Centre of Excellence (ESCoE) Discussion Paper 2020-11, London.

Balk, B.M. (1980), "A Method for Constructing Price Indices for Seasonal Commodities", *Journal of the Royal Statistical Society A* 143, 68-75.

Balk, B.M. (1994), "On the First Step in the Calculation of a Consumer Price Index", paper presented at the first Ottawa Group Meeting, Ottawa, October 31-November 4.

Balk, B.M. (2000), "On Curing the CPI's Substitution and New Goods Bias", Research Paper 0005, Department of Statistical Methods, Statistics Netherlands, Voorburg.

Balk, B.M. (2002), "Price Indexes for Elementary Aggregates: The Sampling Approach", Research Report, Voorburg: Statistics Netherlands.

Balk, B.M. (2008), *Price and Quantity Index Numbers*, New York: Cambridge University Press.

Carli, Gian-Rinaldo, (1804), "Del valore e della proporzione de' metalli monetati", pp. 297-366 in *Scrittori classici italiani di economia politica*, Volume 13, Milano: G.G. Destefanis (originally published in 1764).

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

Coggeshall, F. (1887), “The Arithmetic, Geometric and Harmonic Means”, *Quarterly Journal of Economics* 1, 83-86.

Dalén, J. (1992), “Computing Elementary Aggregates in the Swedish Consumer Price Index,” *Journal of Official Statistics* 8, 129-147.

Dalén, J. (1994), “Sensitivity Analyses for Harmonizing European Consumer Price Iindexes”, pp. 147-171 in *International Conference on Price Iindexes: Papers and Final Report*, First Meeting of the International Working Group on Price Iindexes, November, Ottawa: Statistics Canada.

Davies, G.R. (1924), “The Problem of a Standard Index Number Formula”, *Journal of the American Statistical Association* 19, 180-188.

Davies, G.R. (1932), “Index Numbers in Mathematical Economics”, *Journal of the American Statistical Association* 27, 58-64.

de Haan, J. and F. Krsinich (2014), “Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes,” *Journal of Business and Economic Statistics* 32, 341–358.

de Haan, J. and H. van der Grient (2011), “Eliminating Chain Drift in Price Indexes Based on Scanner Data”, *Journal of Econometrics* 161, 36-46.

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

Diewert, W.E. (1993), “The Early History of Price Index Research”, pp. 33-65 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1995), “Axiomatic and Economic Approaches to Elementary Price Indexes”, Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (1998), “Index Number Issues in the Consumer Price Index”, *Journal of Economic Perspectives* 12:1, 47-58.

Diewert, W.E. (2002), “Harmonized Indexes of Consumer Prices: Their Conceptual Foundations”, *Swiss Journal of Economics and Statistics* 138:4, 547-637.

Diewert, W.E. (2013), “Answers to Questions Arising from the RPI Consultation”, Discussion Paper 13-04, Vancouver School of Economics, University of British Columbia, Vancouver, Canada, V6T 1L4.

Diewert, W.E., J. de Haan and K.J. Fox (2016), “A Newly Identified Source of Potential CPI Bias: Weekly versus Monthly Unit Value Price Indexes”, *Economics Letters* 141, 169-172.

Diewert, W.E. and P. von der Lippe (2010), “Notes on Unit Value Index Bias”, *Journal of Economics and Statistics* 230, 690-708.

Dutot, Charles, (1738), *Réflexions politiques sur les finances et le commerce*, Volume 1, La Haye: Les frères Vaillant et N. Prevost.

Eichhorn, W. (1978), *Functional Equations in Economics*, London: Addison-Wesley.

Evans, B. (2012), "International Comparison of the Formula Effect between the CPI and the RPI", Office for National Statistics (ONS), Newport, UK.

Feenstra, R.C. (1994), "New Product Varieties and the Measurement of International Prices", *American Economic Review* 34, 157-177.

Feenstra, R.C. and M.D. Shapiro (2003), "High Frequency Substitution and the Measurement of Price Indexes", pp. 123-146 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Fisher, I. (1921), "The Best Form of Index Number", *Quarterly Publication of the American Statistical Association* 17, 533-537.

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

Fox, K.J. and I. A Syed (2016), "Price Discounts and the Measurement of Inflation", *Journal of Econometrics* 191, 398-406.

Gudnason, R. (2003), "How do we Measure Inflation? Some Measurement Problems", 7th Ottawa Group Meeting on Price Indices, May 27-29, Paris.

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

Hawkes, W.J. (1997), "Reconciliation of Consumer Price Index Trends in Average Prices for Quasi-Homogeneous Goods Using Scanning Data", Third Meeting of the International Working Group on Price Indices, Statistics Netherlands, Voorburg, April 16-18.

Hawkes, W.J. and F.W. Piotrowski (2003), "Using Scanner Data to Improve the Quality of Measurement in the Consumer Price Index", pp. 17-38 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Hill, R. (2018), Statement Submitted to the House of Lords Economic Affairs Committee Inquiry into the Use of the Retail Price Index, July 21, 2018.

ILO, IMF, OECD, UNECE, Eurostat and World Bank (2004), *Consumer Price Index Manual: Concepts and Practice 2020*, Peter Hill (ed.), Geneva: International Labour Office.

IMF, ILO, Eurostat, UNECE, OECD and World Bank (2020), *Consumer Price Index Manual: Theory and Practice*, Brian Graf (ed.), Washington D.C.: International Monetary Fund.

Ivancic, L., W.E. Diewert and K.J. Fox (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes", *Journal of Econometrics* 161, 24-35.

Ivancic, L. and K.J. Fox (2013), "Understanding Price Variation Across Stores and Supermarket Chains: Some Implications for CPI Aggregation Methods", *Review of Income and Wealth* 59:4, 629-647.

Jevons, W.S., (1865), "The Variation of Prices and the Value of the Currency since 1782", *Journal of the Statistical Society of London* 28, 294-320; reprinted in *Investigations in Currency and Finance* (1884), London: Macmillan and Co., 119-150.

Koskimäki, T. and M. Ylä-Jarkko (2003), "Segmented Markets and CPI Elementary Classifications", 7th Ottawa Group Meeting on Price Indices, Paris, May 27-29.

Koskimäki, T. and Y. Vartia (2001), "Beyond Matched Pairs and Griliches Type Hedonic Methods for Controlling Quality Changes in CPI Subindexes", 6th Ottawa Group Meeting on Price Indices, Canberra, April 2-6.

Lawrence, D. and W.E. Diewert (2006), "Regulating Electricity Networks: The ABC of Setting X in New Zealand", pp. 207-241 in *Performance Measurement and Regulation of Network Utilities*, T. Coelli and D. Lawrence (eds.), Cheltenham: Edward Elgar Publishing.

Levell, P. (2015), "Is the Carli Index Flawed?: Assessing the Case for the New Retail Price Index RPIJ", Journal of the Royal Statistical Society Series A, 178(Part 2), 303-336.

Ley, E. (2003), "Comment", pp. 379-382 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Leyland, J. (2011), "RPI versus CPI—the Definitive Account", *Significance Magazine*. Web address:
https://www.significancemagazine.com/business/159-rpi-versus-cpi-the-definitive-account

Reinsdorf, M. (1994), "The Effect of Price Dispersion on Cost of Living Indexes", *International Economic Review* 35, 137-149.

Reinsdorf, M. (1996), "Constructing Basic Component Indexes for the U.S. CPI from Scanner Data: A Test Using Data on Coffee", BLS Working Paper 277, Bureau of Labor Statistics, Washington D.C., April.

Reinsdorf, M. and B.R. Moulton (1997), "The Construction of Basic Components of Cost of Living Indexes", pp. 397-436 in *The Economics of New Goods*, T. F. Bresnahan and R.J. Gordon (eds.), University of Chicago Press, Chicago.

Reinsdorf, M. and J. Triplett (2009), "A Review of Reviews: Ninety Years of Professional Thinking About the Consumer Price Index", pp. 17-84 in *Price Index Concepts and Measurement*, W.E. Diewert, J. Greenlees and C. Hulten (eds.), Studies in Income and Wealth, Volume 70, Chicago: University of Chicago Press.

Richardson, D.H. (2003), "Scanner Indexes for the Consumer Price Index", pp. 39-65 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Silver, M. (1995), "Elementary Aggregates, Micro-Indices and Scanner Data: Some Issues in the Compilation of Consumer Price Indices", *Review of Income and Wealth* 41, 427-438.

Silver, M. (2010), "The Wrong and Rights of Unit Value Indices", *Review of Income and Wealth* 56:S1, 206-206.

Silver, M. (2011), "An Index Number Formula Problem: the Aggregation of Broadly Comparable Items", *Journal of Official Statistics* 27:4, 1–17.

Silver, M. and S. Heravi (2001), "Scanner Data and the Measurement of Inflation", *The Economic Journal,* 111, F384-F405.

Silver, M. and S. Heravi (2003), "The Measurement of Quality Adjusted Price Changes", pp. 277-316 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Silver, M. and S. Heravi (2005), "A Failure in the Measurement of Inflation: Results from a Hedonic and Matched Experiment using Scanner Data", *Journal of Business and Economic Statistics* 23:3, 269–281.

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

Szulc, B.J. (1987), "Price Iindexes below the Basic Aggregation Level", *Bulletin of Labour Statistics* 2, 9-16.

Theil, H. (1954), *Linear Aggregation of Economic Relations*, Amsterdam: North-Holland.

Triplett, J.E. (2003), "Using Scanner Data in Consumer Price Indexes: Some Neglected Conceptual Considerations", pp. 151-162 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Triplett, J.E. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Directorate for Science, Technology and Industry, DSTI/DOC(2004)9, Paris: OECD.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

Walsh, C.M. (1921a), *The Problem of Estimation*, London: P.S. King & Son.

Walsh, C.M. (1921b), "Discussion", *Journal of the American Statistical Association* 17, 537-544.

White, A.G. (1999), "Measurement Biases in Consumer Price Indexes", *International Statistical Review* 67:3, 301-325.

White, A.G. (2000), "Outlet Types and the Canadian Consumer Price Index", *Canadian Journal of Economics* 33, 488-505.