The Ottawa Group After 30 Years

Erwin Diewert
Vancouver School of Economics
University of British Columbia
6000 Iona Drive
Vancouver B.C.
Canada V6T 1L4
Revised: June 10, 2024.
Email: erwin.diewert@ubc.ca

18th Meeting of the Ottawa Group on Price Indices at Ottawa, Canada, May 13-15, 2024.

Background Paper for the *Special Session: Panel Discussion on the State of Research and Practice of CPI Methodology and Challenges Ahead*. The author thanks Chihiro Shimizu for helpful comments.

1. Introduction

The Ottawa Group on Price Indices (also known as the United Nations International Working Group on Price Indices) had its first meeting in Ottawa on October 31- November 2, 1994. The present meeting is the 18th meeting of the Group and it is also taking place in Ottawa. Here is what the Ottawa Group website has to say about the purpose of the Group:

"The focus of the Group is on applied research, particularly in the area of consumer price indices. The Group examines advantages and disadvantages of various concepts, methods and procedures in the context of realistic operational environments, supported by concrete examples whenever possible. Participants are specialists and practitioners who work for, or are advisers to, statistical agencies in different countries or international organisations. The Group meets about every two years."

The Ottawa Group website has also posted almost all of the papers (540) and presentation slides that were presented at the 18 meetings. This is an extremely valuable source of materials on the theory and practice of Consumer Price Index construction.[1] Over the years, Statistics Canada and the Australian Bureau of Statistics have maintained this website.[2]

How did the Ottawa Group start? It seems that Paul Armknecht (Bureau of Labor Statistics), Bert Balk (Statistics Netherlands) and Bohdan Schultz (Statistics Canada) were having a beer after a meeting of the Joint UNECE/ILO group on Consumer Price Indices which meets in Geneva every two years.[3] They collectively thought that it would be useful to have more theoretical meetings on CPI theory and practice. Bohdan (who was Head of Consumer Price Index Research at Statistics Canada) went back to his boss, Jacob Ryten, and suggested that Statistics Canada convene a meeting in Ottawa that would discuss CPI problems. Jacob agreed and thus the first meeting of this new group was convened in Ottawa in 1994. I was fortunate enough to be invited to this meeting because I had an interest in index number theory and I was also Chair of the Statistics Canada Advisory Group on Price Statistics at the time (and I am still Chair).[4]

The first meeting of the Ottawa Group will be discussed in some detail in section 4 below. However, it is useful to set the stage for the discussion of this first meeting by reviewing alternative approaches to index number theory in section 2 that were available to Statistical Offices in 1994 and then to review the 1989 ILO Manual on the CPI[5] in section 3, which outlined existing "practical" advice to CPI compilers.

The Ottawa Group participants were very influential in producing the 2004 *Consumer Price Index Manual*. Section 5 will discuss this Manual and note some problems with the advice in this document.

Section 6 discusses the current CPI Manuals.

---

[1] The Ottawa Group also deals with producer price indexes but the main focus has been on consumer price indexes.
[2] Maintenance of the website has been turned over to Carsten Boldsen at the UNECE in Geneva.
[3] The origins of this Group was explained by Stoevska (2018) as follows: "In order to exchange views and experiences in the field of price statistics and to identify the best practices, at an initiative led by the International Labour Office (ILO), the first meeting on CPI was jointly organised by UNECE and ILO in 1978."
[4] In addition to the Ottawa Group and the UNECE/ILO Group, other groups which focus on economic measurement problems are the International Association for Research in Income and Wealth, the US NBER-Conference for Research in Income and Wealth, the NBER (National Bureau for Economic Research) Prices and Productivity Program which was started by Zvi Griliches in 1979 (see Berndt (2001), the Society for Economic Measurement (William Barnett), the Economic Measurement Group (Kevin Fox) and the Economic Measurement Group Asia (Chihiro Shimizu).
[5] See Turvey (1989).

Section 7 concludes by discussing some outstanding problems and looks at possible future developments.

Appendix A explains a consumer demand approach to the construction of practical price indexes that can deal with the chain drift problem as well as the problem of quality adjustment. This Appendix also shows that it is not possible to quality adjust the price of a new product in the period when it first appears in the marketplace unless purchaser utility functions are linear in quantities.

2. Approaches to Index Number Theory in 1994

In 1994, there were four main approaches to bilateral index number theory that could be used by price statisticians if price and quantity data on the same products were available for two periods:

- Basket Approaches;
- Stochastic Approaches;
- Test or Axiomatic Approaches and
- Economic Approaches.

We will describe briefly each approach.[6]

Basket Approaches

The fixed basket approach to the construction of a consumer price index is one of the earliest approaches[7] to bilateral index number theory.[8] Lowe (1823; 333-335) was not the first to propose the fixed basket approach, but he laid out the construction of a fixed basket index (and its use) in some detail. Basically, he proposed that the government should form a list of articles of general consumption, which is a representative annual basket $q \equiv (q_1,\ldots,q_N)$, and if the prices of year t were $p^t \equiv (p_{t1},\ldots,p_{tN})$ and the prices of the base year 0 were $p^0 \equiv (p_{01},\ldots,p_{0N})$, then the *Lowe price index* which gives the aggregate level of prices in year t relative to the level of prices in year 1 is $P_{Lo} \equiv p^t{\cdot}q/p^0{\cdot}q$ where $p^t{\cdot}q$ is defined as the sum $\Sigma_{n=1}^N p_{tn}q_n$. Basically, we compare the cost of the same fixed basket of goods and services for the two years in question and take our price index as the ratio of the resulting costs.[9]

As time passed, economists and price statisticians demanded a bit more precision with respect to the specification of the basket vector q. There are two natural choices for the reference basket: the period 0 commodity vector q^0 or the period 1 commodity vector q^1. These two choices lead to the *Laspeyres* (1871) *price index* $P_L \equiv p^t{\cdot}q^0/p^0{\cdot}q^0$ and the *Paasche* (1874) *price index* $P_P \equiv p^t{\cdot}q^t/p^0{\cdot}q^t$.

The problem with the Laspeyres and Paasche index number formulae is that they are equally plausible but in general, they will give different answers. This suggests that if we require a single estimate for the price change between the two periods, then we should take some sort of evenly weighted average of the two

---

[6] For a much more comprehensive description and history of the various approaches to index number theory, see Chapter 1 of Balk (2008).
[7] See Diewert (1993a) and Balk (2008) for the early history of index number theory.
[8] Bilateral index number theory compares price levels in two periods. Multilateral index number theory compares price levels across more than two periods.
[9] Lowe (1823; Appendix page 95) suggested that the commodity basket vector q should be updated every five years.

indexes as our final estimate of price change between periods 0 and 1. Examples of such symmetric averages are the arithmetic mean, which leads to the *Sidgwick* (1883; 68) *Bowley* (1901; 227) *index*[10] , $(1/2)P_L + (1/2)P_P$, and the geometric mean, which leads to the *Fisher* (1922) *ideal index*, P_F, defined as $[P_L P_P]^{1/2}$.

The problem with the arithmetic mean of the Paasche and Laspeyres indexes is that the resulting index does not satisfy the time reversal test. Thus if $P(p^0,p^t,q^0,q^t)$ denotes a generic index number formula between periods 0 and t going forward from period 0, then we would like the formula to satisfy $P(p^t,p^0,q^t,q^0) = 1/P(p^0,p^t,q^0,q^t)$; i.e., if we measure price change going backwards from period t to period 0, then we would like the resulting index, $P(p^t,p^0,q^t,q^0)$, to equal the reciprocal of the index that goes forward from period 0 to t, $P(p^0,p^t,q^0,q^t)$. However, the Fisher index does satisfy the time reversal test. Moreover, it is the *only* index that is a homogeneous symmetric average of the Laspeyres and Paasche price indexes, P_L and P_P, and satisfies the time reversal test.[11]

Thus the basket approach to consumer price indexes leads to the Fisher ideal index as a possible "best" choice of bilateral index number formula.

Stochastic Approaches

The stochastic approach to the determination of the price index can be traced back to the work of Jevons (1865) (1884) and Edgeworth (1888) (1923) over a hundred years ago.

The basic idea behind the unweighted (or more accurately, the evenly weighted) stochastic approach[12] is that each price ratio or price relative, p_{tn}/p_{0n} for n = 1,2,…,N, can be regarded as an estimate of a common inflation rate α between periods 0 and t; i.e., it is assumed that

(1) $p_{tn}/p_{0n} = \alpha + \varepsilon_n$; n = 1,2,…,N

where α is the common inflation rate and the ε_n are random variables with mean 0 and variance σ^2. The least squares estimator for α is the *Carli* (1764) *price index* P_C defined as

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

Unfortunately, P_C does not satisfy the Time Reversal Test, i.e., $P_C(p^t,p^0) \neq 1/ P_C(p^0,p^t)$[13].

Now assume that the logarithm of each price relative, $\ln(p_{tn}/p_{0n})$, is an unbiased estimate of the logarithm of the inflation rate between periods 0 and t, β say. Thus we have:

(3) $\ln(p_{tn}/p_{0n}) = \beta + \varepsilon_n$; n = 1,2,…,N

---

[10] Bowley (1899; 641) appears to have been the first to suggest the use of this index.

[11] See Diewert (1997; 138).

[12] Frisch (1936; 3) called the stochastic approach the atomistic approach.

[13] In fact Fisher (1922; 66) noted that $P_C(p^0,p^1)P_C(p^1,p^0) \geq 1$ unless the period 1 price vector p^1 is proportional to the period 0 price vector p^0; i.e., Fisher showed that the Carli index has a definite upward bias. He urged statistical agencies not to use this formula.

where $\beta \equiv \ln\alpha$ and the ε_n are independently distributed random variables with mean 0 and variance σ^2. The least squares or maximum likelihood estimator for β is the logarithm of the geometric mean of the price relatives. Hence the corresponding estimate for the common inflation rate α is the *Jevons* (1865) *price index* P_J defined as:

$$(4)\ P_J(p^0,p^t) \equiv \prod_{n=1}^{N} (p_{tn}/p_{0n})^{1/N}.$$

The Jevons price index P_J does satisfy the Time Reversal Test and hence is much more satisfactory than the Carli index P_C. However, both the Jevons and Carli price indexes suffer from an important flaw: each price relative p_{tn}/p_{0n} is regarded as being equally important and is given an equal weight in the index number formulae (2) and (4).[14]

Theil (1967; 136-137) proposed a solution to the lack of weighting in (2) and (4). He argued as follows. Suppose we draw price relatives at random in such a way that each dollar of expenditure in the base period has an equal chance of being selected. Then the probability that we will draw the nth price relative is equal to $s_{0n} \equiv p_{0n}q_{0n}/p^0\cdot q^0$, the period 0 expenditure share for commodity n. Thus the overall mean (period 0 weighted) logarithmic price change is $\sum_{n=1}^{N} s_{0n} \ln(p_{tn}/p_{0n})$. Now repeat the above mental experiment and draw price relatives at random in such a way that each dollar of expenditure in period t has an equal probability of being selected. This leads to the overall mean (period t weighted) logarithmic price change of $\sum_{n=1}^{N} s_{tn} \ln(p_{tn}/p_{0n})$. Each of these measures of overall logarithmic price change seems equally valid so we could argue for taking a symmetric average of the two measures in order to obtain a final single measure of overall logarithmic price change. Theil argued that a nice symmetric index number formula can be obtained if we make the probability of selection for the nth price relative equal to the arithmetic average of the period 0 and 1 expenditure shares for commodity n. Using these probabilities of selection, Theil's final measure of overall logarithmic price change was

$$(5)\ \ln P_T(p^0,p^t,q^0,q^t) \equiv \sum_{n=1}^{N} (1/2)(s_{0n} + s_{tn}) \ln(p_{tn}/p_{0n}).$$

Taking the exponential of both sides of (5) leads to the *Törnqvist-Theil price index*, $P_T(p^0,p^t,q^0,q^t)$.[15] There were many more stochastic approaches to index number theory that were available to price statisticians in 1994 but for our purposes, we will regard $P_T(p^0,p^t,q^0,q^t)$ as a "best" index number formula from the viewpoint of the stochastic approach to index number theory.[16]

The Test Approach to Bilateral Index Number Theory

[14] Walsh (1901; 104) (1921; 82-83), Fisher (1911; 194-196) and Keynes (1930; 71-81)vigorously criticized the lack of weighting in the unweighted stochastic approach to index number theory.

[15] Leo Törnqvist (1936) probably had this formula in mind but it was not explicit in his 1936 paper. It was explicit in Törnqvist and Törnqvist (1937).

[16] For additional materials on the stochastic approach, see Selvanathan and Rao (1994), Clements, Izan and Selvanathan (2006) and Diewert (2010). However, there are many more stochastic approaches to index number theory. For example, the estimation of hedonic regressions can be regarded as a stochastic approach to index number theory; see Court (1939) who introduced the term "hedonic regression" and Chapter 4 in Berndt (1991) who provided additional references to this literature around the time of the first Ottawa Group meeting. This approach was used by Ellen Dulberger (1989) to quality adjust the price of computers in the US. Her methodology was picked up by the Bureau of Economic Analysis around this time in order to quality adjust the US price series for computers.

The test approach to index number theory looks at various “reasonable” properties of index number formulae for a generic bilateral price index of the form $P(p^0,p^t,q^0,q^t)$. For example, a *strong identity test* is $P(p^0,p^t,q^0,q^t) = 1$ if $p^0 = p^t = p$ for all p, q^0 and q^t. A *weak identity test* (it is easier to satisfy) is $P(p^0,p^t,q^0,q^t) = 1$ if $p^0 = p^t = p$ and $q^0 = q^t = q$ for all p and q. Thus the weak identity test requires the index number formula to equal unity (which means that there is no aggregate price change), provided that period t prices *and* quantities, p^t and q^t, are equal to their period 0 counterparts, p^0 and q^0.

The test approach to bilateral index number theory is mainly due to Walsh (1901) (1921) (1924), Fisher (1911) (1922), Eichhorn (1976) and Eichhorn and Voeller (1976). Diewert (1992) and Balk (2008) showed that the Fisher price index $P_F(p^0,p^t,q^0,q^t)$ defined above satisfies more “reasonable” tests than other indexes in common use. Thus our candidate for “best” bilateral price index from the viewpoint of the test approach is the Fisher price index.

The Economic Approach to Index Number Theory

The economic approach to index number theory[17] works as follows: suppose all purchasers of the N products in scope have the same preference or utility function, f(q),[18] where $q \equiv [q_1,\ldots,q_N]$ is a generic vector of aggregate purchases over a time period and $Q^t \equiv f(q^t)$ is the *period t aggregate quantity* that corresponds to the period t vector of purchases, q^t. It is further assumed that all purchasers of the N products face the same period t price vector p^t and they collectively choose q^t as a solution to the following period t utility maximization problem:

(6) $\max_q \{f(q)\ ;\ p^t\cdot q \le e^t\ ;\ q \ge 0_N\} \equiv Q^t$

where e^t is observed period t aggregate expenditures on the N products.[19] The unit cost function that corresponds to the utility function f(q), c(p), is defined as the minimum cost of achieving the utility level 1. If purchasers face the period t vector of prices p^t, we have:

(7) $c(p^t) \equiv \min_q \{p^t\cdot q : f(q) = 1;\ q \ge 0_N\} \equiv P^t$.

Thus the *period t aggregate price level*, P^t, is equal to the unit cost $c(p^t)$ and we have a decomposition of period t expenditure into the product of P^t and Q^t:[20]

(8) $e^t = p^t\cdot q^t = c(p^t)f(q^t) = P^tQ^t$.

Using this approach, we find that the bilateral price index that compares the prices of period t to the prices of period 0 is equal to $c(p^t)/c(p^0) = P^t/P^0$ (this is the Konüs (1924) *cost of living index* between periods 0 and t when we have linearly homogeneous utility functions) and the corresponding quantity index is equal to $f(q^t)/f(q^0) = Q^t/Q^0$.

---

[17] Frisch (1936: 10) called the economic approach the “functional approach”. The founder of the economic approach was Konüs (1924).

[18] It is assumed that f(q) is a linearly homogeneous, nondecreasing and concave function of q over the set of nonnegative consumption vectors, $q \ge 0_N$ where 0_N is a vector of zeros of dimension N.

[19] Thus $e^t = p^t\cdot q^t$. In practice, p^t is the period t vector of unit value prices and the corresponding q^t is a vector of total purchases of the N products during period t.

[20] For the details of this decomposition , see Diewert (1976) (1987).

The practical aspect of this theory comes into play at this point. If we make certain specific assumptions about the functional form for either f(q) or c(p), we can show that $c(p^t)/c(p^0)$ is equal to a known bilateral index number formula. This is the theory of *exact index numbers* that was developed by Konüs and Byushgens (1926)[21], Afriat (1972), Pollak (1983)[22] and Diewert (1976). For example, suppose that the functional form for the unit cost function is $c(p) \equiv p\cdot\beta = \Sigma_{n=1}^{N} p_n\beta_n$ where the β_n are positive constants. Then $q^t = \beta Q^t$, $q^0 = \beta Q^0$, $P^0 = p^0\cdot\beta$ and $P^t = p^t\cdot\beta$. This is the case of *Leontief preferences* where purchasers consume the commodities in fixed proportions over time. In this case, the Laspeyres and Paasche indexes between periods 0 and t will be equal to the true cost of living index, $c(p^t)/c(p^0) = p^t\cdot q^0/p^t\cdot q^0$ $(= P_L) = p^t\cdot q^t/p^0\cdot q^t$ $(= P_P) = p^t\cdot\beta/p^0\cdot\beta = P^t/P^0$.

As a second example, suppose that the unit cost function is equal to $c(p) \equiv [p\cdot Bp]^{1/2}$ where the N by N symmetric matrix of coefficients B satisfies certain regularity conditions.[23] Then Konüs and Byushgens (1926) and Diewert (1976) showed that the cost of living index between periods 0 and t, $c(p^t)/c(p^0)$, is exactly equal to the Fisher index $P_F = [P_LP_P]^{1/2}$. Diewert (1976) went on to show that the above functional form for the unit cost function is a *flexible functional form*; i.e., it can approximate and arbitrary twice differentiable linearly homogeneous unit cost function around a vector of positive prices. This means that the Fisher price index can accommodate a wide range of substitution possibilities between the products whereas the Paasche and Laspeyres price indexes in isolation cannot do this.

For our next example, suppose the logarithm of the unit cost function, $\ln c^*(p)$, is the sum of a constant plus linear and quadratic terms in the logarithms of prices. This unit cost function is called the translog unit cost function. [24] This functional form for a unit cost function is also a flexible functional form. Diewert (1976) showed that the Törnqvist-Theil price index, $P_T(p^0,p^t,q^0,q^t)$ defined by (5) above, is exactly equal to the ratio of the translog unit cost functions, $c^*(p^t)/c^*(p^0)$.

From the viewpoint of the economic approach to index number theory, we can view the Fisher and Törnqvist-Theil price indexes as being equally good.

The assumptions made for the economic approach are very strong and one can see why many price statisticians are skeptical about the usefulness of the economic approach to index number theory. Two justifications for the use of the economic approach to index number theory can be offered:

- When a product goes on sale at a discounted price, usually consumption of that product increases relative to the consumption of other products; i.e., a Consumer Price Index should probably take into account substitution effects. The economic approach to index number theory does take substitution effects into account whereas the other approaches are silent on this issue.
- When attempting to adjust prices for quality change, one cannot avoid comparing the *utility* or *usefulness* of a new product with existing products; i.e., it seems that the economic approach to index number theory is likely to be useful when adjusting for quality change.

---

[21] For a translation and commentary on this paper, see Diewert and Zelenyuk (2024).
[22] Pollak did the research for this paper during the 1970s as a consultant for the US Bureau of Labor Statistics.
[23] These conditions are spelled out in Diewert (1976). Essentially, the matrix B must have one positive eigenvalue with a positive eigenvector. The remaining eigenvalues must be equal to 0 or be negative.
[24] This functional form is due to Christensen, Jorgenson and Lau (1971). See Diewert (1976) (1987) for the details of the exactness of the translog case.

To summarize the above results, we see that the Fisher index $P_F(p^0,p^t,q^0,q^t)$ emerges as a "best" choice of formula using the basket, test and economic approaches and the Törnqvist-Theil index, $P_T(p^0,p^t,q^0,q^t)$ emerges as a "best" choice using the stochastic and economic approaches. Diewert (1978) suggested that in most situations, it would not matter which of these two indexes was used since he showed that these two indexes approximated each other to the second order around a point where $p^0 = p^t$ and $q^0 = q^t$.

Index Number Theory at the First Stage of Aggregation

There is a problem with the bilateral price indexes of the form $P(p^0,p^t,q^0,q^t)$ that were defined above. The above formulae assumed that all of the transactions of a product within the accounting period can be represented by a *single price* and a *single quantity*. It is natural to let the single quantity be the *sum* of the quantities sold. But then, if we want the single price times the single quantity to equal the *value* of transactions for the product in the period, the single price *must* equal the value of transactions divided by the sum of quantities purchased or sold; i.e., the single price must equal a *unit value*.[25]

Note that there is a problem with the unit value solution to the formation of period prices and quantities that could be used in an index number formula. If there is substantial inflation within the accounting period, unit values will give a much higher weight to transactions that occur near the end of the period compared to transactions that occurred near the beginning; it is as if the end of period transactions are being *artificially quality adjusted* to be more valuable than the beginning of the period transactions.[26] It is difficult to deal with this potential problem but it does show that economic measurement is easier if general inflation is low. An obvious solution to this artificial implicit weighting problem is to choose the accounting period to be small enough so that the general inflation within the period is small enough to be ignored. This is precisely the solution suggested by the index number theorist Fisher[27] and the great measurement economist Hicks: the length of the accounting period should be the Hicksian "week".[28] Of course, in most situations, it is not practical to implement this solution.

---

[25] The early index number theorists Walsh (1901; 96) (1921; 88), Fisher (1922; 318) and Davies (1924; 96) all suggested unit values as the prices that should be inserted into a bilateral index number formula. Walsh nicely sums up the case for unit values as follows: "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 principle 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 (1921; 88).

[26] See Hill (1996) for a discussion of measurement problems under conditions of high inflation.

[27] "Essentially the same problem enters, however, whenever, as is usually the case, the data for prices and quantities with which we start are averages instead of being the original market quotations. Throughout this book, 'the price' of any commodity or 'the quantity' of it for any one year was assumed given. But what is such a price or quantity? Sometimes it is a single quotation for January 1 or July 1, but usually it is an average of several quotations scattered through the year. The question arises: On what principle should this average be constructed? The *practical* answer is *any* kind of average since, ordinarily, the variations during a year, so far, at least, as prices are concerned, are too little to make any perceptible difference in the result, whatever kind of average is used. Otherwise, there would be ground for subdividing the year into quarters or months until we reach a small enough period to be considered practically a point." Irving Fisher (1922; 318).

[28] "I shall define a week as that period of time during which variations in price can be neglected." John R. Hicks (1946; 122).

Index Number Theory when Only Price Information is Available

The approaches to index number theory that could be used when *only* price information is available are the *stochastic appro*ach and the *test approach*. The basket approach cannot be used because without quantity information, there are no baskets. The economic approach requires information on prices and quantities.

When discussing the stochastic approach, we have already mentioned two indexes that have been used when only price information is available: the Carli index $P_C(p^0,p^t)$ defined by (2) and the Jevons index $P_J(p^0,p^t)$ defined by (4). Indexes that depend on prices alone are called *elementary indexes* in the index number literature. Another commonly used elementary index is the Dutot (1738) index which is defined as follows:

(9) $P_D(p^0,p^t) \equiv (1/N)\Sigma_{n=1}^N p_{tn}/(1/N)\Sigma_{n=1}^N p_{0n}$.

Thus the period t price level is P^t defined as the period t average price $(1/N)\Sigma_{n=1}^N p_{tn}$ and the period 0 price level P^0 is defined as $(1/N)\Sigma_{n=1}^N p_{0n}$ and the Dutot price index is simple the ratio of these two average price levels.

The Jevons, Carli and Dutot indexes are the most commonly used bilateral indexes that have been used by National Statistical Offices. Which of these three indexes is "best"? We can attempt to answer this question by looking at the test approach to elementary indexes.

We applied the test approach to the Carli index by noting that this elementary index number formula did not satisfy the Time Reversal Test and in fact, violated this test with an upward bias. The Dutot index does not satisfy Fisher's (1922) Commensurability Test; i.e., it is not invariant to changes in the units of measurement. Thus it should only be used if the products in scope have the same unit of measurement. Dalén (1992) initiated a more systematic analysis of the test approach to elementary indexes and ended up endorsing the Jevons index because it satisfied more tests than the Carli and Dutot indexes. In particular, the Jevons index satisfies the following Circularity Test (10) and the weaker Multiperiod Identity Test (11):[29]

(10) $P(p^1,p^2)P(p^2,p^3) = P(p^1,p^3)$;
(11) $P(p^1,p^2)P(p^2,p^3)P(p^3,p^1) = 1$.

Another motivation for the use of the Jevons formula when only prices were available came from the empirical investigations of Bohdan Szulc (Schultz): Szulc (1983; 552-554) found that the upward bias that accumulated from the use of chained Carli indexes could be very high in situations where prices "bounce" to use his terminology. Thus Statistics Canada was one of the first Statistical Agencies to switch from the use of the Carli formula to the Dutot and Jevons formula for elementary indexes.

Walsh introduce his multiperiod identity test in order to provide a measure of an index number formula's discrepancy between the left hand side of (10) and the right hand side of (10). The overall measure of price

---

[29] Westergaard (1890; 218-219) formulated the Circularity Test; the name for the test is due to Fisher (1922; 270). Westergaard noted that the Jevons index was the only index which satisfied property (10) for the indexes which he studied. Walsh (1901; 389) (1924; 506) proposed the Test defined by (11). Diewert (1993a; 40) proposed the name Multiperiod Identity Test for this test. These Tests were originally proposed for the more general indexes of the form $P(p^1,p^2,q^1,q^2)$ but they can be specialized to the case where $P(p^1,p^2,q^1,q^2)$ depends only on p^1 and p^2.

change going from period 1 to 3 on the left hand side *chains* together price change between periods 1 and 2 and between periods 2 and 3 in order to arrive at a measure of price change between periods 1 and 3 whereas the right hand side of (10) measure the price change between periods 1 and 3 directly. Frisch (1936; 7) attributed the concept of chained indexes (the left hand side of (10)) to Marshall (1887). Frisch (1936; 9) termed any difference between the two sides of (10) as the *chain drift* of the index number formula. The chain drift problem is a problem that is still very much with us today.

The above material gives a brief summary of index number *theory* as it existed in 1994.[30] There is one major problem with the above theory: it directly or implicitly assumes that the underlying prices and quantities, p_{tn} and q_{tn}, are positive. This means that price ratios of the form p_{tn}/p_{1n} always exist as finite numbers and this enables the above theories applied to comparisons of prices between any two periods to be based on the use of these matched prices. At higher levels of aggregation, prices and quantities are in general positive, but at lower levels of aggregation, the introduction of new goods and services means that many prices are not matched and many of the above approaches to index number theory do not address this lack of matching.

In the following section, we look at the practical advice on constructing a Consumer Price Index (CPI) that was available to price statisticians in 1994.

3. The 1989 Turvey Consumer Price Indices Manual

In 1989, The International Labour Organization (ILO) published *Consumer Price Indices: An ILO Manual* written primarily by Ralph Turvey (1989), who was the Director of the Labour Information and Statistics Department at the ILO in Geneva. In addition to Turvey's main text, there are 8 Appendices in this very useful book. Appendices 2 to 8 are articles that are reprinted from the ILO Bulletin of Labour Statistics. We find it convenient to review two of these Appendices before reviewing the main Manual.

Appendix 1 has the title "Resolution concerning consumer price indices". This document was a Resolution made by the Fourteenth International Conference of Labour Statisticians, which met at the ILO Headquarters in Geneva, October 28 to November 6, 1987. The Resolution was intended to provide principles and methods in order to construct a CPI as well as guide lines and standards that are generally accepted as constituting good statistical practice. Here is a general statement on the purpose and construction of a CPI:

"The purpose of a consumer price index is to measure changes over time in the general level of prices of goods and services that a reference population acquire, use or pay for consumption. A consumer price index is estimated as a series of summary measures of the period-to-period proportional change in the prices of a fixed set of consumer goods and services of constant quantity and characteristics, acquired, used or paid for by the reference population. Each summary measure is constructed as a weighted average of a large number of elementary aggregate indices. Each of the elementary aggregate indices is estimated using a sample of prices for a defined set of goods and services obtained in, or by residents of, a specific region from a given set of outlets or other sources of consumption goods and services." Turvey (1989; 124).

Elementary indexes will be defined shortly in the Resolution. The exact nature of the aggregate weights is not spelled out in detail in the above quotation. Note that product prices are expected to come from retail outlets rather than from actual household expenditures.

[30] As was mentioned earlier, a much more complete history is given in Balk (2008).

Here is how the Resolution defined elementary aggregates:

"In defining elementary aggregates (in terms of kinds of goods or services, types of outlets and regions), the following principles should be observed:
(a) related goods or services which are thought to display similar price movements should be grouped together in an elementary aggregate;
(b) goods or services whose prices might reasonably be expected to move markedly differently should not be grouped together in the same elementary aggregate;
(c) elementary aggregates should be distinguished whenever weights (including regional or outlet weights) are available or can be estimated;
(d) such regional or outlet weights should be used in calculating the index even when separate regional or outlet-type sub-indices are not required;
(e) elementary aggregates should be described so that any good or service can be unambiguously assigned to the appropriate elementary aggregate.
In the calculation of elementary aggregate indices, consideration should be given to the possible use of geometric means." Turvey (1989; 125).

The Resolution recognized that weights may not be available at the elementary level but if they are available, they should be used. The precise functional form that should be used to aggregate prices at the elementary level was not spelled out but there was a preference for the use of geometric means.

The Resolution describes the type of weights that should be used at higher stages of aggregation as follows:

"In deriving the weights of the elementary aggregates, a household expenditure survey is usually the main source of data. As far as resources permit, such surveys should be representative of household size, income level, regional location, socio-economic group and any other factors which may have a bearing on household expenditure patterns. The period of the survey should be a normal one (or temporary abnormalities should be adjusted in determining the weighting pattern) and should preferably cover a whole year if seasonal variations in expenditure patterns are important." Turvey (1989; 126).

The important points to note in the above description of weights is that they should be *annual* weights. Thus annual weights should be used that are *representative* for monthly (or quarterly) purchases of consumer goods and services made by households.

Some possible problems emerge from the above description of ILO CPI methodology:

- There is a probable lack of weighting at the elementary level.
- Monthly prices will mainly be collected by retail outlet surveys but expenditures and expenditure weights will mainly be collected by different household surveys and the resulting surveys may be inconsistent.
- The existence of strongly seasonal products (products that are only available at certain seasons of the year) is not consistent with the use of annual weights.

Of course, at the time of the writing of the Turvey Manual, scanner data was just beginning to come into existence and so price statisticians faced with the task of constructing a CPI were limited in what they could do with the resources at hand.

Here is what the Resolution had to say about the frequency for changing the upper level weights:

"The weights should be examined periodically, and particularly if economic circumstances have changed significantly, to ascertain whether they still reflect current expenditure or consumption patterns. The weights should be revised or adjusted if the review shows that this is not the case. In any case, they should be revised at least once every ten years." Turvey (1989; 126).

The prices that appear in the elementary aggregates are *sampled prices* from the set of products that are in scope for the particular elementary category of products. The Resolution recognized that there was a problem with disappearing products. The proposed solution to a disappearing product price was to replace it with the price of a similar product:

"Substitutions will be necessary when priced items disappear permanently from the outlet(s) in which they are priced. … Evaluations of the difference in characteristics and decisions on how to use substitute prices in the index should, to the extent possible, be based on solid, empirical evidence of the market valuation of the difference in characteristics between the original and the substitute items. A number of techniques and data sources may be used to approximate this market valuation. In the absence of a satisfactory estimate of the specific adjustment for the difference in characteristics, a choice must be made between an assumption of no change and an assumption that the price difference is simply and wholly a reflection of the difference in characteristics. Under the former assumption, the price for the substitute should be compared directly with that of the item for which it is substituted; this assumption can be made only when the items are fairly similar. Where the whole price difference is taken as a reflection of the difference in characteristics, the index should be constructed by linking the series for the substitute to that of the item for which it is substituted." Turvey (1989; 128).

It can be seen that the problem of finding a substitute product price for a missing product is associated with a quality adjustment problem. The above procedures for making the required quality adjustments are not entirely satisfactory.

The broad outlines of CPI theory and practice at the time of the first Ottawa Group meeting were laid out in the above Resolution. There is some ambiguity on exactly what the upper level weights should be and what formula should be used to aggregate prices when quantity information is not available. Appendix 7 by Bohdan Szulc in the Turvey Manual[31] addresses these questions.

Appendix 7 has the title "Price indices below the basic aggregation level".

Szulc on pages 167-168 of the Turvey Manual lays out the basic methodology for representing a Lowe index using annual expenditure weights for a past reference year "a" and monthly elementary price indexes for say N categories of goods and services. Since this methodology is still used today, it is worthwhile to explain it in some detail.

Suppose initially that there are N goods and services in scope for the index. All prices are assumed to be positive. Let the prices for month t be given by $p^t \equiv [p_{t1},\ldots,p_{tN}]$ with the base month prices defined as $p^0 \equiv [p_{01},\ldots,p_{0N}]$. We assume that we can define an annual price $p^a \equiv [p_{a1},\ldots,p_{aN}]$ for the same list of products for a past year indexed by "a". There is a companion annual vector of quantities, $q^a \equiv [q_{a1},\ldots,q_{aN}]$. These annual data can be used to form annual expenditure shares, s_{an}, defined as follows:

[31] Appendix 7 is a re-publication of Szulc (1987).

(12) $s_{an} \equiv p_{an}q_{an}/p^a \cdot q^a$; $\quad n = 1,\ldots,N.$

Using month 0 as the base for the monthly price index and using q^a as the annual basket, the *Lowe index* for the level of prices in month t relative to month 0 is defined as follows:

$$
\begin{aligned}
(13)\ P_{Lo}(p^0,p^t,q^a) &\equiv p^t \cdot q^a/p^0 \cdot q^a \qquad t = 0,1, \ldots \\
&= \Sigma_{n=1}^N\, p_{tn}q_{an}/p^0 \cdot q^a \\
&= \Sigma_{n=1}^N\, (p_{tn}/p_{on})p_{on}q_{an}/p^0 \cdot q^a \\
&= \Sigma_{n=1}^N\, (p_{tn}/p_{on})s_{oan}
\end{aligned}
$$

The *hybrid expenditure shares*[32] s_{oan} which appear in the last line of (13) are defined as follows:

(14) $s_{oan} \equiv p_{on}q_{an}/p^0 \cdot q^a$; $\quad n = 1,...,N.$

Thus the hybrid shares use the prices of month 0 and the quantities of year a. Thus the Lowe basket type index can also be expressed as a share weighted average of the monthly relative prices for the N products, p_{tn}/p_{on}, for $n = 1,\ldots,N$.

The annual expenditure shares s_{an} defined by (12) can be "price updated" to month 0 by multiplying the nth annual share s_{an} by the price index (p_{on}/p_{an}) to obtain the updated "shares" s_{oan}^*:

$$
\begin{aligned}
(15)\ s_{oan}^* &\equiv s_{an}(p_{on}/p_{an}) \qquad t = 0,1, \ldots \\
&= (p_{on}/p_{an})p_{an}q_{an}/p^a \cdot q^a \\
&= p_{0n}q_{an}/p^a \cdot q^a\ .
\end{aligned}
$$

However, the above "shares" do not sum to one in general. Note that:

(16) $\Sigma_{n=1}^N\, s_{oan}^* = p^0 \cdot q^a/p^a \cdot q^a.$

Define the "true" price updated annual expenditure shares, s_{oan}^{**}, by dividing the shares s_{oan}^* defined by (15) by $p^0 \cdot q^a/p^a \cdot q^a$:

$$
\begin{aligned}
(17)\ s_{oan}^{**} &\equiv s_{oan}^*/(\Sigma_{m=1}^N\, s_{oam}^*) && t = 0,1, \ldots \\
&= (p_{0n}q_{an}/p^a \cdot q^a)/(p^0 \cdot q^a/p^a \cdot q^a) && \text{using (15) and (16)} \\
&= p_{0n}q_{an}/p^0 \cdot q^a \\
&= s_{oan} && \text{using (14).}
\end{aligned}
$$

Thus under the above assumptions, the Lowe index defined by (13) can be defined in terms of the hybrid shares s_{oan} defined by (14) or in terms of the price updated shares defined by (17); i.e., we have:

(18) $P_{Lo}(p^0,p^t,q^a) \equiv p^t \cdot q^a/p^0 \cdot q^a = \Sigma_{n=1}^N\, (p_{tn}/p_{on})s_{oan} = \Sigma_{n=1}^N\, (p_{tn}/p_{on})s_{oan}^{**}$; $\quad t = 0,1,\ldots$.

Note that the information required to implement the last expression for the Lowe index in (18) consists of: (i) estimates of the annual shares, s_{an} defined by (12); (ii) the monthly price relatives p_{tn}/p_{on} and (iii) the

[32] Szulc refers to these shares as hybrid-value shares; see Turvey (1989; 167-169).

price relatives p_{on}/p_{an} that compare the annual prices p_{an} for the products in the base year to the monthly prices in the base month p_{0n}. In practice, the N price ratios p_{tn}/p_{on} are not "true" product price ratios; instead they are elementary (bilateral) price indexes for a category of very similar products. Thus the various representations for the Lowe index given by equations (18) no longer hold exactly; they are only approximate equalities.[33]

Szulc goes on to describe explicit formulae for constructing elementary indexes or micro-indices using his terminology. He introduced this topic with the following paragraph:

"There is an abundant literature, both theoretical and descriptive, on the computation of consumer price indices above the basic aggregation level, but little is written about their derivation below that level. In this respect, the index makers resemble those chefs who only allow their dishes to be presented to patrons at a certain stage of preparation, without showing how they have been mixed and simmered in the kitchen. The reserve in explaining these details does not imply that the meals are unhealthy or tasteless, and a similar conclusion holds for price indices. On the other hand, the early stages of preparation do impact on final results and are of interest, at least to some specialists." Turvey (1989; 170-171).

On pages 171-172, Szulc derives 6 elementary indexes that measure price change between periods 0 and 1. As usual, let the positive price vectors for N products in period t be $p^t \equiv [p_{t1},\ldots,p_{tN}]$ for t = 0,1. His first approach to exhibiting useful elementary indexes works by first defining *price levels* for periods 0 and 1. Thus define the period t Arithmetic, Geometric and Harmonic *price levels*, $P_A{}^t$, $P_G{}^t$ and $P_H{}^t$, as follows:

(19) $P_A{}^t \equiv \Sigma_{n=1}{}^N (p_{tn}/N)$; $P_G{}^t \equiv (\Pi_{n=1}{}^N p_{tn})^{1/N}$; $P_H{}^t \equiv [\Sigma_{n=1}{}^N (1/N)(p_{tn})^{-1}]^{-1}$; t = 0,1.

The corresponding price indexes between periods 0 and 1, $P_A(p^0,p^1)$, $P_G(p^0,p^1)$ and $P_H(p^0,p^1)$, are defined as ratios of the above price levels:

(20) $P_A(p^0,p^1) \equiv P_A{}^1/P_A{}^0$; $P_G(p^0,p^1) \equiv P_G{}^1/P_G{}^0$ and $P_H(p^0,p^1) \equiv P_H{}^1/P_H{}^0$.

In his second approach to the derivation of elementary indexes, he took the arithmetic, geometric mean of the N price ratios directly in order to define $P_A{}^*(p^0,p^1)$, $P_G{}^*(p^0,p^1)$ and $P_H{}^*(p^0,p^1)$:

(21) $P_A{}^*(p^0,p^1) \equiv \Sigma_{n=1}{}^N (1/N)(p_{1n}/p_{0n})$; $P_G{}^*(p^0,p^1) \equiv \Pi_{n=1}{}^N (p_{1n}/p_{0n})^{1/N}$; $P_H{}^*(p^0,p^1)) \equiv [\Sigma_{n=1}{}^N (1/N)(p_{1n}/p_{0n})^{-1}]^{-1}$.

Szulc showed that $P_G(p^0,p^1) = P_G{}^*(p^0,p^1)$, which in turn is equal to the Jevons index $P_J(p^0,p^1)$ defined by (4) above. Note also that $P_A(p^0,p^1)$ is equal to the Dutot index $P_D(p^0,p^1)$ defined earlier by (9) and $P_A{}^*(p^0,p^1)$ is equal to the Carli index $P_C(p^0,p^1)$ defined earlier by (2). Szulc also showed that:[34]

(22) $P_A{}^*(p^0,p^1) \geq P_G{}^*(p^0,p^1) \geq P_H{}^*(p^0,p^1)$.

Szulc made the following observations about the properties of the 5 elementary indexes[35] that he considered above:

---

[33] The approximation of the price ratio p_{on}/p_{an} by an elementary index counterpart is likely to be problematic. The problem of zero prices also creates problems for this methodology.

[34] See Turvey (1989; 173). The inequality follows from an inequality in Hardy, Littlewood and Polya (1934; 26).

[35] Remember that $P_G(p^0,p^1) = P_G{}^*(p^0,p^1)$ so that there are only 5 indexes instead of 6.

- The Dutot index perhaps gives too much weight to higher priced products and becomes problematic if the products being aggregated are heterogeneous.
- The Carli index is likely to give a higher measure of price change than the other indexes and he noted that the upward bias in a chained Carli index could be quite spectacular if prices "bounced".[36]
- He liked the fact that the Dutot[37] and Jevons indexes were transitive; i.e., they satisfied the circularity test (10) above.

This completes our discussion of Appendices 1 and 7 of the Turvey CPI Manual. We turn now to a discussion of the main text in this Manual.

Of course, Turvey endorsed the materials in the Appendices but he added a lot of additional useful materials on various CPI topics. We discuss some of these materials below.

Turvey explains the purpose of the Manual as follows:

"This manual is aimed at practising statisticians who have to construct or revise a consumer price index (CPI). It reflects the international resolution on the subject, reprinted in Appendix 1, but goes beyond it in discussing matters of detail which could not be covered in such a brief document. It is also designed to help the users of consumer price indices, including students of economics, who need to learn about the problems and limitations of these indices. Finally, it is addressed to governments which need to know about the resources required by their statisticians to produce a reliable index.

The manual deals with the practice of consumer price index numbers and does not attempt to survey the academic literature on the subject. Much of that extensive and fascinating literature is irrelevant for the purposes of this manual. One reason is that no compiler of a consumer price index, whether it be monthly or quarterly, can hope to obtain new weights more than once a year at the most, and the data used to compute new weights always refer to the past rather than to the present, whereas much of the literature deals with other types of index." Turvey (1989; 1).

The first paragraph in the above quotation could also describe why I have written the present document!

The second paragraph in the above quotation is also important. Turvey is quite correct to note that much of index number theory appeared to be irrelevant in 1989 because traditional index number theory assumed that current price and quantity data were available to the price statistician whereas this is not the case when the official statistician is asked to produce an up to date Consumer Price Index when real time quantity (or value data) were simply not available to the statistician. This lack of quantity situation started to change in the 1980s when retail outlets started to record all product sales electronically and they were able to store this transactions data. It took time for price statisticians to realize that the academic index number theory that assumed the availability of price and quantity data was no longer irrelevant. It also took time for private companies to realize that it was a good thing for them to contribute their scanner data to national statistical agencies so that more accurate price indexes could be produced.

---

[36] He referred to Szulc (1983) where he demonstrated this upward bias.

[37] He noted that Canada mostly used a chained Dutot index because it "has a long tradition and is considered to be more easily understood by the general public than other transitive index formulae, such as the ratios of equi-weighted harmonic or geometric mean prices. The reasons for using this particular transitive micro-index formula are not that strong, though, and in future discussions other factors might be taken into account as well." Turvey (1989; 175).

Turvey (1989; 4-8) discussed the various purposes that a CPI could be used for.

Turvey (1989; 8) noted several reasons why a CPI could differ from a national accounts price deflator for household consumption.

Turvey discussed the *national* (or permanent resident) versus the *domestic perspective* for a consumer price index as indicated in the following quotation:

"Sales in a region or purchases by its residents? The households within any region may make some of their purchases outside that region and some of the household purchases made within the region may be made by households resident outside the region. In either case, there will be a difference between observing the prices paid within the region and the prices paid by residents of the region. This raises questions concerning both the purposes of the index, the subject of this chapter, and about the sampling aspects of index construction, discussed in a later chapter." Turvey (1989; 10).

The national perspective looks at the household consumption of permanent residents (and excludes tourist consumption expenditures and the consumption expenditures of temporary residents) while the domestic perspective looks at the sales of consumer goods and services made by producers located in the country. The latter perspective is the approach taken in the international System of National Accounts (which has a production perspective) while the former perspective is the usual perspective taken by producers of a CPI. This distinction is related to why a CPI could differ from a national accounts deflator for household consumption. However, note that typically, CPI statisticians, using the methodology explained in the Turvey Manual, collected their micro prices from retail shops rather than from the actual expenditures of resident households. A problem is that consumer goods and services that are purchased from domestic retailers include the purchases of tourists, domestic businesses or of government agencies. Also when taking the national perspective to the construction of a CPI, permanent residents purchase goods and services from abroad and these purchases cannot be tracked using the sales of domestic retailers.[38]

Here is what Turvey had to say about the inclusion of household production of consumer goods and services:

"No one doubts that for a number of purposes the addition of the imputed value of any own-account production and of any income in kind should be added to a measure of money income or consumption to obtain a measure of total income or consumption. It may also be useful to include a measure of the value of government services provided free. If a deflator for such a total measure is required, then this deflator should obviously include the imputed prices of these imputed values. But when the sales to be deflated or the incomes to be deflated, evaluated or determined include no value imputations, then the price index should not include them either. Hence, whether or not to include imputed items should depend on what is the most important purpose for which the consumer price index is to be used." Turvey (1989; 12).

---

[38] Turvey (1989; 14) later commented on the problems associated with including or excluding household expenditures made abroad: "The resolution of the Fourteenth International Conference of Labour Statisticians specifies that the extent to which expenditure abroad is included in the index should be made clear. The issue of principle arises, as already noted, with respect to expenditure on holidays abroad and cross-border shopping. For some purposes, the whole of the cost of holidays is relevant, while for the analysis of inflation, for example, it is only their domestic component (e.g. travel on a national airline) which is of interest. The same applies to items which some households buy on cross-border shopping expeditions. As with all these issues, there are practical considerations as well. The only feasible way of measuring the cost of overseas holidays may be to collect the prices of package holidays, and these often involve the problem of seasonal unavailability. Cross-border price collection may be impracticable."

The above paragraph raises an important reason for having separate indexes: one that is largely free of imputations that central banks could use to monitor domestic household inflation and another that would use imputations in order to better measure actual household consumption.

Turvey (1989; 12-13) discussed some of the problems associated with several categories of "difficult to measure" goods and services like insurance and financial services. With respect to property insurance, he discussed the gross premiums versus the net premiums (premiums paid less the value of claims) approaches in some detail. This is an important issue which has not been settled.[39]

Turvey discussed the problems associated with matching *payments* for items of household consumption to their *use*:[40]

"Consumption expenditure can be conceived and measured in three ways which it is important to distinguish. The recommendation defines them as follows:

- Acquisition indicates that the total value of all goods and services delivered during a given period, irrespective of whether they were wholly paid for or not during the period, should be taken into account.
- Use indicates that the total value of all goods and services actually consumed during a given period should be taken into account.
- Payment indicates that the total payments made for goods and services during a given period, without regard to whether they were delivered or not, should be taken into account.

Note that differences between the three concepts of consumption are not merely a matter of timing. If payment follows acquisition, interest may be charged in addition to the equivalent of the cash price. When use extends over several years, the value of this use will reflect the price level of those years, not the price at the date of acquisition." Turvey (1989; 15-16).

Turvey (1989; 16-24) addressed the problems associated with measuring the price of Owner Occupied Housing (OOH) in a CPI. His discussion is excellent. After defining three broad approaches to this measurement problem, he presents the following menu of seven alternative treatments of OOH:

"Formulating the three questions previously listed in more detail now yields the following set of more specific, alternative questions.
(A) Net acquisitions. What is the change through time in the total purchase value of a sample of new owner-occupied dwellings similar to the new owner-occupied dwellings acquired by consumers in the reference period?
(Bl) User cost (1). What is the change through time in the mortgage interest and conventional depreciation at replacement cost in respect of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?
(B2) User cost (2). What is the change through time in the opportunity cost of the invested capital value, plus depreciation, less accruing capital gains, in respect of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?
(B3) User cost (3). What is the change through time in the estimated rental value of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?

[39] I return to this issue in the final section of this paper.
[40] Turvey (1989; 15) made the following important point about this topic: "Note that differences between the three concepts of consumption are not merely a matter of timing. If payment follows acquisition, interest may be charged in addition to the equivalent of the cash price. When use extends over several years, the value of this use will reflect the price level of those years, not the price at the date of acquisition."

(C1) Payment (1). What is the change through time in the cash outlays on down payments on purchases, mortgage interest and repayments in respect of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?
(C2) Payment (2). What is the change through time in the cash outlays on mortgage interest and repayments in respect of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?
(C3) Payment (3). What is the change through time in the cash outlays on mortgage interest, excluding repayment elements, in respect of a sample of owner-occupied dwellings similar to consumers' owner-occupied dwellings in the reference period?" Turvey (1989; 17-19).

Turvey recognized that the purpose of the index would determine which approach to pricing OOH should be chosen. For example, a CPI that central bankers could use to measure general inflation might want use the acquisitions approach due to its apparent lack of use of imputations,[41] national income accounts might to use a user cost approach and governments might want to use a payments approach to index pensions and transfer payments to households. On page 22, Turvey lists the data requirements for implementing each of his seven approaches. Turvey's discussion of OOH is in general excellent.[42] This section of the Turvey Manual makes a strong case for having more than one CPI to serve different purposes.

Turvey argued that consumer durables should be treated in the same way that OOH is treated in the index:[43]

"If these arguments are not accepted, the case for treating consumer durables and owner-occupied dwellings, as far as practicable, in the same way is a strong one." Turvey (1989; 25).

Turvey addressed the following question: Should the index relate to a point in time or a period?

"This choice is evidently more important with a quarterly index than with a monthly one, though, as in the rest of this manual, the exposition runs in terms of a monthly index. If used for deflating income, expenditure or sales, the index should obviously relate to the period of time to which the money flow in question relates. For economic analysis, where the index will be used in conjunction with other economic statistics, most of which relate to a period rather than to a point in time, it seems appropriate that the consumer price index should do the same." Turvey (1989; 25).

Turvey went on to indicate that practical considerations would often force price statisticians to collect prices at a point in time.[44]

---

[41] The acquisitions approach when applied to housing is not straightforward. The problem is that housing property consists of a structure and a plot of land that the structure sits on. A sale of a house with a structure which is not new between households simply cancels out and should not affect the overall CPI. Thus the acquisitions approach should concentrate on the sales of properties which have a *new structure* on the land plot. But if the seller of the property is also a household, then the land part of property value should also not affect the overall CPI. Thus the price statistician must make an *imputation* to decompose the purchase price of the property into its land and structure components. The European Economic Union's Harmonized Index of Consumer Prices (HICP) uses the acquisitions approach and it has struggled with the problem of how to deal with OOH. The solution thus far has been to exclude OOH from the HICP.
[42] A reservation I have about his analysis of OOH is that he did not recognize that a housing property has a *structure component* and a *land component*. Depreciation applies to the structure component but not to the land component. Thus on page 19, Turvey derives a standard user cost for a housing property but he applies a depreciation rate d to the entire property instead of applying it only to the structure portion of property value. In addition to property location, structure size and structure age being important determinants of property value, the size of the underlying land plot is also an important property price determining characteristic.
[43] "If these arguments are not accepted, the case for treating consumer durables and owner-occupied dwellings, as far as practicable, in the same way is a strong one." Turvey (1989; 25).
[44] In recent times, businesses have realized that they can practice a form of price discrimination by using "dynamic pricing"; i.e., by changing prices rapidly. In 1989, it was possible to argue that a single monthly price was likely to be

Chapters 3 and 4 of Turvey deal with the choice and weighting of elementary aggregates and it covers much the same ground as was covered in Appendix 7 but with added detail; e.g., Turvey presents materials on the sampling of products, sources of data and the selection of representative products.

Chapter 5 is on the details of how to collect prices in practice. The material on quality adjustment and replacement of disappearing products on pages 76-82 is of particular interest. Here is how Turvey introduced this topic:

"Quality judgements take two forms. On the one hand, there is the search for a variety which is of the same quality as the one to be replaced, so that its price can be used instead of the old one. On the other hand, there is the evaluation of any difference in quality between the new and the old variety. In both cases, differences which are relevant to quality as seen by consumers have to be distinguished from irrelevant differences. In the first case, a variety has to be sought for which the relevant differences are minimal. In the second case, the new variety is chosen on other grounds and the problem is to put a monetary value on any differences which are not minimal." Turvey (1989; 77).

On pages 80-82, Turvey describes the key features of a hedonic regression, where the price of a product is regressed on its price determining characteristics. A hedonic regression can be used to provide quality adjusted prices for new products. He noted that the use of this technique is absolutely necessary in the case of house price indexes since each property is basically a unique product:[45]

"In the field of consumer price indices, the most promising use of the technique relates to the prices of existing dwellings. On the one hand, it is necessary because the dwellings sold in any period will hardly ever be the same dwellings sold in the preceding or reference period, so the need for it is particularly great. On the other hand, a reassuringly large number of price observations may be obtainable with descriptions and measurements of the relevant characteristics. For example, a British study uses about 12,000 house price observations per month, with the following variables:

- House type: detached, semi-detached, terraced, bungalow, flat.
- Number of: habitable rooms, bathrooms, separate toilets, garages, garage spaces.
- Presence of a garden, of a plot of 1 acre or more.
- Central heating: full, partial, none.
- Freehold.
- Location (12 regions).
- Age of property in years." Turvey (1989; 81).

Chapter 6 on computation of the CPI is largely an elaboration of the Appendix 7 material due to Szulc. However, there is some additional materials on how to deal with missing prices and seasonal products.

On page 90, Turvey observed that the geometric mean of the price relatives (ratios) is the ratio of the geometric mean price levels as was explained in Appendix 7. However, Turvey's comments on this property of geometric means is worth quoting:

---

representative for prices of the product during the entire month but now this argument is less convincing. Using a monthly unit value price is likely to be much more representative than a single price quote.

[45] The property characteristics listed in the quotation below are very reasonable. In my work on housing hedonics with coauthors, we found that the actual floor space of the structure and the area of the land plot are important price determining characteristics. We also used splines on these characteristics instead of just entering floor space and plot size as linear variables. The treatment of condominium apartments is more complicated. See Diewert, de Haan and Hendricks (2015), Diewert and Shimizu (2015) (2017) (2022) and Burnett-Isaacs, Huang and Diewert (2020).

"In view of this superiority, it is not surprising that many statisticians regard the use of geometric means as the best solution. Why, then, are they so seldom used? One reason is that they make the calculations difficult, but this argument loses its validity once computers are used for calculating the index. A second reason is that it may be feared that their use is too difficult to explain to users of the index. But most indices have features which are difficult to explain, and in any case, the degree of complexity that is acceptable is growing, through time, in most countries. Statisticians should have the courage of their convictions." Turvey (1989; 90-92).

On pages 92-99, Turvey deals with the problem of missing observations. He suggests four methods to deal with this problem:

"When an individual observation is unavoidably missing, the calculations should omit the corresponding price from the data set with which current prices are compared, so that like is compared with like. In other words, the two sets of prices must be "matched". However, matching can also be achieved by using an imputed price, calculated in one of the following three ways:
(a) by carrying forward the previous observation, thus assuming no price change, the simplest procedure, which is clearly acceptable only when there is not much inflation;
(b) by assuming that the price would have moved in the same proportion as those prices within the elementary aggregate which were recorded;
(c) a missing price may also be imputed by using an observation from another, similar outlet which is not included in the regular price collection." Turvey (1989; 92).

Of course, replacing a missing price with the price of an alternative product can lead to a quality adjustment problem as is acknowledged by Turvey who discussed the advantages and disadvantages of the above four methods.

On pages 102-107, Turvey attempts to deal with the problem of strongly seasonal products. These are products which are only available for certain months of the year. This is where the methodology of using a fixed annual basket in the context of a monthly price index runs into difficulties. Turvey recognized the difficulties:[46]

"We now come to a problem where there seems to be no absolutely correct answer. Different methods are used in different countries, yielding dissimilar results. This is not because of differences in the purposes to be served by the consumer price index, but because of various practical approaches to a purely technical problem. This problem extends beyond seasonally varying prices, and arises when some goods or services are totally unavailable at certain times during the year, or are only available to a very limited extent when not in season, so that meaningful prices cannot be observed. The problem, a significant one, can arise with fruits, vegetables, meat, fish, sports goods, package holidays and some kinds of clothing. The essence of the problem is that if one single set of weights is used, fictitious prices must be imputed for those months when there is no price to observe. Alternatively, if different weights are used for different months, reflecting the varying availabilities of items, the meaning of month-to-month changes in the index becomes unclear." Turvey (1989; 103).

[46] A decade earlier, Turvey (1979) set a set of fictitious seasonal data to many national statistical offices around the world and asked them to use their usual procedures for dealing with seasonal data in order to construct a consumer price index for this data set. He got back about 20 answers. There was no consensus on a suitable method and the answers showed a great deal of variation.

Turvey's suggested solutions to the problem of constructing a monthly subindex for strongly seasonal data using existing CPI methodology did not get around the fundamental difficulties mentioned above. However, Turvey did have one very sensible suggestion for producing a rolling year annual index:[47]

"There is one procedure akin to seasonal adjustment which escapes both the need to impute fictitious prices and the alternative problem of interpreting month-to-month movements. This is the calculation of a 12-month centred moving average. Each month it compares the current 12-month cost of buying 12 reference-year monthly baskets with the total cost of the 12 reference-year baskets. From a practical point of view, such a moving average is not very useful, since it entails a publication delay six months longer than usual. However, it can serve as a standard for judging other methods." Turvey (1989; 103).

The problem with this method is that central banks need an index that is sensitive to month to month fluctuations in prices. As the data for the current month become available, the rate of change in the rolling year index depends on not only the price movements of the current month prices with the previous month prices but it also depends on the monthly prices lagged backwards 12 months. Thus the rolling year index will not necessarily alert central bankers to an uptick in inflation. However, the rolling year indexes are well suited for indexation purposes. This is another argument for statistical offices to produce more than one CPI.

In the end, Turvey came to the following conclusion:

"The easiest solution would be to avoid the problem by totally omitting all goods and services not available in all 12 months." Turvey (1989; 106).

Turvey concluded Chapter 6 with a nice discussion on the pros and cons of chaining. He was in favour of chaining annually but not within the year:

"Chaining more frequently than annually is usually impracticable and in any case is undesirable. This is because a chained index usually fails the circularity test of reverting to its original value when some prices, after having changed, return to their original values, with quantities moving inversely to prices. Such oscillatory price and quantity behaviour, if it occurs at all, is most likely to be seasonal, displaying an annual pattern. The advantage of an annually chained index is that it follows the evolution of the pattern of consumption. It is thus more representative than an index using an unchanging set of weights." Turvey (1989; 110).

Chapter 7 is on sources of error and Chapter 8 is on publication issues.

This completes our review of the Turvey CPI Manual. We turn now to a discussion of the first Ottawa Group Meeting.

4. The First Ottawa Group Meeting

---

[47] Mudgett (1955) and Stone (1956) suggested this method for forming an annual index. Diewert (1983) generalized their approach to the comparison of a rolling year consisting of 12 consecutive months of data with the data of a base year which is precisely the method suggested by Turvey. Pages 400-407 of the 2004 *Consumer Price Index Manual* give more details on this method. In the 2004 Manual, it was suggested that this method was a pure index number method for performing seasonal adjustment: "Thus rolling year indices offer statistical agencies an objective and reproducible method of seasonal adjustment that can compete with existing time series methods of seasonal adjustment." ILO, IMF, OECD, Eurostat, UN and World Bank (2004; 403).

As was mentioned, the first Meeting of the Ottawa Group took place in Ottawa over October 31 to November 2, 1994. Jacob Ryten of Statistics Canada, the Chair of the Meeting, provided an introduction to the Meeting. The quotation below summarizes his opening remarks:

"In opening the session, the Chair explained that the purpose of this working group is to bring together an independent forum of specialists from different countries to exchange ideas on crucial problems of measuring price change and to propose concrete solutions. The choice of the proposed topics for discussion originated from a long-standing debate about possible bias in the CPI. In the seventies, while inflation was relatively high, it was commonly believed that there was a downward bias in the CPI. At that time, media reported that the CPI was really too low and that actual inflation should have been higher. In fact, the 1982 Conference on Price Level Measurement organised by Statistics Canada, dealt indirectly with this issue. In recent years the debate on bias has re-emerged, however, notwithstanding that the inflation is low, it is now believed, at least in North America, that the bias is upward and that inflation is overestimated. Proper discussion of this complex issue cannot be done without addressing several detailed questions such as: the sampling error of the CPI, the treatment of substitution of goods, the index formula and the formula aggregation at the macro level. To make significant progress on these issues, it was agreed that the discussion for this first meeting should be devoted to only two main issues: the micro level aggregation and its macro effects, and how to detect and estimate the bias of the consumer price index." Ottawa Group (1994; 2).

Bohdan Schultz's paper at the 1982 Statistics Canada Conference[48] showed that at the elementary index level of aggregation, the use of the chained Carli formula could lead to tremendous upward bias if prices were volatile and of course, at the macro level, a fixed basket type index is subject to possible (upward) substitution bias. Thus there was a need to devise methods to either measure these potential biases or to mitigate them.

Ryten concluded his opening remarks by listing some related research areas that the meeting participants should address:

"The group has also identified a wide range of issues related to this topic including:

- Relevance of the CPI as an indicator of the Cost of Living Index,
- Harmonisation of European price indices,
- Importance of agreeing on definition and concept of cost of living vs. fixed basket,
- Defining a wider measurement of inflation (ex. whole economy price index),
- Need for another economic approach to monitor inflation." Ottawa Group (1994; 2).

Below are some comments on the papers that were presented at the first meeting of the Ottawa Group.

Paper 1 by Marta Haworth (1994) described some of the problems the UK encountered when it moved from sampling products by central office choice to probability sampling. She also discussed problems that occurred when validating the collected data.

Paper 2, by Paul Armknecht, Brent Moulton and Kenneth Stewart (1994), summarized the efforts of the U.S. Bureau of Labor Statistics (BLS) to reduce possible bias at the lowest level of aggregation where only price information is available and at higher levels of aggregation when expenditure information is also available. The authors refer to the very important research of Marshall Reinsdorf (1993) who compared the trends in US average prices for relatively homogeneous product groups with the corresponding trends in official BLS price indexes for the various product groups. Reinsdorf found that the official indexes increased substantially faster than the average price indexes. For example, he found that the official index

---

[48] This paper was published as Szulc (1983).

for food showed average annual increases during the 1980's of 4.2% per year while the weighted mean of average prices grew at only 2.1% per year. He attributed these differences to *outlet substitution bias*; i.e., the effects of consumers shifting their purchases from higher cost outlets to lower cost outlets. It turned out that the bias Reinsdorf observed was not due to outlet substitution bias. It was later discovered by Moulton (1993), Reinsdorf and Moulton (1997) and Reinsdorf (1998) that the bias was due to *elementary formula bias*; i.e., the use of the Carli formula at the elementary level. Greenlees later noted this early BLS research ultimately led to the use of the Jevons formula at the elementary level:[49]

> "Moulton (1993) and the subsequent paper by Reinsdorf and Moulton (1997) demonstrated that a geometric mean, unlike the arithmetic mean, is not vulnerable to the bias described in the previous paragraph. Moulton therefore suggested that the BLS consider moving to a geometric mean formula in the CPI." Greenlees (2006; 24).

Thus the pioneering work of Reinsdorf eventually led the BLS to switch to the Jevons formula.

Armknecht, Moulton and Stewart (1994; 2) also mentioned the work of Manser and McDonald (1988) and Aizcorbe and Jackman (1993) in estimating *upper level substitution bias*. Greenlees later described this BLS research as follows:

> "In the wake of the Reinsdorf (1993) paper and the considerable interest it generated, the BLS devoted the December 1993 issue of the Monthly Labor Review to four articles on consumer price measurement issues written by economists in the Office of Prices and Living Conditions. Notable among these was the paper by Ana Aizcorbe and Patrick Jackman, who compared indexes based on the CPI's Laspeyres formula to indexes employing the "superlative" Fisher Ideal and Törnqvist formulas for the period 1982 to 1991. That paper was the first to construct superlative indexes using detailed CPI series defined by geographic area and item category, and it provided several estimates of the upward substitution bias resulting from use of the fixed-weight Laspeyres formula to aggregate those series. These Aizcorbe-Jackman results became the basis for many subsequent estimates, including those by the Boskin Commission, of what came to be called "upper level" substitution bias." Greenlees (2006; 24).

Prior to the discovery of superlative index number formulae, researchers attempted to measure the substation bias of a Laspeyres or Lowe index (relative to a Cost of Living Index) by estimating systems of consumer demand functions, which is a tricky business. The Manser and McDonald paper is an example of this approach. Aizcorbe and Jackman (1993) showed that it was much simpler to simply compare the fixed basket index with the corresponding Fisher or Törnqvist indexes.

The above BLS research along with the work of Szulc (1983) (1987) indicated that there was a substantial amount of research prior to the Boskin Commission Report[50] into possible biases in the "standard" two stage Lowe index that was described in the Turvey CPI Manual.

A third area of BLS research mentioned in the Ottawa Group 1994 Paper 2 is the idea of performing *quality adjustment* at the first stage of aggregation. The authors of Paper 2 referred to Griliches and Cockburn (1994) who suggested that a brand name drug should be regarded as being the same as a generic counterpart.

---

[49] "Moulton (1993) and the subsequent paper by Reinsdorf and Moulton (1997) demonstrated that a geometric mean, unlike the arithmetic mean, is not vulnerable to the bias described in the previous paragraph. Moulton therefore suggested that the BLS consider moving to a geometric mean formula in the CPI." Greenlees (2006; 24). Thus the BLS, like Statistics Canada, independently moved to the use of the Jevons index at the elementary level of aggregation.

[50] See Boskin, Dulberger, Gordon, Griliches, and Jorgenson (1996) (1998). For an excellent review of the effect on the BLS of the Boskin Commission report, see Greenlees (2006). For my own attempts to measure the various biases that were identified in the Boskin Commission Report, see Diewert (1995a) (1996) (1998).

Thus instead of treating the generic and brand name drug as separate products, treat them as a single product so that the overall period price of the product is the unit value price defined over both products.

An overall view on Paper 2 is that it provides us with some evidence of the problems associated with the two stage methodology that is explained in the Turvey Manual. It provided support for the use of the Jevons formula at the lowest level of aggregation. It also provided an introduction to the sources of bias and their magnitude that was the focus of the Boskin Commission Report which appeared in 1996.

Paper 3 was authored by Sellwood (1994). This paper is very interesting for a number of reasons:

- It provided insight into the formation of Eurostat's Harmonized Index of Consumer Prices (HICP). Sellwood suggested that the main use of the HICP should be to measure inflation across the member countries in the European Union in a way that would be useful to Central Banks.[51] He argued that *imputed* prices for Owner Occupied Housing and Household Production should not appear in a HICP. Since the purpose of a national CPI might be different from the Central Bank perspective, Sellwood suggested that the HICP would be a supplement to national CPIs rather than a replacement.
- The main purpose of the paper was to give guidance to the HICP on whether to use the Carli or Dutot formula at the elementary level. Sellwood initially thought that once the purpose of the index (what is to be estimated) was known, then it would become clear whether Carli or Dutot would be the preferred choice. However, he noted that this strategy did not prove to be successful.
- He noted that the Dutot index gave too much weight to higher priced items and that this seemed to be a conclusive argument in favour of the Carli index. Thus he appealed to the test approach to bilateral indexes in coming out in favour of Carli over Dutot. But one could counter that the Dutot index satisfies the circularity and time reversal tests whereas the Carli fails both tests.[52] However, Sellwood's support for the Carli index was not strong.
- He was very doubtful about the use of the economic approach to the construction of a CPI.[53]

The last point needs some elaboration. The economic statisticians present at the first Ottawa Group meeting came from North America, Europe and Australia.[54] The North American representatives mostly favoured the economic approach to index number theory[55] while the European and Australian representatives mostly favoured the Lowe index approach at the higher levels of aggregation and the basket, stochastic and test approaches in general.

Paper 4 was authored by Bohdan Schultz (Szulc) (1994). This very important paper followed up on his earlier papers, Szulc (1983) (1987) but in the present paper, he used Statistics Canada monthly micro data

[51] For a discussion of HICP methodology, see Diewert (2002).
[52] It is interesting that Sellwood referred to the work of Dalén (1994) who was also doing research on the methodology for the HICP. We will review this paper by Dalén shortly. Dalén made a strong case for the use of the Jevons formula but Sellwood does not even mention the Jevons index as an alternative elementary index.
[53] "The classical model of a consumer maximising utility by applying a rational decision process to balance costs and benefits seems remote from common experience." Sellwood (1994; 3).
[54] It should be noted that over time, the participants in the Ottawa Group meeting have come from the above three continents plus Asia, Africa and South America.
[55] Triplett (1999) was a strong advocate of the economic approach to index number theory. At one point in time, he criticized me for writing papers that gave support to other approaches to index number theory.

for over 50 commodities, collected from December 1988 to January 1994 in the province of Ontario. He compared 6 alternative elementary index formulae:

"Six micro-index formulae have been taken into consideration in the study, of which three are ratios of mean prices collected in the base and observation times:

- a ratio of equiweighted arithmetic mean prices (RAMP),
- a ratio of equiweighted geometric mean prices (RGMP), and
- a ratio of equiweighted harmonic mean prices (RHMP).

Three other formulae are means of price relatives (expressing price change from the base to observation times for individual items):

- an equiweighted arithmetic mean of price relatives (AMPR),
- an equiweighted geometric mean of price relatives (GMPR), and
- an equiweighted harmonic mean of price relatives (HMPR)." Schultz (1994; 4).

After noting that he calculated both fixed base and chained indexes using the above formulae, he summarized his results as follows:

"Next, chains of these indices were calculated on the December 1988 time base, alternatively with monthly, annual and quinquennial linking. Differences in numerical results are stunning, in some cases they are tenfold and more. This means that the choice of micro-index formulae is a very important decision, likely more important than choosing a macro-index formula at higher aggregation levels. Another major conclusion from the study is that linking micro-indices, particularly frequent linking, greatly increases the effect of using alternative formulae. Consequently micro-index formulae should not be assessed solely according to their behaviour in direct price comparisons, but also, if not primarily, according to their comportment in chain price comparisons. Non-transitive formulae can play havoc with chain index numbers when relative prices tend to bounce." Schultz (1994; 2).

Schultz also explained why Statistics Canada switched from using Dutot indexes to the use of Jevons indexes:

"Until 1994, micro-indices in the Canadian CPI have been calculated by monthly linking of the month-to-month ratios of arithmetic mean prices. As transitive formulae, the latter do not systematically bias the results, which was confirmed by the study. Nevertheless, a switch towards ratios of geometric mean prices was decided with the CPI basket update that took place in January 1995, to avoid problems with those categories of products that exhibit a broad spectrum of prices." Schultz (1994; 2).

Schultz noted that both the Dutot and Jevons formula satisfied the Transitivity (or Circularity) Test and he noted that there was no systematic difference between the two formulae using his Ontario data set. However, there was concern that the Dutot formula gave too much weight to products that had high prices.[56]

It is interesting to see that Statistics Canada and the Bureau of Labor Statistics both independently came to the conclusion that the Jevons formula was "best" at lower levels of aggregation.

Paper 5 was authored by Bert Balk and it has the title "On the First Step of the Calculation of a Consumer Price Index". It can be viewed a follow up paper to Sellwood's paper. The quotation below describes the purpose of his paper.

---

[56] Another important problem with the Dutot formula is that the index is not invariant to changes in the units of measurement. Thus the Dutot formula should only be used if the products are all measured in the same units.

"Sellwood (1994) considered the reasons why it has proved so difficult to resolve the problem of which micro-index formula to use in compiling a CPI. He suggests that "there is a lack of an agreed point of reference on what is to be measured". The present paper tries to add to the discussion by carefully specifying the object of estimation. In my view, a CPI can best be viewed as a group price index; that is a price index which is an average of household specific price indices. This implies that the form of the household specific price indices determines the form of the CPI. As will be shown, this has certain consequences for the form for the price indices at the lowest level of commodity aggregation. In the majority of cases, however, the available information is not sufficient to calculate these indices. They have to be approximated. Various assumptions and/or sampling considerations come into play and give rise to various micro-index formulas. It seems that the search for a micro-index formula which can be applied universally is misdirected, As a by-product we obtain the conclusion that it is hard to formally justify the use of a geometric mean." Balk (1994; 2).

Balk's initial object of estimation was a Laspeyres index. His paper is brilliant in many ways. However, in the end, he shows that all of the micro-indexes that are commonly used are very difficult to justify, given that we have specified various target indexes that ideally we would like to estimate. This is a sobering conclusion but it was not very helpful to national price statisticians who had to choose an elementary index in order to estimate their CPI.[57]

Paper 6 was authored by Jörgen Dalén (1994). This paper and the earlier paper by Dalén (1992) are very important for two reasons:

- Dalén derived easy to understand numerical relationships between the various elementary index number formulae and
- Dalén systematically looked at the properties or tests that the various elementary indexes satisfied; i.e., he initiated a sytematic *test approach* to the study of elementary indexes.

Thus far, we have defined four elementary indexes that are discussed in the elementary index number literature: (i) the Carli index $P_C(p^0,p^1)$ defined above by (2); (ii) the Jevons index $P_J(p^0,p^1)$ defined by (4); (iii) the Dutot index $P_D(p^0,p^1)$ defined by (9) and (iv) the Harmonic index $P_H^*(p^0,p^1)$ defined by (21). We also noted that Szulc (1987; 12) established the following inequalities between three of these indexes:[58]

(23) $P_H^*(p^0,p^1) \leq P_J(p^0,p^1) \leq P_C(p^0,p^1)$.

The fifth elementary index number formula that is discussed in the literature is the geometric average of the Carli and harmonic formulae $P_{CSWD}(p^0,p^1)$, which was suggested by Carruthers, Sellwood and Ward (1980; 25) and Dalén (1992; 140):

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

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.

---

[57] My own view is that the test approach is the most helpful approach in choosing an elementary index.

[58] See the inequalities (22). The inequalities in (23) will be strict provided that the period 0 vector of prices, p^0, is not proportional to the period 1 vector of prices, p^1.

Dalén (1994; 151) had an interesting justification for the three indexes which appear in definition (24). He observed that the Laspeyres index collapses down to the Carli index if the base period expenditure shares are all equal and the Paasche index collapses down to the Harmonic index if the period 1 expenditure shares are all equal. Thus if expenditure shares are equal in both periods 0 and 1, we can view $P_{CSWD}(p^0,p^1)$ as an approximation to the Fisher index. This is an important result. Many National Statistical Offices want their CPI to be a basket type index and they would like their elementary index to be consistent with a basket index. It can be argued that the bilateral Fisher index is a kind of *generalized symmetric basket index* since it is the geometric mean of two equally relevant basket type indexes. At the elementary level, we have no information on expenditure shares so assume that these unknown expenditure shares are equal and we obtain Dalén's approximations to the Laspeyres and Paasche indexes, $P_C(p^0,p^1)$ and $P_H(p^0,p^1)$. Thus $P_{CSWD}(p^0,p^1)$ can be interpreted an approximate Fisher index and hence is an approximation to a *generalized basket type index*.

Dalén (1994; 150) also gave a very clear explanation for the upward bias inherent in the Carli formula; i.e., he explained in simple terms why $P_C(p^0,p^1)P_C(p^1,p^0) > 1$ unless p^1 was proportional to p^0 in which case $P_C(p^0,p^1)P_C(p^1,p^0) = 1$.

More importantly, Carruthers, Sellwood and Ward (1980) and Dalén (1992) developed useful approximate formula which could explain why there were systematic differences between the five elementary index number formulae that were defined above. Basically, these approximations allow us to measure the distance between the three indexes which appear in the inequalities (23) and the above authors showed that to the accuracy of various second order Taylor series approximations, $P_J(p^0,p^1) \approx P_{CSWD}(p^0,p^1) \approx P_D(p^0,p^1)$. Since $P_{CSWD}(p^0,p^1)$ is an approximate generalized basket type index and since $P_J(p^0,p^1)$ approximates $P_{CSWD}(p^0,p^1)$ to the second order around an equal price point, we have a justification for the Jevons index as an approximation to a generalized basket type index.

Finally, Dalén (1992) initiated the *test approach to elementary indexes* by looking at eight different tests that could be applied to a bilateral elementary index. The Jevons index ended up satisfying all of these tests. This was an important result.

Paper 7 was authored by Keith Woolford (1994). This paper also looks at the issues surrounding the choice of an elementary index number formula and the consistency of this choice with choice of the higher level index formula which applies when price and expenditure information is available. Woolford considered the big three elementary index formulae that are used by statistical agencies: the Jevons, Carli and Dutot. He noted (with approval) that the Carli and Dutot indexes were consistent with a basket type index.[59] Woolford liked the Dutot index because it could deal with price bouncing behavior (i.e., it satisfied the circularity test) and it was consistent with a macro basket type index. However, he noted that higher priced items got too much weight in the index and he was worried that the index did not depend on price relatives. In the end, he did not end up with a clearly expressed opinion on which formula was best.

[59] As noted by Dalén above, if expenditure shares are equal in period 1, the Laspeyres index collapses into a Carli index. If quantities are equal to same constant in both periods, then the Laspeyres index is equal to the Dutot index. Triplett (1998) also derived these approximations and discussed the consistency of various elementary indexes with various approaches to the construction of index numbers. Triplett was a vigorous defender of the economic approach to the construction of consumer price indexes but he was well aware of some of the weaknesses of the COLI approach In particular, he noted that if the accounting period is short, consumption will not coincide with the acquisition of consumer goods that can be stored; i.e., he recognized the stockpiling problem. He also stressed the importance of search costs and consumer habit formation. See Triplett (2003) for additional discussion of these points.

In the second half of his paper, Woolford used Australian Bureau of Statistics data to construct the big three elementary indexes for one Australian city:

"To provide some indication of the differences and their impact on aggregate index outcomes, existing price data was used to construct indexes for the Fresh fruit and vegetables sub-group for one city using, in turn, each of the three approaches to derive indexes for the twenty three elementary aggregates (with index aggregation using weighted arithmetic means)." Woolford (1994; 8).

The result of this experiment was totally in line with the predicted inequalities established by Carruthers, Sellwood and Ward (1980) and Dalén (1992): the Jevons and Dutot indexes were approximately equal and the Carli indexes showed much higher rates of inflation. Since the ABS basically used the Carli index, this was an interesting result.

In the discussion of his paper, Woolford was not willing to concede that the Carli index might have an upward bias. He argued that as long as the base was not changed, the Carli index satisfied Walsh's (1901; 389) (1924; 506) Multiperiod Identity Test provided that the base remained constant. This is an interesting argument, which is also true. Consider an elementary index $P(p^1,p^2)$ that is defined over 5 periods. Suppose that $p^5 = p^1$ and the elementary index satisfies the identity test $P(p^1,p^1) = 1$. Then the sequence of 5 *fixed base index numbers* is the following one:

(25) $P(p^1,p^1) = 1;\ P(p^1,p^2)\ ;\ P(p^1,p^3)\ ;\ P(p^1,p^4)\ ;\ P(p^1,p^5) = P(p^1,p^1) = 1.$

Thus the Multiperiod Identity Test is satisfied if fixed base (or direct) index numbers are used. However, typically, the base period is changed due to sample updating or other considerations. Suppose this change of base occurs at the end of period 3. Then the new sequence of index values (with $p^5 = p^1$) is the following one:

(26) $P(p^1,p^1) = 1;\ P(p^1,p^2)\ ;\ P(p^1,p^3)\ ;\ P(p^1,p^3)P(p^3,p^4)\ ;\ P(p^1,p^3)P(p^3,p^5) = P(p^1,p^3)P(p^3,p^1) \geq 1$

where the last inequality in (26) follows if $P(p^1,p^2)$ is the Carli index. Thus as soon as the original base period is updated to a new base, the upward bias of the Carli formula shows up in a very obvious way.

Paper 8 was authored by Alain Saglio (1994). This paper made use of two years of scanner data collected by the Nielsen Corporation on sales of chocolate bars in France. Saglio computed an index of average prices and compared it to a corresponding Laspeyres index. He found that the average price index decreased 1.6% per year while the Laspeyres index decreased only 0.2% per year. He attributed the difference between the two indexes to substitution effects; French households switched to lower cost outlets and brands and the Laspeyres index was not able to pick up these substitution effects.

This paper introduced *scanner data* to the Ottawa Group of statisticians.[60] The implications of the existence of scanner data were far reaching. With the advent of scanner data, it became possible to compute

[60] Early studies by price statisticians using scanner data include Silver (1995), Dalén (1997), de Haan and Opperdoes (1997), Ioannides and Silver (1997), Lowe (1998) and Reinsdorf (1999). Hawkes and Piotrowski (2003) and Fixler, Greenlees, Fenwick, Lowe and Silver (2003) provide an excellent overviews on the use of scanner data to improve measurement of the CPI.

elementary indexes that utilize price and quantity information. One implication of the existence of scanner data was that the Turvey CPI Manual needed an update to accommodate this new source of information. Of course, it turned out that the use of scanner data led to new problems.[61]

There were two room documents which were published in the proceedings of the conference and listed on the Ottawa Group website. The room document by Carruthers, Sellwood and Ward (1994) has already been referred to in the discussion of Dalén's paper and it is a reprint of their 1980 journal article. These authors provided useful approximations to the differences between the various elementary indexes used in practice.

The second room document was Diewert (1995a). This paper covered the following five topics associated with the use of elementary indexes:

- The use of unit values at the very first stage of aggregation;
- The statistics literature on the properties of the various elementary indexes used in practice;
- The test approach to elementary indexes;
- Rough estimates of the possible biases in a CPI and
- The implications of scanner data for elementary indexes.

On the first and last topics, following Walsh (1901; 96) and Davies (1924; 183), Diewert (1995a; 22) advocated the use of unit values to represent the prices that would be used in an elementary index number formula. The companion quantity to the unit value price of a product is the total quantity purchased during the period by the group of consumers in scope for the index. Of course, in order to compute a unit value for a narrowly defined product, it is necessary to have price and quantity data for the transactions in scope for the product. This ties in with the last topic: the availability of scanner data means that one can now use a superlative index (or other target index that depends on prices and quantities) as a target index at the elementary level.

If the transactions in scope for the index were taken from retail outlets, then Diewert, like Saglio, argued for the use of shop specific unit values. But he realized that determining the scope of a unit value was not a straightforward choice:

"However, if individual outlet data on transactions were not available or were considered to be too detailed, then unit values for a homogeneous commodity over all outlets in a market area might form the lowest level of aggregation. Some further discussion on the concept of a unit value for a homogeneous commodity seems warranted. Saglio (1994) noted that the unit value or average price of a homogeneous commodity could be distinguished by: (i) its point of purchase (outlet effect); (ii) the various competing brands or product lines of the commodity that are being sold at an outlet; e.g., Cadbury versus Hershey chocolate bars (brand effect) and (iii) the various package sizes at which the commodity is sold (packaging effect). Thus finely classifying unit values on the basis of outlets, brands and packages should in principle be done, if the requisite data were available. However, it may turn out that empirically, some of this fineness of classification is not required." Diewert (1995a; 22).

Treating each product in an outlet in the same area as a separate product may not be required if the local outlets in scope sell their products at approximately the same prices. The finer the product classification is, the less will be product matches across time periods. The general conclusion that emerges is that the choice

---

[61] Two of the biggest problems are the *existence of product sales* and *rapid product turnover*. The reader should note that the elementary index number theory that has been presented thus far assumed that all prices were positive; i.e., there were no missing prices and unmatched products. Zero prices create obvious problems in bilateral price indexes.

of a product classification and the calculation of the corresponding monthly unit values is somewhat artistic.[62]

Diewert asked: how will statistical agencies access scanner data from retailers?

"In general, firms now process information on their costs and sales using computers so that summary information is available to managers on a monthly basis. Detailed information on prices and quantities could be extracted from this information base in many cases. In some cases, firms might be persuaded to provide information on prices and quantities to the Statistical Agency instead of filling out numerous forms.

The existence of private firms compiling detailed price and quantity information leads to an interesting dilemma for Statistical Agencies: (i) should the Agency buy the data from the information processing firm or (ii) should the Agency set up its own information processing subsidiary to compete against the private firm? Silver (1995) points out that the first alternative will lead to a loss of control by the Statistical Agency in its data collection activities. However, alternative (ii) may lead to charges that the Agency is providing unfair competition to the private sector. More public discussion on these issues seems to be required. However, it is clear that eventually, Statistical Agencies will be forced to join the electronic highway in one form or another." Diewert (1995a; 24).

Over time, it has been the case that many private firms have made their scanner data available to National Statistical Offices. In addition, marketing research firms have collected consumer price and quantity data from individual households and they have sold this information to economists who have constructed various consumer price indexes.

On topic 2 listed above, Diewert (1995a; 26-30) reviewed the relationships between various elementary indexes that were established by Carruthers, Sellwood and Ward (1994) and Dalén (1992).

On topic 3, Diewert (1995a; 6-16) extended Dalén's (1992) *axiomatic* or *test approach to elementary indexes*.[63] He took on board the tests that Dalén studied and added additional tests that were counterparts to Diewert's (1992) list of 20 tests that were satisfied by the Fisher index. These tests were for a bilateral price index of the form $P(p^0,p^1,q^0,q^1)$ and thus not all of them applied to a bilateral price index of the form $P(p^0,p^1)$. In the end, Diewert (1995a; 6-16) showed that the Jevons index $P_J(p^0,p^1)$ satisfied some 16 tests and the Dutot index $P_D(p^0,p^1)$ satisfied 15 of the 16 tests.[64]

On topic 4, Diewert summarized the available evidence on bias estimates as follows:

[62] Ivancic and Fox addressed the aggregation over unit values issue by running a weighted time product dummy hedonic regression over different product and store outlet classifications of the data. They summarized their results using Australian scanner data as follows: "A hedonic regression framework is used to test for item homogeneity across four supermarket chains and across stores within each of these supermarket chains. We find empirical support for the aggregation of prices across stores which belong to the same supermarket chain. Support was also found for the aggregation of prices across three of the four supermarket chains." Ivancic and Fox (2011a; 1). This paper was later published as Ivancic and Fox (2013a). Their basic idea is useful in terms of the recent literature on how to cluster scanner data. For example, suppose we have data on a product for a number of periods and we wish to cluster the data into more homogeneous sub-product groupings. We could run a time product hedonic regression and then look at the magnitudes of the product quality parameters. Products could be classified into subgroups depending on the closeness of their product parameters. If price and quantity data were available, then a weighted time product dummy regression should be run.

[63] See also Eichhorn (1978) on tests for indexes of the form $P(p^0,p^1)$.

[64] The Dutot index does not satisfy the commensurability test; i.e., it is not invariant to changes in the units of measurement. Thus it should only be used when the products in scope are measured in the same units.

"Summarizing the empirical evidence reviewed in this section and the previous one, we see that it is likely that in recent years, a typical official CPI has a .2% per year commodity substitution bias, a .25% per year outlet substitution bias, a linking bias of perhaps .1% per year and a new goods bias of at least .25% per year; i.e., an upward bias of at least .8% per year. If the Statistical Agency is also making use of a biased elementary price index formula, this will add an additional upward bias to the official index. The reader will note that all of the 5 above sources of bias were regarded as being additive, an assumption which is probably approximately correct." Diewert (1995a; 35).

These bias estimates were very similar to the bias estimates made by the Boskin Commission in 1996. This is not too surprising since we were all using the same raw materials on estimates of bias. Looking back at the above estimates, they seem reasonable for the time period except that the estimate for outlet substitution was probably too high.

At the end of the paper, Diewert made the following recommendations:

- "Statistical Agencies should follow the emphatic advice of Irving Fisher [1922; 29-30] and avoid the use of the Carli arithmetic mean of price relatives formula (2) to form elementary price aggregates.
- (ii) If information on quantities is not available at the elementary or basic level, either the geometric price index (4) advocated by Jevons or the average price index (9) suggested by Dutot should be used. Axiomatic justifications for these two indexes were provided in sections 3 and 4 and (weak) economic justifications were presented in section 5.
- (iii) At the level of the individual outlet, the best elementary average price for a homogeneous commodity would seem to be its unit value: the value of units sold during the sample period divided by the total quantity sold. If outlet unit values are available, then in aggregating over outlets, there is no need to restrict ourselves to using the Jevons or Dutot formulae to construct elementary prices. From the viewpoint of economic theory, it seems preferable to use the Fisher ideal price index in this second stage of elementary aggregation.
- (iv) Values and quantities should be sampled rather than just prices. Sampling values and quantities will greatly reduce the new introductions bias.
- (v) Statistical Agencies should consider either purchasing electronic point of sale data from firms currently processing these data or the Agencies should set up Divisions which would compete in this area.
- (vi) Recent economic history will have to be rewritten in view of the substantial outlet substitution and elementary price index biases that Reinsdorf and Moulton have uncovered." Diewert (1995a; 37-38).

Jacob Ryten in his closing remarks noted the need to have further meetings of the Ottawa Group and he provided the following list of ten topics that could be discussed in future meetings:

- "How to establish price indices for difficult areas such as insurance and gambling fees payable to a state agent.
- Product quality adjustment and the use of hedonic methods.
- The necessary steps for the harmonization of CPI.
- Clarifying concepts (e.g. What is a cost of living index?).
- Treatment of new products and outlets in price indices.
- Treatment of durable goods in the CPI.
- Index formulae at macro level, including the linking problem.
- Organization and techniques related to price surveys.
- Linkage between temporal and spatial price comparisons.
- Problem of seasonality and price indices.
- Measurement of inflation." Ottawa Group (1994).

The above topics were indeed discussed in the next 16 meetings of the Ottawa Group. The above topics are still with us today.

It can be seen that the first meeting of the Ottawa Group presented evidence that the Turvey paradigm for a CPI needed to be updated. Thus the stage was set for the production of a revised CPI Manual. This is the topic of the next section of this paper.

5. The First Ottawa Group CPI Manual

The need for a revised CPI Manual was recognized by members of the Ottawa Group in 1998:

"In September 1998 a meeting of international organisations was convened to discuss the possibility of updating a manual on the compilation of CPI. This meeting agreed that such a manual should be developed. As part of the process and as a member of the newly formed Inter-secretariat Working Group on Price Statistics (IWGPS), the OECD undertook to co-ordinate and maintain the chapter outline for the new manual on CPI. This document represents the latest version of the CPI manual outline." Obst (1999).

However, it took six years before the *Consumer Price Index Manual: Theory and Practice* was finally finished in 2004.[65]

The editor of the 2004 Manual was Peter Hill. Here is a listing of the chapters and their authors:

- Preface: Peter Hill, Paul Armknecht and W. Erwin Diewert
- Reader's guide: Peter Hill
- 1 An introduction to consumer price index methodology: Peter Hill
- 2 Uses of consumer price indices: Peter Hill
- 3 Concepts and scope: Peter Hill and Fenella Maitland-Smith
- 4 Expenditure weights and their sources: Valentina Stoevska and Carsten Boldsen
- 5 Sampling: Jorgen Dalen, A. Sylvester Young and Bert Balk
- 6 Price collection: David Fenwick
- 7 Adjusting for quality change: Mick Silver
- 8 Item substitution, sample space and new products: Mick Silver
- 9 Calculating consumer price indices in practice: Carsten Boldsen and Peter Hill
- 10 Some special cases: Keith Woolford, David Fenwick
- 11 Errors and bias: John Greenlees and Bert Balk
- 12 Organization and management: David Fenwick
- 13 Publication, dissemination and user relations: Tom Griffin
- 14 The system of price statistics Kimberly Zieschang
- 15 Basic index number theory: W. Erwin Diewert
- 16 The axiomatic and stochastic approaches to index number theory: W. Erwin Diewert
- 17 The economic approach to index number theory: The single-household case: W. Erwin Diewert
- 18 The economic approach to index number theory: The many-household case: W. Erwin Diewert
- 19 Price indices using an artificial data set: W. Erwin Diewert
- 20 Elementary indices: W. Erwin Diewert

[65] See ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004).

- 21 Quality change and hedonics: Mick Silver
- 22 The treatment of seasonal products: W. Erwin Diewert
- 23 Durables and user costs: W. Erwin Diewert .

The theoretical chapters in the 2004 Manual simply reflected the state of index number theory in 1994 that was covered in sections 2-4 above with some additional detail. Thus the Turvey Manual touched on methods for dealing with quality change by discussing hedonic regressions but Chapter 21 looked at methods for dealing with quality change in more systematic fashion. Similarly, The Turvey Manual discussed the treatment of seasonal products and user costs for housing but Chapters 22 and 23 discussed these topics in much more detail. In Chapter 22, the use of Rolling Year Mudgett-Stone indexes and the use of year over year monthly indexes (with month specific weights) was suggested as providing checks on the seasonal product components of the official month to month CPI. However, the Chapter ended up on a pessimistic note with the following conclusion on dealing with strongly seasonal products in a month to month CPI:

"It is evident that more research needs to be carried out on the problems associated with the index number treatment of seasonal commodities. There is, as yet no consensus on what is best practice in this area." ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004; 417).

Chapter 23 on the treatment of durable goods in general (and of Owner Occupied Housing in particular) provided a more comprehensive treatment of possible approaches to the treatment of durable goods than was provided in the Turvey Manual. In particular, the 2004 Manual noted that the user cost for a dwelling unit should be decomposed into separate user cost terms for the land and structure components of the property. Chapter 23 also noted that the acquisitions approach would tend to give a much smaller weight to OOH than the user cost and rental equivalence approaches. Both Manuals came to the conclusion that more than one approach to the treatment of OOH was required in order to address different purposes for a CPI. However, it is interesting that National Statistical Offices have not provided alternative indexes for OOH on regular basis.

6. The Recent Ottawa Group CPI Manuals

After the publication of the 2004 CPI Manual, some problems with the advice in this Manual started to emerge. The problem of choosing between fixed base or direct indexes versus chained indexes was discussed in the 2004 Manual:

"Hill (1993; 388), drawing on the earlier research of Szulc (1983) and Hill (1988; 136–137), noted that it is not appropriate to use the chain system when prices oscillate or bounce. This phenomenon can occur in the context of regular seasonal fluctuations or in the context of price wars. However, in the context of roughly monotonically changing prices and quantities, Hill (1993; 389) recommended the use of chained symmetrically weighted indices (see paragraphs 15.18 to 15.32). The Fisher and Walsh indices are examples of symmetrically weighted indices." ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004; 281).

The 2004 Manual went on to give some further guidance on this topic:

"It is possible to be a little more precise about the conditions under which to chain or not to chain. Basically, chaining is advisable if the prices and quantities pertaining to adjacent periods are *more similar* than the prices and quantities of more distant periods, since this strategy will lead to a narrowing of the spread between the Paasche and Laspeyres indices at each link. Of course, one needs a measure of how similar the prices and quantities are pertaining to two periods. The similarity measures could be relative ones or absolute ones. In the case of absolute comparisons, two vectors of the same dimension are similar if they are identical and dissimilar otherwise. In the case of relative comparisons, two vectors are similar if they are proportional and dissimilar if they are non-proportional. Once a similarity measure has been defined, the prices and quantities of each period can be compared to each other using this measure, and a "tree" or path that links all of the observations can be constructed where the most similar observations are compared with each other using a bilateral index number formula." ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004; 281).

The 2004 Manual referred to a 2002 discussion paper by Diewert for explicit measures of absolute and relative price dissimilarity.[66] Normally, prices in adjacent periods would tend to be the most similar, so in practice, it was thought that chaining would normally be "best".

Ivancic and Fox questioned the above advice in a 2011 Ottawa Group paper:

"Chaining is used in index number construction to update weights and link new items into an index. However, chained indexes can suffer from, sometimes substantial, drift. The Consumer Price Index Manual (ILO, 2004) recommends the use of dissimilarity indexes to determine when chaining is appropriate. This study provides the first empirical application of dissimilarity indexes in this context. We find that dissimilarity indexes do not appear to be sufficient to resolve the issue of when to chain." Ivancic and Fox (2011b; 1).

Of course, it was well known that chaining a Carli, Laspeyres or Paasche index often resulted in massive chain drift if prices "bounce". However, Ivancic and Fox used superlative indexes to do the bilateral linking. The amount of chain drift in using superlative chained indexes was thought to be small. They summarized their empirical results as follows:

"The ILO (2004) states that chaining is appropriate when 'the prices and quantities pertaining to adjacent periods are more similar than the prices and quantities of more distant periods' (p. 281). Results are presented in tables 1 and 2. We find that when the chained dissimilarity indexes are compared with their direct counterparts there are very few circumstances — in total only 9 out of 76 — where both the direct price and quantity dissimilarity indexes are less than the chained price and quantity dissimilarity indexes. Chaining was found to be appropriate in the majority of cases — 47 out of 76 cases. In the remaining 20 cases there was no clear evidence on the issue of chaining." Ivancic and Fox (2011b; 5).

For their data set, chaining was appropriate for 47 out of 76 cases, satisfactory for 20 cases and not satisfactory for 9 cases. Their results indicated that chaining superlative indexes could lead to chain drift in some cases.

A very important prior paper appeared in Kevin Fox's Economic Measurement Workshop in 2008 by Jan de Haan (2008). He showed that the chain drift problem could be a big one, even if superlative indexes were used to perform the bilateral links. The scanner data he used were based on a sample of more than 100 supermarkets that belonged to one of the largest supermarket chains in the Netherlands. His data set contained weekly observations on turnover and quantities sold for hundreds of detergents and covered the period week 1 in 2005 to week 35 in 2008 (191 weeks). The average number of matched products in each consecutive two week period was 53 and the range of matched products varied between 43 and 63. The range of product price variation and the corresponding changes in sales of the products was shocking. Figure 1a in his paper plotted the price of Product XXX Tablets over the 191 weeks and Figure 1b plotted the

[66] This paper was eventually published as Diewert (2009). A problem with Diewert's suggested measures of relative price dissimilarity is that he assumed that all prices were positive (or that the dissimilarity measure could only be applied to products that were present in the two periods being compared).

corresponding weekly quantities sold. The "normal" price of Product XXX was about 6.5 Euros. The product went on sale for 12 of the 191 weeks with a sale price that was approximately one half of the regular price. As the price went down, the sales of Product XXX went from virtually 0 to over 3000 for 7 of the 12 weeks that XXX Tablets went on sale. The stockpiling problem that Triplett (2003) identified was a real problem! Needless to say, the chain drift problem with the Dutch data set was massive: de Haan showed that the chained Fisher and chained Törnqvist indexes ended up in week 191 at 4.95% and 7.43% of the week 1 value of these indexes. The chained Jevons index using bilateral matched products at each link ended up at 76.65% of the week 1 index level.[67]

Typically the chain drift that results from the use of a chained superlative index will be downward as was the case with de Haan's data. Downward chain drift occurs due to purchasers stockpiling goods when they are on sale. But upward chain drift can also occur. Feenstra and Shapiro (2003) found upward chain drift in the Törnqvist formula using a scanner data set. Persons (1928; 100-105) had an extensive discussion of the chain drift problem with the Fisher index and he gave a numerical example on page 102 of his paper that showed how upward chain drift could occur.[68]

De Haan discussed several methods that he thought could be applied to reduce the chain drift problem. The simplest method is to simply increase the length of the period:

"Extending the observation period from a week to a month reduces the 'noise' in the data and most likely lowers or possibly removes chain drift. Figure 7 shows monthly chained (matched goods) price index numbers. Clearly they are much more reasonable than the weekly chained index numbers shown earlier. According to the Jevons index the prices of detergents declined by approximately 10% whereas the Fisher and Törnqvist suggest that prices did not change significantly. Despite the smoothing of the data the Fisher and Törnqvist indexes are still highly volatile, especially during the first half of the period." Jan de Haan (2008; 18).

Thus increasing the length of the period from a week to a month did indeed eliminate the tremendous downward chain drift. But will we reduce potential chain drift even more if we increase the length of the period to a quarter?

"As mentioned earlier, most statistical agencies compile CPIs on a monthly basis. This does not mean that monthly chaining would be the optimal choice. Figure 9 shows what happens when we extend the observation period to a quarter and employ quarterly chaining. Measured by the quarterly chained Fisher and Törnqvist, prices have risen by some 20% over the entire period, largely as a result of strong increases during the third and fourth quarter. This is a lot more than measured by the monthly chained Fisher and Törnqvist indexes. Should we conclude that the monthly chained indexes still exhibit downward drift or alternatively that the quarterly chained indexes exhibit upward drift? A priori we expect the use of quarterly unit value and quantity data to remove most of the short term 'noise' so that quarterly chaining of Fisher and Törnqvist indexes leads to approximately unbiased results. But then we would expect chained Fisher and Törnqvist indexes to be similar, which is not the case here: at the end of the period, in quarter 14, the chained Törnqvist is 5.41 points below the chained Fisher." Jan de Haan (2008; 21-22).

---

[67] The chained Jevons index no longer satisfies the circularity test when there are missing products but it obviously is less subject to chain drift than the superlative chained indexes that use matched products in each bilateral link.

[68] Persons (1928; 102) explained that it was *incomplete adjustment* that caused the Fisher chained index to climb above the corresponding fixed base index in his example. Ludwig von Auer (2019) (2024) proposed a similar theory. Note that Triplett (2003) noted that *habit formation* (a form of incomplete adjustment) along with *stockpiling behavior* caused difficulties with the standard economic approach to index number theory. Triplett's third difficulty with the standard approach was the existence of *search costs* for consumers.

The above results are troubling. The difference between the quarterly and monthly indexes may be due to unit value bias; i.e., high priced products get more weight in the unit value for a period than low priced products. A quarter may be too long a period to use a unit value as a representative price for the quarter.[69]

Another obvious solution to the chain drift problem is to use a fixed base.[70] However, it turns out that for many elementary strata (like detergents), there is a rapid turnover of new and disappearing products. Thus as time goes on, there are fewer and fewer product matches of current period prices with the corresponding base period prices and thus the resulting fixed base index becomes more unreliable. This is the *de Haan problem*: how to find an index number formula that is free from chain drift (i.e., satisfies the circularity test) but also can deal with the problem of product churn and missing prices.

A possible solution to the de Haan problem emerged in the Ottawa Group paper by Ivancic, Fox and Diewert (2009).[71] These authors suggested the use of a *multilateral index* over a window of consecutive periods.[72] Within the window, the indexes satisfy the circularity test and as a result, there can be no chain drift within the window of time periods. Ivancic, Fox and Diewert suggested the use of two multilateral indexes, the GEKS index[73] and the Weighted Time Product Dummy index.[74]

One way of explaining how the GEKS index works as follows. Suppose we have price and quantity data for T periods. Let $P_F(r,t)$ be the *matched product bilateral Fisher index* which compares the prices of period t to the prices of period r (the base period). One could choose period 1 as the base period and form the following sequence of price levels relative to period 1: $P_F(1,1) = 1$, $P_F(1,2)$, $P_F(1,3)$, ..., $P_F(1,T)$. But one could also use period 2 as the base period and use the following sequence of price levels to measure the price level of each period relative to the period 2 level: $P_F(2,1)$, $P_F(2,2) = 1$, $P_F(2,3)$, ..., $P_F(2,T)$. Each period could be chosen as the base period and thus we end up with T alternative series of Fisher price levels. Since each of these sequences of price levels is equally plausible, Gini (1931) suggested that it would be appropriate to take the geometric average of these alternative price levels in order to determine the final set of price levels. Thus the (unnormalized) *GEKS price level*s for periods t = 1,2,...,T are defined as follows:

(27) $p_{GEKS}^t \equiv [\prod_{r=1}^T P_F(r,t)]^{1/T}$.

Note that all time periods are treated in a symmetric manner in the above definition. The GEKS *normalized price level* P_{GEKS}^t are obtained by normalizing the above price levels so that the period 1 index is equal to 1. Thus we have the following definition for the period t normalized price levels P_{GEKS}^t:

(28) $P_{GEKS}^t \equiv p_{GEKS}^t/p_{GEKS}^1$; t = 1,…,T.

---

[69] However, de Haan's result may not be "representative". Ivancic, Fox and Diewert (2009) found that volatility of a chained superlative index decreased as they went from weekly to monthly to quarterly indexes. When they aggregated over stores, the volatility also decreased.

[70] Recall our discussion of Woolford (1994).

[71] This paper was eventually published as Ivancic, Diewert and Fox (2011). The published paper dropped the material on the weighted Time Product Dummy multilateral method.

[72] The use of multilateral indexes in the time series context dates back to Persons (1921) and Fisher (1922; 297-308), Gini (1931) and Balk (1980) (1981).

[73] See Gini (1931), Eltetö and Köves (1964) and Szulc (1964).

[74] Rao (1995) (2005; 574) was the first to consider this model using expenditure share weights. However, Balk (1980; 70) suggested this class of models much earlier using somewhat different weights. This model with missing prices is discussed in some detail in Diewert (2004) (2023).

If one chooses the matched product Törnqvist-Theil bilateral index number $P_T(r,t)$ instead of the Fisher formula $P_F(r,t)$ in definitions (27) and (28), one obtains the (normalized) *GEKS-Törnqvist price level*s P_{GEKS-T}^t for t = 1,…,T.[75]

Explaining how the Weighted Time Product Dummy multilateral method works when there are missing products is more complicated.[76] Recall the notation that was introduced in section 2 above for the period t price and quantity vectors, $p^t \equiv [p_{t1},\ldots,p_{tN}]$ and $q^t \equiv [q_{t1},\ldots,q_{tN}]$, and the period t expenditure share vector, $s^t \equiv [s_{t1},\ldots,s_{tN}]$. If product n in period t is missing, then define $p_{tn} \equiv q_{tn} \equiv s_{tn} \equiv 0$. Suppose we have a window of T periods. Let S(t) be the set of products n that are purchased in period t and let $S^*(n)$ be the set of time periods t such that product n was purchased. We assume that each product is sold in at least one time period and there is some product overlap. The underlying hedonic model is given by the approximate equalities (29):

$$(29)\ p_{tn} \approx \pi_t\alpha_n\ ; \qquad t = 1,\ldots,T;\ n\in S(t).$$

Thus the model assumes that product prices vary in a proportional manner (approximately) over time. The parameter π_t is interpreted as the period t *price level* for the products in scope and the parameter α_n is a *quality adjustment parameter* which is equal to the (relative) utility that an extra unit of product n gives to purchasers of the products in scope. Take the logarithms of both sides of equations (29) and add error terms ε_{tn}. We obtain the following linear regression model:

$$(30)\ \ln p_{tn} = \rho_t + \beta_n + \varepsilon_{tn}\ ; \qquad t = 1,\ldots,T;\ n\in S(t).$$

where $\rho_t \equiv \ln\pi_t$ for t = 1,…,T and $\beta_n \equiv \ln\alpha_n$ for n = 1,…,N. Estimates for the parameters ρ_t and β_n can be obtained by solving the following least squares minimization problem:

$$(31)\ \min_{\rho,\beta}\ \{\Sigma_{t=1}^T\ \Sigma_{n\in S(t)}\ [\ln p_{tn} - \rho_t - \beta_n]^2\}.$$

In order to obtain a unique solution to (31), we need to add a normalization on the parameters. A natural normalization in the present context is to set the period 1 price level π_1 equal to 1. Thus we set the logarithm of π_1 equal to 0:

$$(32)\ \rho_1 = 0.$$

We assume that there is a sufficient amount of product overlap between the periods for the solution to (31) with the normalization (32) , ρ_1^* (= 0), $\rho_2^*,\ldots,\rho_T^*$, $\beta_1^*,\ldots,\beta_N^*$, to be unique. With these estimates in hand, we can obtain estimates for the price levels π_t and quality adjustment parameters α_n by exponentiating the ρ_t^* and β_n^*:

$$(33)\ \pi_t^* \equiv \exp[\rho_t^*]\ ;\ t = 1,\ldots,T;\ \alpha_n^* \equiv \exp[\beta_n^*]\ ;\ n = 1,\ldots,N.$$

[75] This multilateral index is also known as the CCDI index. This multilateral index was used by Inklaar and Diewert (2016). It is an adaptation of the distance function approach used by Caves, Christensen and Diewert (1982) to the price index context.
[76] For a much more complete description of the model with references to the literature, see Diewert (2023).

Define period t expenditures on the N products as $e^t \equiv p^t \cdot q^t = \Sigma_{n\in S(t)} p_{tn}q_{tn}$ for t = 1,…,T. The *Time Product Dummy period t aggregate price and quantity levels*, $P_{TPD}{}^t$ and $Q_{TPD}{}^t$ are defined as follows:

(34) $P_{TPD}{}^t \equiv \pi_t{}^*$; $Q_{TPD}{}^t \equiv e^t/P_{TPD}{}^t$; t = 1,…,T.

The above definitions do not make use of the estimated $\alpha_n{}^*$. However, if quantity data are available, it is possible to use these estimated quality adjustment parameters to *directly* define the period t aggregate quantity level, $Q_{TPD}{}^{t*}$, and then *indirectly* define the corresponding period t price level, $P_{TPD}{}^{t*}$, by deflating period t expenditures e^t by $Q_{TPD}{}^{t*}$:

(35) $Q_{TPD}{}^{t*} \equiv \Sigma_{n=1}{}^N \alpha_n{}^* q_{tn} = \Sigma_{n\in S(t)} \alpha_n{}^* q_{tn}$; $P_{TPD}{}^{t*} \equiv e^t/Q_{TPD}{}^{t*}$; t = 1,...,T.

If the errors ε_{tn} in the linear regression (30) are all equal to 0, so that equations (29) hold exactly, then $P_{TPD}{}^t$ will equal $P_{TPD}{}^{t*}$ and $Q_{TPD}{}^t$ will equal $Q_{TPD}{}^{t*}$ for all t. However, in general, the following inequalities will hold:[77]

(36) $P_{TPD}{}^{t*} \leq P_{TPD}{}^t$; $Q_{TPD}{}^{t*} \geq Q_{TPD}{}^t$; t = 1,...,T.

The Time Product Dummy hedonic regression model described above has the advantage that it can be implemented using price information alone.[78] However, it has the disadvantage that it does not take into account the economic importance of each product. If price and quantity information are available, then it is better to make use of this additional information which leads us to the *Weighted Time Product Dummy hedonic regression model*. We continue to assume that prices are approximately proportional over time; i.e., assume that assumptions (29) hold. Consider the following weighted least squares minimization problem:

(37) $\min_{\rho,\beta}\ \{\Sigma_{t=1}{}^T \Sigma_{n\in S(t)} s_{tn}[\ln p_{tn} - \rho_t - \beta_n]^2\}$.

As was the case with the minimization problem defined by (31), solutions to (37) will not be unique. Thus again, impose the normalization $\rho_1 = 0$ on the unknown parameters which appear in (37). Again assuming a sufficient degree of product overlap, the solution to (37) with the normalization (32), which we denote by $\rho_1{}^*$ (= 0), $\rho_2{}^*,\ldots,\rho_T{}^*$, $\beta_1{}^*,\ldots,\beta_N{}^*$, will be unique. With these new estimates in hand, repeat the algebra in equations (33)-(36) and we obtain the directly estimated Weighted Time Product Dummy period t price and quantity levels, $P_{WTPD}{}^t$ and $Q_{WTPD}{}^t$, and the indirectly estimated Weighted Time Product Dummy period t price and quantity levels, $P_{WTPD}{}^{t*}$ and $Q_{WTPD}{}^{t*}$. Ivancic, Fox and Diewert (2009) estimated the directly estimated Weighted Time Product Dummy indexes using their Australian data set.

The multilateral methods described above have a property that is troublesome: if a product is present in only one period of the window of T observations, then it has no effect on the resulting price levels.[79] This means that new products entering the marketplace in period T have no effect on the multilateral indexes.

---

[77] These inequalities were derived in a more general model by de Haan (2004) (2010) and de Haan and Krsinich (2018; 763); see Diewert (2023) for more details.

[78] The price levels $P_{TPD}{}^t$ in equations (34) can be estimated by using price information alone.

[79] That this property holds for the GEKS and CCDI multilateral indexes is obvious since these indexes depend on bilateral indexes which in turn use only the prices of matched products in the two periods under consideration. If a product appears in only one period, it cannot be a matched product. That this property holds for the TPD and WTPD price levels is not so obvious; it was first established by Diewert (2004).

The Australian data set was described as follows:

"We use a scanner data set collected by A.C. Nielsen, which contains information on four supermarket chains located in one of the major capital cites in Australia. In total, over 100 stores are included in this data set with these stores accounting for approximately 80% of grocery sales in this city; …. The data set contains 65 weeks of data, collected between February 1997 and April 1998. Information on 19 different supermarket item categories, such as bread, biscuits and soft drinks are included. A large number of observations on transactions exist for all item categories, with a minimum of 225,789 observations for the item category "butter" and a maximum of 2,639,642 observations for the item category "juices". An observation here refers to the average weekly price (weekly unit value) and total weekly quantity sold of each item transacted in each store in each week." Ivancic, Fox and Diewert (2009; 12).

The authors demonstrate the upward chain drift of the Laspeyres index for their data set and show that increasing the length of the period from a week to a month to a quarter did reduce the upward bias:

"The impact of time aggregation is extremely pronounced when chained indexes are used. This is particularly true for the Laspeyres index, where a number of price change estimates appear to explode as the frequency of chaining increases. For example, table 5 shows that Laspeyres price change estimates for the item category toilet paper based on quarterly, monthly and weekly time aggregation (with no item aggregation over stores) range from a somewhat reasonable (106.71−100 =) 6.71% (quarterly, fixed basket) to a massive (11,955−100 =) 11,855% (weekly, fixed basket) over the 15 month period." Ivancic, Fox and Diewert (2009; 13).

The authors also demonstrate the likely downward bias of chained Fisher indexes:

"The Fisher index appears to be relatively less affected by time aggregation than the Laspeyres and Paasche index. Despite this, even the Fisher index shows a degree of variation which seems to be a cause for concern. For example, from Table 7, the Fisher flexible-basket chained estimates of price change for the item category toilet paper (no item aggregation over stores) were calculated at (100.43−100 =) 0.43%, (98.61−100 =) −1.39% and (79.86−100 =) −20.14% for quarterly, monthly and weekly time aggregation respectively. Overall, our results show that with item aggregation over stores and using the flexible-basket chained Fisher index, on average the absolute difference between weekly and quarterly price change estimates is approximately 8%. When we look at indexes where items have been disaggregated over stores (Table 7), the average absolute difference increases to approximately 14%." Ivancic, Fox and Diewert (2009; 13-14).

The authors summarized their results as follows:

"What conclusions can be drawn from the results of our experiments with scanner data? Our tentative conclusions are as follows:

- The use of weekly chained index numbers, even those based on superlative index number formulae, is not recommended due to the erratic nature of the resulting indexes.
- Fixed base or direct comparisons of a current period with a base period seem to give reasonably reliable results, at least using monthly or quarterly data. However, these fixed base comparisons suffer from the problems associated with new and disappearing goods; i.e., over time, it becomes increasingly difficult to match items." Ivancic, Fox and Diewert (2009; 17).

Here is the solution that they offered to the chain drift problem that affected monthly superlative indexes: choose a window length of 13 months (so that strongly seasonal products enter the indexes) and compute a GEKS multilateral index for the initial 13 months. If the index is not revisable, then these 13 price levels become the "permanent" index levels. When the data for month 14 become available, compute a new set of GEKS multilateral price levels for months 2 to 14. Use the ratio of the month 14 price level to the month 13 price level to update the month 13 level of the permanent index. When the month 15 data become available, compute a new set of GEKS multilateral price levels for months 3 to 15. Use the ratio of the month 15 price level to the month 14 price level to update the month 14 level of the permanent index. This is their *Rolling Year GEKS methodology*. They compared their RYGEKS index with the GEKS index that was constructed over the entire 15 month data period:

"It is of interest to compare the GEKS and RYGEKS series as this will give us some indication of whether the RYGEKS index is sensitive to the length of window chosen. It will also indicate whether a 13 month window is long enough to provide us with a stable price series. We compare the monthly GEKS price indexes with our RYGEKS indexes. The results show that there is very little difference between the standard GEKS and RYGEKS series, with plots of the GEKS and RYGEKS series sitting virtually on top of each other." Ivancic, Fox and Diewert (2009; 25).

The authors also used Weighted Time Product Dummy price levels to replace their GEKS price levels and found that the resulting series approximated the corresponding GEKS series fairly well. Their overall conclusion was that there was little difference between the two methods:

"The results indicate that statistical agencies that use scanner data may be able to use either the GEKS or CPD approach to estimate drift free estimates of price change with some confidence as both methods give very similar results. One issue to be considered is that of temporal fixity. With traditional multilateral index number methods, index numbers are generated not only for the current period but also for all past periods in the domain of definition of the multilateral index. Thus a drawback of traditional multilateral indexes applied in the time series context is that they violate temporal fixity, which means that when a time period is added to the multilateral index the index number results for previous periods may change. With our recommended "rolling year GEKS" approach we avoid this problem of having to make constant revisions to past values of the index as the data for a new period become available." Ivancic, Fox and Diewert (2009; 30).

Ivancic, Diewert and Fox (2011) suggested that the movement of the rolling window indexes for the last two periods in the new window be linked to the last index value generated by the previous window. However Krsinich (2016) suggested that the movement of the indexes generated by the new window be linked to the previous window index value for the second period in the previous window. Krsinich called this a *window splice* as opposed to the IDF *movement splice*. De Haan (2015; 27) suggested that perhaps the linking period should be in the middle of the old window which the Australian Bureau of Statistics (2016; 12) termed a *half splice*. Ivancic, Diewert and Fox (2011; 33) suggested that the *geometric average* of all links for the last period in the new window to the observations in the old window could be used as the linking factor. Diewert and Fox (2021) looked at these alternative linking strategies *Average* or *mean linking* seems to be the safest strategy from a statistical point of view. Of course, another strategy would be to use an *ever expanding window*.[80]

Unfortunately, the Rolling Window methodology does not completely eliminate the chain drift problem; the method does not pass the circularity test as soon as a new window is added. The Rolling Window study by Ivancic, Fox and Diewert utilized data for only 16 months. With longer time series, it seems that the chain drift problem can still show up; see Fox, Levell and O'Connell (2024). However, the Rolling Window methodology certainly reduces the chain drift problem.

In order to deal with the availability of scanner data and the associated chain drift problem that could sometimes occur using chained superlative indexes, it became clear that the 2004 CPI Manual needed to be

[80] This method is applied in the predicted share price similarity linking method that is described in Diewert (2023) and in Diewert and Shimizu's (2024) application of an ever expanding window version of the Weighted Time Product Dummy multilateral method. The Appendix to this paper also deals with this method of linking.

updated. Thus in 2015, work began on an updated Manual which appeared in 2020: *Consumer Price Index Manual: Concepts and Methods 2020*.[81] The editor of this new manual was Brian Graf and the lead institution was the IMF. The purpose of the 2020 Manual was explained as follows:

"The Manual is intended for the benefit of agencies that compile CPIs, as well as users of CPI data. It explains in some detail the methods that are recommended for use to calculate a CPI. A separate companion publication, *Consumer Price Index Theory*, explains the underlying economic and statistical theory on which the methods are based." IMF/ILO/UNECE/Eurostat/OECD/The World Bank (2020; xi).

The companion Manual, *Consumer Price Index Theory*, has not been finalized due to the rapid pace of new CPI theoretical research which has led to revisions which incorporate the new research. Draft chapters of the CPI Theory Manual are available on the IMF CPI website.

The 2020 CPI Manual was intended as an updated version of the 2004 CPI Manual with the addition of some new chapters:

"The current Manual represents an update of the 2004 Manual published by the International Labour Organization (ILO). Individual authors were recruited to review and update each chapter. Some chapters required extensive updating and rewriting, while others needed only minimal updating from the 2004 version of the manual. Two new chapters have been added on scanner data and updating Consumer Price Index (CPI) weights. … The Manual benefited from the experience of several experts responsible for updating the individual chapters. The authors included: Paul Armknecht, IMF (Retired); Corinne Becker, Swiss FSO; David Fenwick, UK ONS (Retired); Jan de Haan, Statistics Netherlands; Brian Graf, IMF, editor of the Manual; Claude Lamboray, Eurostat; Maria Mantcheva, IMF (Retired); Valentina Stoevska, ILO; Marcel van Kints, ABS; Mick Silver, IMF (Retired); and Jan Walschots, Statistics Netherlands (Retired). The Manual has also benefited from valuable contributions by many other experts who served as primary reviewers for individual chapters, including: Badria Al-Aadi, NCSI (Oman); Carsten Boldsen, UNECE; Rob Cage, Bureau of Labor Statistics (United States); Barra Casey, CSO (Ireland); Ronald Johnson, Expert (external reviewer); Patrick Kelly, Statistics South Africa; Brent Moulton, Expert (external reviewer); Ragnhild Nygaard, Statistics Norway; Niall O'Hanlon, IMF; Federico Polidoro, ISTAT (Italy); Rafael Posse, INEGI (Mexico); Yunita Rusanti (BPS-Statistics Indonesia); and V. Thuy, GSO (Vietnam)." IMF/ILO/UNECE/Eurostat/OECD/The World Bank (2020; xv).

There are 14 chapters and 7 Appendices in the 2020 Manual. Most of the chapters follow up on similar chapters in the 2004 Manual but Chapter 10 on Scanner Data is a valuable entirely new chapter. This chapter also covers the use of multilateral indexes.

Here is a listing of the seven Appendices which also contain new material:

- The Harmonised Index of Consumer Prices (European Union);
- Classification of Individual Consumption According to Purpose (COICOP 1999);
- Classification of Individual Consumption According to Purpose (COICOP 2018);
- Resolution Concerning Consumer Price Indices Adopted by the Seventeenth International Conference of Labour Statisticians;
- Spatial Comparisons of Consumer Prices, Purchasing Power Parities and International Comparisons;
- Some Basic Index Number Formulas and
- The Consumer Price Index Research Agenda.

[81] See IMF/ILO/UNECE/Eurostat/OECD/The World Bank (2020). We will refer to this document as the *2020 Manual*.

We will discuss the materials in Appendix 7 on the Consumer Price Index Research Agenda in the following section.

7. Looking Ahead

Appendix 7 of the 2020 CPI Manual has an excellent listing of research topics that Ottawa Group researchers could (and should) focus on for the future:[82]

- Scanner data;
- Web scraped prices;[83]
- Price updating of expenditure weights;[84]
- The use of administrative data to form indexes;
- Plutocratic versus democratic weighting of households;
- The use of credit card data to form household specific price indexes;[85]
- Calculating elementary indexes using expenditure weights;
- Quality adjustment;[86]
- The treatment of seasonal products;
- The use of target price indices for the CPI;[87]
- On the choice of formula for calculating higher level price Indices;[88]
- The use of long-term and short-term Links;[89]

[82] See pages 455-459 in IMF/ILO/UNECE/Eurostat/OECD/The World Bank (2020).
[83] See Cavallo (2017) on this topic.
[84] Recall the discussion around equations (12)-(17) in section 3 on price updating of expenditure weights.
[85] See Cavallo (2020).
[86] "Quality adjustment is a cross-cutting issue that continues to pose challenges. Quality-adjustment issues have been noted previously, but there is a general need to provide better guidance on the treatment of quality changes. In particular, NSOs continue to struggle with measuring quality changes for clothing, cars, telecommunication equipment, multipurpose information technology devices, computers, and, in general, products with high churn. Sharing of practical experiences implementing methods and best practices is needed." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 456).
[87] "It is useful to have a measurable target index for the CPI. The target index can be used as guidance for deciding calculation methods and practices for the regular CPI and for measuring potential bias. Empirical research can address issues including identifying potential target indices (for example, Walsh, Fisher, Törnqvist, or Constant Elasticity of Substitution [CES]) and how to apply these formulas in practice." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 457).
[88] "Arithmetic aggregation is used by almost all countries for calculation of higher-level price indices. Is this the best solution? What are the alternatives, with regard to geometric aggregation or aggregation by other types of averages or by use of indices that apply explicit estimates of substitution elasticities, such as the CES/Lloyd–Moulton price index?" IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 457).
[89] "The Use of Long-Term and Short-Term Links The long-term/short-term link approach has been used in Sweden for many years and has now also been adopted by the United States. This method facilitates calculating the long-term links of the CPI by the use of superlative index number formulas. The use of long-term and short-term links was mentioned in the paper Estimating the Benefits and Costs of New and Disappearing Products (Diewert and Feenstra 2017) as the best way to get around the problem of dated weighting information. A growing number of countries will begin to explore this approach. Research, discussion, and practical experiences can be shared to identify the advantages and disadvantages of this method." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 457). The work by Diewert and Feenstra (2017) (2022) is mostly about applying an econometric approach to the estimation of preferences

- Retrospective Calculations of Superlative Price Indexes;[90]
- Consumer price indexes for different groups and geographic areas;
- Measuring "hard to measure" services;[91]
- Insurance and financial services;[92]
- Owner Occupied Housing;
- Digitalization;[93]
- Well being and sustainability.

It can be seen that many of the topics listed above were already flagged in the Turvey CPI Manual as areas that required further research.

The 2020 Manual had the following interesting comments on the last topic listed above:

"Measures of well-being attract much interest from policymakers, media, and the public. Ongoing research has focused on how time allocation information is needed to better measure household welfare change (Stiglitz and others, Report on the Measurement of Economic Performance and Social Progress, 2009). As described previously, the Reinsdorf and Schreyer working paper not only focuses on the welfare effects of the digital economy but also outlines key problems in using the CPI in measuring economic well being in general. The paper identifies three reasons why the CPI will overestimate the cost of living and hence underestimate progress in real welfare: (1) insufficient adjustment for quality changes; (2) delayed inclusion of truly new products; and (3) disregarding the appearance and

---

(and to the calculation of the resulting price and quantity indexes) in situations where there are new and disappearing products. A simplified version of this econometric methodology is pursued in the Appendix to this paper.

[90] "Retrospective calculations of superlative price indices are very useful for analytical purposes and to serve as a benchmark to assess the quality of the CPI and quantify potential bias. A limited number of NSOs have begun to compile superlative price indices. A sharing of experience can be used to develop best practices". IMF/ILO/UNECE /Eurostat /OECD/The World Bank (2020; 457).

[91] "The treatment of telecommunication services in the CPI continues to create issues for compilers. All these issues raise user concerns that the CPI becomes less representative and reliable. Where relevant, work on services should be coordinated with the Voorburg Group on Service Statistics." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 457). For recent work on telecom services, see Abdirahman, Coyle, Heys and W. Stewart (2020) (2022).

[92] "Insurance and financial services continue to pose measurement issues for the CPI. With the update of the Classification of Individual Consumption According to Purpose (COICOP) to the 2018 version (COICOP 2018), a separate division (Division 12) was created for insurance and financial services. The net versus gross approaches have not been fully reconciled. There are also problems of choosing appropriate deflators for the payment of premiums. More discussion and research are needed to guide compilers on the appropriate measurement of these services." IMF /ILO/UNECE/Eurostat /OECD/The World Bank (2020; 457). For theoretical approaches to measuring property insurance services and gambling services, see Diewert (1993b; 415-427) (1995b). For theoretical approaches to measuring financial services, see Diewert (2014) and Diewert, Fixler and Zieschang (2016).

[93] "The Organisation for Economic Co-operation and Development Statistics Working Paper, "Measuring Consumer Inflation in a Digital Economy" (Reinsdorf and Schreyer 2020), discusses in detail the problems in measuring the welfare effects of the digital economy, including the effects of services provided for free (or at least without any direct payment) on the internet. According to this working paper, there is a significant need to identify how CPIs can better reflect and incorporate digital goods and services, and work is needed to clarify the conceptual issues and develop methods that better measure the digital economy in the CPI. These issues include, but are not limited to, defining and identifying the goods and services, including different types of internet purchases, services for free, and shared economy services." IMF /ILO/UNECE/Eurostat /OECD/The World Bank (2020; 458). For related research on this topic see Brynjolfsson, Collis, Diewert, Eggers and Fox (2019) and Diewert, Fox and Schreyer (2022).

use of free products. Solving these issues involves addressing both conceptual and practical measurement issues. In a cost of living index context, the theoretically correct way of including truly new products and products offered for free would be the use of estimated "reservation" or "shadow" prices. This can be done in theory and in research studies. However, for the regular production of the monthly CPI, this is usually not feasible and other approaches must be implemented." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 458).

In the Appendix to this paper, we suggest a simplified version of the Diewert-Feenstra estimation of reservation prices methodology that could be used by National Statistical Offices in real time.

The 2020 CPI Manual concluded with the following observations on the problems of measuring welfare when there are free goods and services:

"The issue of a CPI for measuring economic wellbeing is not restricted to the effects of digitalization, but also includes further discussion on the coverage of the CPI and the treatment of different types of goods and services provided for free, potentially including public goods and services such as education, health, safety, or parks. The issue relates to the discussion of cost of goods indices versus cost of living indices, and conditional versus unconditional cost of living indices, where additional experiences and guidance would be useful. It may be useful to invite experts from other areas of official statistics to discuss measuring welfare and economic well-being." IMF/ILO/UNECE/Eurostat /OECD/The World Bank (2020; 458).

Goods and services that are provided to the public free of charge by governments need to be valued at prices that indicate the value of these services to households receiving these goods and services. For possible solutions to these complex valuation problems, see Atkinson (2005), Schreyer (2010) (2012) and Diewert (2018) and the references in these papers.

Diewert (2001) had the opportunity to speculate on "the shape of things to come at statistical agencies". He speculated that:

- Firms will submit price and quantity data on their sales of consumer products to the NSO via the internet. This has come to pass.
- Diewert was less sanguine about NSOs obtaining price and quantity data directly from households.[94] However, some market research firms have persuaded a sample of households to provide their data on household purchases to these firms, who in turn have made their data available to researchers and NSOs (at a price). A potential problem with these credit and debit card data on household purchases is that there may be *no specific product code* to go together with the purchases; the product categories may be highly aggregated. However, some current household specific information on prices and quantities is better than no information.
- Diewert (2001; 108) noted that web scraping of prices would lead to more accurate CPIs. In particular, information on used durable goods (like cars) would lead to better information on consumer durable depreciation rates and hence facilitate the user cost approach to the treatment of durable purchases.

[94] "It is possible that a similar revolution in household data collection could also occur as money for household transactions is replaced by credit and debit card transactions. This switch to plastic away from money opens up the possibility of collecting detailed price and quantity data on individual households' electronically rather than by the old diary and recall methods. If this comes to pass, again as in the producer case, there will be a vast improvement in the quality of household data. I am less certain that this rosy prospect actually will materialize because of concerns about privacy" Diewert (2001; 107).

- Diewert thought that it would be likely that statistical agencies would produce families of indexes. In particular, he thought that NSOs would produce at least 2 CPIs: (i) a Nonrevisable CPI and (ii) A Revisable CPI.[95] He also thought that different CPIs would be produced to suit different purposes.[96] For the most part, this prediction has not come about. However, the papers by Becker (2024) and Abe and Inakura (2024) at this Ottawa Group Meeting indicate that interest in alternative CPIs is growing.
- Diewert (2001; 108-109) also thought that the existence of strongly seasonal products would lead NSOs to produce at least two indexes for seasonal product strata: one index that focussed on year over year price change in the same month (this index would use seasonal baskets) and another index that would attempt to measure month to month price change. This prediction has also not materialized.
- Diewert (2001; 111) thought that many of the unresolved problems associated with the use of hedonic regression models would be solved. This seems to be the case; see Triplett (2004) and section 2 of Diewert and Shimizu (2024).

Finally, Diewert speculated that once information on the household allocation of time became available, many theoretical innovations in modeling the behavior of households would be incorporated into the design of consumer price indexes. Three possible innovations were singled out.

(i) *Implementation of Becker's (1965) theory of the allocation of time*.

"In Becker's model of consumer behavior, households combine their time with market goods and services to produce finally demanded "commodities" that yield direct utility. For example, a consumer combines the services of a bed with time to produce "sleep utility". The theory also takes into account the disutility of time spent working on the external market and the disutility of the time spent commuting to work. The advantage in implementing this theory is that it will give a more complete picture of household activities: the time costs spent on each consumption activity will be valued at some opportunity cost of time and added to cost of purchasing goods and services from the marketplace." Diewert (2001; 109-110).

Some progress has been made in implementing and extending Becker's framework; see Schreyer and Diewert (2014) and Diewert, Fox and Schreyer (2017) and the references in these papers. Note that incorporating the time constraint into standard consumer theory will enable us to better value the contribution of free goods and services to improving the standard of living of households.

(ii) *Implementation of an extended version of Becker's model to cover household nonmarket production*.

[95] "The present single CPI (Consumer Price Index) may be replaced by two indexes: one that is timely (like the present index) and another that would be produced with a lag so that current quantity information could be used (to reduce substitution bias) and quality adjustments could be incorporated. This is a recommendation of the Boskin Commission." Diewert (2001; 108).

[96] "Multiple CPIs may be calculated that reflect different index number purposes or methodologies. For example, some users may feel that a rental equivalence or user cost approach to major consumer durable goods like housing is more appropriate than a money purchases or acquisitions cost approach. Thus statistical agencies may provide alternative indexes that reflect the two approaches. Similarly, some users may want a CPI that has incorporated hedonic quality adjustments or adjustments for increases in the size of consumer choice sets. On the other hand, other users may regard such adjustments as lacking in objectivity and reproducibility and demand a CPI without such adjustments. Finally, some users may want a price index for the domestic purchases of consumer goods and services while other users may want a price index that reflects domestic sales." Diewert (2001; 108). Note that this recommendation was already made by Turvey (1989).

"Becker's model of consumer behavior was concerned with how the household combines purchases of consumer goods and services with its time to "produce" final "commodities" that are demanded by that household to satisfy wants. However, in recent years, as self employment and contracting out of services have grown, many households are producing goods and services at home that are sold to other users. This home production for market sale is not taken into account in Becker's model and so it needs to be extended. The implications of this extension for the cost of living index are profound. Instead of just collecting information on typical consumer goods and services like food, clothing, housing etc., the extended COL to cover household market production would have to include production type inputs like materials (if a product was being made at home) or office equipment (if a business service was being provided) and various traditional consumer purchases (like heating fuel, telephone services, transportation, home computers etc.) would have to be allocated between business and personal use. In addition to putting these business intermediate inputs into the scope of the COL, it would be necessary to account for the outputs produced as well." Diewert (2001; 110).

For some additional references to the literature on measuring household production, see Abraham and Mackie (2005), Hill (2009) and Schreyer and Diewert (2014).

(iii) *Implementation of an extended version of Becker's model to medical economics*.

"Illness, disease and accidents reduce capabilities. For example, if I break my leg, playing tennis or jogging is not feasible. If my vision deteriorates (and I do not get new glasses or have an eye operation), then reading a book or watching television may not be feasible. In the context of Becker's theory of the allocation of time, accident or disease adds extra constraints to the consumer's utility maximization problem and of course, this addition of constraints will reduce welfare. Conversely, certain medical treatments will treat the disease or illness and will remove or lessen these extra constraints, thus adding to consumer welfare. This extension of Becker's theory opens up the possibility of providing welfare based evaluations of the effects of certain medical treatments." Diewert (2001; 110).

It can be seen that obtaining information on the household allocation of time is the key to improving the measurement of the cost of living and household welfare. However, progress in implementing household time surveys has been slow and it will probably continue to be slow. The basic problem is that it is not conceptually simple to measure the household allocation of time across various activities. It is possible to engage in child care while also to listen to music or to watch television. And how exactly are households able to record accurately the time spent on alternative activities? Measuring a household's (digital) monetary transactions is conceptually much simpler.

Two additional interesting problematic measurement areas that were addressed by members of the Ottawa Group are the following topics:

- Measuring household welfare and the construction of consumer price indexes under pandemic conditions and
- Measuring the effects of environmental change.

On the first topic, see Diewert and Fox (2022a) (2022b) and for an introduction to the second topic, see Brandt, Schreyer and Zipperer (2014).

Finally, I would like to draw attention to three papers that were presented at this year's Ottawa Group Meeting. These papers address some of the problems associated with constructing multiple CPIs for different uses and with the treatment of "hard to measure goods and services".

The first paper was by Corinne Becker (2024) and she showed how different CPIs could be constructed for different socio-economic groups; i.e., she constructed alternative CPIs for households stratified by age, family composition and income. She also pointed out that a "traditional" CPI covered only about 60% of total household expenditures: taxes and various social insurance charges made up the remaining 40% of household expenditures.

The second paper was by Naohito Abe and Noriko Inakura (2024) and it is similar to Becker's paper in focussing on the construction of CPIs for separate household groups. Abe and Inakura also focus on the difference between total household consumption and the portion of consumption that is covered by a traditional CPI. Traditional CPIs cover what is called Household Final Consumption in the international System of National Accounts. Actual Consumption adds to Household Final Consumption the free and highly subsidized goods and services provided by governments to households, such as health care, education, daycare and subsidized housing.

The third paper was by Pavel Belchev (2024) and it dealt with alternative approaches to the measurement of real gambling expenditures in a CPI.

The measurement of the *real* quantity of gambling and insurance products is a very difficult task. Even the measurement of *nominal* gambling and insurance expenditures is controversial. Many if not most economists and statisticians lean towards taking a net approach to these expenditures which involve risk and uncertainty. Gross gambling expenditures equal the amount wagered over the period under consideration. Net gambling expenditures equal gross expenditures less winnings or prize money. Similarly when measuring property insurance services, one could use *gross premiums* paid over the reference period or *net premiums* equal to gross premiums less the value of insurance claims paid back to policy holders.

It is possible to make a case for the use of *gross premiums* and *gross wagers* as useful *nominal measures* of insurance and gambling services, particularly when dealing with household specific indexes. We will first consider the case for gross wagers.

Suppose that there are two households (or homogeneous household groups) and the value of their gambling expenditures in a given period are V_{G1} and V_{G2}. Suppose Household 1 is made up of losers and Household 2 is made up of winners, with the value of prize winnings by Household 2 equal to V_{W2}. There is a gambling business sector that collects the wagers and distributes the prize money. There are 3 consolidated transactions made by the households with the gambling sector. The Table below shows these transactions.

Table 1: Transactions between Households and the Gambling Sector

	Household 1	Household 2	Gambling Sector
Flow Transaction	V_{G1}		V_{G1}
Flow Transaction		V_{G2}	V_{G2}
Balance Sheet Transaction		V_{W2}	V_{W2}

The interpretation of the above table is that households gain utility by gambling and the nominal value of gambling (equal to the amounts wagered during the period) for Household j is V_{Gj} for j = 1,2. At the end of the period, Household 2 gets a transfer equal to the winnings V_{W2} and this amount is added to Household 2's Balance Sheet. *Total Household Comprehensive Income* is defined as household expenditures on gambling, $V_{G1} + V_{G2}$, plus net increase in household wealth, V_{W2}. If the gambling sector distributes winnings in a costless manner, $V_{G1} + V_{G2}$ will equal winnings V_{W2} and thus Comprehensive Income[97] is equal to two times gambling expenditures; i.e., we have what looks like a potential double counting problem. The alternative accounting approach is to subtract line 4 in the above Table from line 3 and we obtain the *net approach* to the treatment of gambling expenditures. The problem with this approach is that while the value of gambling expenditures to Household 1 would be equal to $V_{G1} > 0$, the net value of gambling expenditures to Household 2 would be equal to $V_{G2} - V_{W2} < 0$. This implies that the marginal utility to Household 2 of spending money on gambling is negative, which does not seem to be a sensible conclusion. Thus the net approach to the treatment of gambling expenditures breaks down when we disaggregate the household sector into winners and losers. Whether it is reasonable to adopt the gross approach to gambling expenditures requires additional research and discussion.

We can take a similar approach to the treatment of property insurance expenditures by following the monetary transactions associate with insurance expenditures. However, it is now necessary to introduce a *Nature* or *Environment Sector* into the economy. For our present purpose, this sector creates *property damage* during the reference period with a nominal value equal to V_D. We disaggregate the insured household population during the reference period into Group 1 households who make no damage claims and into Group 2 households who suffer property damage equal to V_D. Group j Households pay gross premiums equal to V_{Gj} for j = 1,2. Let V_C be the value of claims made by Group 2 households. There are 4 transactions in this simplified insurance model. The counterpart to Table 1 is now Table 2:

Table 2: Transactions between the Household, Insurance and Nature Sectors

	Household 1	Household 2	Insurance Sector	Nature Sector
Flow Transaction	V_{G1}		V_{G1}	
Flow Transaction		V_{G2}	V_{G2}	
Balance Sheet Transaction		V_C	V_C	
Balance Sheet Transaction		$-V_D$		$-V_D$

If the Insurance Sector were not present, then all of the entries in the above Table would vanish except the transaction in the last row; i.e., Nature would cause the value of Household 2's assets to drop by V_D and Comprehensive Net Income of the economy would drop by V_D. However, when an Insurance Sector is added to the economy, gross premiums allow this sector to create an asset, V_C, which can fully or partially offset the asset loss V_D. Each household pays a gross premium and each household gets a positive increment of utility from this payment, whether damage occurs or not. The Net (Insurance) Comprehensive Income of Household 2 is equal to gross premium expenditures V_{G2} plus net increase in wealth which is equal to the value of insurance claims V_C less the value of environmental damage V_D. If $V_C = V_D$, then total Net Comprehensive Income is equal to the value of Gross Premium payments, $V_{G1} + V_{G2}$. Contrast this *gross premiums* approach to the corresponding *net premiums* approach to property insurance services which subtracts V_C from V_{G2} and thus Household 2's *net expenditure* on insurance becomes $V_{G2} - V_C < 0$, which is awkward, because it implies that the marginal utility of spending money on insurance is negative. Thus the Gross Premiums approach to the measurement of insurance services seems to be more reasonable. We note that if this approach were accepted, then nominal GDP growth in most countries would be increased by a considerable amount since investment in housing structures has been very large during the past two

[97] See Diewert and Fox (2024) on the concept and history of Comprehensive Net Income.

decades in most countries and thus gross property premiums have grown considerably more than net premiums.[98]

My overall conclusion is on the Ottawa Group Meetings is that these meetings have been instrumental in improving consumer price indexes over the years.

Appendix: The Consumer Demand Approach to the Construction of Multiperiod Indexes

A1. The Estimation of Consumer Demand Functions using the Unit Cost Function

We consider the case where all purchasers (or consumers) of a specified group of N related products have the same preferences over a window of T periods. Let the period t aggregate quantity vector be $q^t \equiv [q_{t1},\ldots,q_{tN}]$ for t = 1,…,T. If product n in period t is not available, then $q_{tn} \equiv 0$. The period t price vector which corresponds to q^t is $p^t \equiv [p_{t1},\ldots,p_{tN}]$ for t = 1,…,T. If product n is available in period t, then p_{tn} is the observed (unit value) price for product n. If product n is not available in period t, then p_{tn} should be the (unobserved) Hicksian reservation price p_{tn}^* that is just high enough to deter purchasers from buying product n in period t.[99] For now, if product n is not available in period t, we set p_{tn} equal to 0.

Let the consumer's utility function be f(q) where $q \equiv [q_1,\ldots,q_N]$. We assume f(q) is once a differentiable, linearly homogeneous, nondecreasing and concave function of the components of q. Then it can be shown that consumer preferences can be represented by the dual unit cost function, c(p), defined as follows for positive price vectors $p \equiv [p_1,\ldots,p_N]$:

$$\text{(A1)}\quad c(p) \equiv \min_q \{p\cdot q : f(q) = 1;\ q \geq 0_N\}$$

where 0_N is a vector of zeros of dimension N. Suppose for the moment, that there are no missing products in period t. Then if all purchasers are maximizing the utility function f(q) subject to their budget constraints and all purchasers are facing the same observed period t price vector p^t, then the observed period t aggregate quantity vector q^t will satisfy the following system of equations:[100]

$$\text{(A2)}\quad q^t = \nabla c(p^t)u^t\ ; \qquad t = 1,\ldots,T$$

where observed period t expenditure $e^t \equiv p^t\cdot q^t \equiv \Sigma_{n=1}^N p_{tn}q_{tn}$, unobserved utility is u^t and $\nabla c(p^t) \equiv [\partial c(p^t)/\partial p_1,\ldots, \partial c(p^t)/\partial p_N]$ is the vector of first order partial derivatives of c(p) evaluated at $p = p^t$. We also have the following equations:

$$\text{(A3)}\quad e^t = c(p^t)u^t\ ; \qquad t = 1,\ldots,T.$$

[98] This still leaves open the difficult issues surrounding the measurement of real gambling and insurance services. In particular, the real measures should be sensitive to changes in risk. For an introduction to the theoretical literature on this topic, see Diewert (1993b) (1995b).
[99] Hicks (1940; 114) introduced the concept of a reservation price into the economics literature.
[100] See Diewert (1974; 110-113). This follows from Shephard's Lemma.

Now divide q^t by e^t and $\nabla c(p^t)u^t$ by $c(p^t)u^t$ and after a bit of rearrangement, we obtain the following system of *consumer demand functions*:

$$(A4)\ q^t = e^t\nabla c(p^t)/c(p^t)\ ; \qquad t = 1,\ldots,T.$$

At this point, a functional form for c(p) can be postulated, the partial derivatives of $c(p^t)$ can be calculated, error terms to equations (A4) can be added and the resulting system of estimating equations can be used to provide estimates for the unknown parameters which define the postulated unit cost function c(p).[101] Once c(p) is known, the period t *multilateral or multiperiod price and quantity levels*, P^t and Q^t, are defined as follows:

$$(A5)\ P^t \equiv c(p^t);\ Q^t \equiv e^t/c(p^t)\ ; \qquad t = 1,\ldots,T.$$

A problem with the above price levels P^t defined by (A5) is that the corresponding price indexes, P^t/P^1, are not invariant to changes in the units of measurement and hence they are not suitable for official use. This problem can be addressed by converting the system of estimating equations defined by (A4) into share equations. Thus define the *expenditure share of product n in period t*, s_{tn}, as follows:

$$(A6)\ s_{tn} \equiv p_{tn}q_{tn}/p^t\cdot q^t = p_{tn}q_{tn}/e^t\ ; \qquad n = 1,\ldots,N\ ;\ t = 1,\ldots,T.$$

Multiply both sides of equation n in (A4) by p_{tn} and divide by e^t for n = 1,…,N and t = 1,…,T. Equations (A4) become equations (A7):

$$(A7)\ s_{tn} = p_{tn}[\partial c(p^t)/\partial p_n]/c(p^t)\ ; \qquad n = 1,\ldots,N\ ;\ t = 1,\ldots,T.$$

Errors can be added to the right hand sides of equations (A6) and the unknown parameters in the chosen functional form for c(p) can be estimated.[102] The functional form should be chosen so that the estimated unit costs $c(p^t)$ for t = 1,…,T are uniquely determined up to a scale factor and so that the estimated P^t/P^1 are invariant to changes in the units of measurement for the individual products.

Rather than specifying an explicit error structure for equations (A7), we will assume that the unknown parameters which characterize the chosen unit cost function c(p) are estimated by solving the least squares minimization problem (A8) below with respect to the choice of these parameters:[103]

$$(A8)\ \min_{\text{parameters of } c(p)}\ \{\Sigma_{t=1}^{T}\ \Sigma_{n=1}^{N}\ [s_{tn} - p_{tn}[\partial c(p^t)/\partial p_n]/c(p^t)]^2\}.$$

The above estimation strategy is fine if there are no missing products. But as soon as there are missing products, it is necessary to replace the missing prices by reservation prices which become unknown parameters which have to be estimated. Thus suppose product n in period t is missing and the unknown reservation price is $p_{tn}^* > 0$. Then the corresponding expenditure share $s_{tn} = 0$. We could set $p_{tn} = 0$ and then the term $s_{tn} - p_{tn}[\partial c(p^t)/\partial p_n]$is equal to 0 as well and we can simply omit this term in the sum of terms

[101] It will turn out that a normalization on one of the unknown parameters in c(p) will have to be made in order to identify all of the parameters in c(p).

[102] Again, it is necessary to impose a normalization on the parameters in c(p) in order to identify all of the unknown parameters.

[103] As usual, we will require a normalization on these parameters in order to identify all of the unknown parameters in c(p).

defined by (A8). However, the unknown reservation price p_{tn}^* still appears in the remaining terms that involve partial derivatives of the unit cost function. Thus econometric estimation of the unknown parameters in the case of missing observations becomes extremely difficult if not impossible.[104]

Since the cost function approach to the estimation of reservation prices fails when there are missing products, we turn to the direct estimation of consumer utility functions, which involves the estimation of systems of inverse demand functions. The usual system of consumer demand functions give quantities as functions of expenditures and prices; the system of inverse demand functions give prices as functions of quantities consumed and expenditure.

A2. The Estimation of Inverse Demand Functions

Assume for the moment that there are no missing products in each period. The inverse demand function estimation methodology starts with the assumption that the observed period t quantity vector q^t is a solution to the following period t utility maximization problem:

$$\text{(A9)} \max_q \{f(q) : p^t \cdot q = e^t ; q \geq 0_N\} ; \qquad t = 1,\ldots,T.$$

The first order conditions for the observed q^t to solve the period t utility maximization problem are the following conditions:

$$\text{(A10)} \nabla f(q^t) = \lambda_t p^t ; \qquad t = 1,\ldots,T;$$
$$\text{(A11)} p^t \cdot q^t = e^t ; \qquad t = 1,\ldots,T.$$

Take the inner product of both sides of (A10) with q^t and solve the resulting equation for the Lagrange multiplier λ_t. We find that

$$\text{(A12)} \lambda_t = q^t \cdot \nabla f(q^t)/e^t \qquad t = 1,\ldots,T$$
$$= f(q^t)/e^t$$

where the second line in (A12) follows from Euler's Theorem on homogeneous functions which (using our assumption that f(q) is linearly homogeneous in q) implies that $f(q^t) = q^t \cdot \nabla f(q^t) = \Sigma_{n=1}^N q_{tn} \partial f(q^t)/\partial q_n$ for t = 1,…,T. Substitute λ_t defined by (A12) into equations (A10) and after a bit of rearrangement, we obtain the following system of estimating equations:

$$\text{(A13)} p^t = e^t \nabla f(q^t)/f(q^t) ; \qquad t = 1,\ldots,T.$$

Equations (A13) define a *system of inverse demand functions*. We could assume a suitable functional form for the utility function f(q), add error terms of the right hand sides of these equations and use the resulting system of equations as estimating equations to determine the unknown parameters that characterize the function f(q).[105] Once the parameters have been estimated, we can calculate period t aggregate quantities Q^t and the corresponding price levels P^t as follows:

[104] Hausman (1996) (1999) used variants of this cost function methodology to estimate reservation prices but it is not known how he solved this problem. Estimation is possible if the unit cost function is linear function of prices but in this case, finite reservation prices do not exist.
[105] We also require a normalization on the parameters that define f(q) in order to obtain a unique function.

(A14) $Q^t \equiv f(q^t)$; $P^t \equiv e^t/f(q^t)$; $\quad t = 1,\ldots,T.$

Again, there is a problem with the above strategy: we would like the estimated relative price levels P^t/P^1 to be invariant to changes in the units of measurement for the N products in scope. This will not be the case if we estimate the system of inverse demand functions using equations (A13). As in section A1 above, we solve this problem by estimating the system of share equations that corresponds to (A13). Thus multiply both sides of equation n for period t in (A13) by q_{tn}/e^t and we obtain the following system of *inverse demand function share equations*:

(A15) $s_{tn} = q_{tn}[\partial f(q^t)/\partial q_n]/f(q^t)$; $\quad n = 1,\ldots,N ; t = 1,\ldots,T.$

Equations (A15) are valid even if there are missing prices and quantities. If product n in period t is not available, then $q_{tn} = s_{tn} = 0$ and equation t, n in (A15) becomes 0 = 0. This is an important point and explains why estimating a system of inverse demand functions is possible when there are missing products and why it is not possible to estimate a system of regular demand functions when there are missing products: when product n in period t is missing, we know that $q_{tn} = 0$ whereas we do not know what the companion reservation price p_{tn} is.

Rather than specifying an explicit error structure for equations (A15) after adding error terms,[106] we will simply assume that the unknown parameters which characterize the chosen utility function f(q) are estimated by solving the least squares minimization problem (A16) below with respect to the choice of these parameters:[107]

(A16) $\min_{\text{parameters of } f(q)} \{\Sigma_{t=1}^T \Sigma_{n=1}^N [s_{tn} - q_{tn}[\partial f(q^t)/\partial q_n]/f(q^t)]^2\}.$

Once the unknown parameters characterizing f(q) have been estimated, we can calculate period t aggregate quantities Q^t and the corresponding price levels P^t using definitions (A14).

If product n in period t is missing, then $q_{tn} = 0$ and $s_{tn} = 0$ as well. Thus the term involving s_{tn} in the minimization problem (A16) simply drops out of the summation. But that is all that happens: the remaining terms in the summation are not affected when q_{tn} equals zero. The unknown reservation prices p_{tn}^* which appeared in the minimization problem defined by (A8) when there were missing prices do not appear anywhere in the minimization problem defined by (A16). Note also that the unknown reservation prices p_{tn}^* for missing products do not appear in definitions (A14) either. Recall that in the main text, we defined S(t) as the set of products that were purchased in period t. When there are missing products, it can be seen that the double summation that appears in (A16) can be replaced by the following summation of terms that excludes missing products:

(A17) $\Sigma_{t=1}^T \Sigma_{n=1}^N \{s_{tn} - q_{tn}[\partial f(q^t)/\partial q_n]/f(q^t)\}^2 = \Sigma_{t=1}^T \Sigma_{n\in S(t)} \{s_{tn} - q_{tn}[\partial f(q^t)/\partial q_n]/f(q^t)\}^2$.

[106] It is usual in estimating systems of consumer demand equations to assume no missing prices and also to assume that the errors in the N equations pertaining to a single period are correlated so that a variance covariance matrix with N(N+1)/2 unknown parameters also needs to be estimated. In our present context where there will typically be hundreds of products, this strategy becomes unworkable. The existence of missing prices also creates difficulties for the standard consumer demand systems methodology.

[107] As usual, we will require a normalization on these parameters in order to identify all of the unknown parameters in f(q).

The equality in (A17) is useful in setting up the independent and dependent variables that correspond to the nonlinear least squares regression model when we have missing products. It is more convenient to set up the single equation nonlinear regression problem that corresponds to the least squares minimization problem (A16) than to set up the corresponding problem that excludes the shares that correspond to missing products.

Once the solution to the least squares minimization problem has been obtained, then if product n in period t is missing, the reservation price for this product p_{tn}^* can be estimated as follows:

$$(A18)\ p_{tn}^* \equiv e^t[\partial f(q^t)/\partial q_n]/f(q^t).$$

The bottom line is this: it is virtually impossible to estimate systems of direct consumer demand functions when there are missing prices but it is reasonably straightforward to estimate systems of inverse demand functions. In the following sections, we will work through the algebra presented in this section for specific functional forms for f(q).

A3. The Econometric Estimation of Linear Preferences

Our first example of the methodology explained in the previous section is the case where the utility function is a homogeneous linear function of the quantities consumed. Thus we assume that $f(q,\alpha)$ has the following functional form:

$$(A19)\ f(q,\alpha) \equiv \Sigma_{n=1}^N \alpha_n q_n = \alpha\cdot q.$$

The least squares minimization problem (A16) becomes the following problem:

$$(A20)\ \min_{\alpha\text{'s}}\ \{\Sigma_{t=1}^T \Sigma_{n=1}^N [s_{tn} - \alpha_n q_{tn}/\alpha\cdot q^t]^2\}.$$

If $\alpha^* \equiv [\alpha_1^*,\ldots, \alpha_N^*]$ is a solution to (A20), then it can be seen that $\lambda\alpha^*$ is also a solution to (A20) where λ is any positive number. This non-uniqueness always occur when we attempt to estimate utility functions. The scale of utility is arbitrary so we need to impose at least one normalization on the estimated parameters in order to obtain a cardinal measure of utility. There is another problem with the minimization problem defined by (A20): it can be the case that there is no solution to (A20). For example, suppose that there are only 2 periods and 2 products in scope. Suppose further that product 1 is only available in period 1 and product 2 is only available in period 2. In this case, there are only 2 independent estimating equations in the nonlinear minimization problem defined by (A20):

$$(A21)\ s_{11} = \alpha_1 q_{11}/(\alpha_1 q_{11} + \alpha_2 q_{12}) = \alpha_1 q_{11}/(\alpha_1 q_{11} + \alpha_2 0) = 1;$$
$$(A22)\ s_{22} = \alpha_2 q_{22}/(\alpha_1 q_{21} + \alpha_2 q_{22}) = \alpha_2 q_{22}/(\alpha_1 0 + \alpha_2 q_{22}) = 1.$$

It can be seen that it is not possible to obtain estimates for the quality adjustment parameters α_1 and α_2 in this situation. We need some product overlap between the periods in order to obtain solutions to (A20).

In order to solve the problems of non-uniqueness and non-existence, we assume that there is a product that is present in all T periods and we assume that each product in scope is purchased in at least one period. We reorder products if necessary and let product 1 be the always present product. Thus we have:

(A23) $q_{t1} > 0$; $t = 1,\ldots,T$.

We make the following (preliminary) normalization on the α_n:

(A24) $\alpha_1 \equiv 1$.

Using the above assumptions, the solution $\alpha^* \equiv [\alpha_1^*,\ldots, \alpha_N^*]$ to (A20) is unique with $\alpha_1^* = 1$. With this solution in hand, we can define the period t aggregate quantity levels Q^t and the corresponding period t price levels P^t as follows:

(A25) $Q^t \equiv \alpha^*\cdot q^t$; $P^t \equiv e^t/\alpha^*\cdot q^t$; $t = 1,\ldots,T$.

The above price and quantity levels are unique where we have used the normalization $\alpha_1^* \equiv 1$. Recall that if α^* is a solution to (A20), then $\lambda\alpha^*$ is also a solution to (A20). Thus we can scale the vector α^* by a positive number λ^* so that $\lambda^*\alpha^*$ is also a solution to (A20) but this scaling makes the resulting P^{1*} equal to 1. The scaled $P^{t*} \equiv e^t/\lambda^*\alpha^*\cdot q^t$ for $t = 1,\ldots,T$ are the Least Squares Linear Preferences (LSLP) multilateral price indexes that are the result of this model.

Are these LSLP price indexes that are based on the estimation of a linear utility function acceptable to National Statistical Offices? It seems that they should be acceptable since NSOs are already estimating linear preferences when they use Geary (1958) Khamis (1970) multilateral indexes to construct portions of their CPIs. Many offices also use hedonic time dummy regression models to quality adjust products that quickly appear and then disappear. Hedonic regression models with time dummies are also based on the (implicit) assumption of linear preferences.[108]

In the following section, we will look at the test properties of the LSLP system of price and quantity levels, P^t and Q^t, defined by (A25).

A4. The Test Properties of the LSLP Multilateral Price and Quantity Levels

We make the assumptions on the underlying price and quantity data that were made in the previous section. If we ignore the normalization (A24) for the moment (it was used to get a unique solution to the least squares minimization problem defined by (A20)), then the first order conditions for solving (A20) simplify to the following N equations which determine a solution vector α^*:

(A26) $\Sigma_{t=1}^{T} [s_{tn} - (\alpha_n q_{tn}/\alpha\cdot q^t)][(q_{tn}/\alpha\cdot q^t)][1 - (\alpha_n q_{tn}/\alpha\cdot q^t)] = 0$; $n = 1,\ldots,N$.

The terms $q_{tn}/\alpha\cdot q^t$ and $[1 - (\alpha_n q_{tn}/\alpha\cdot q^t)]$ are nonnegative and act as weights for the term $s_{tn} - (\alpha_n q_{tn}/\alpha\cdot q^t)$. If $q_{tn} = 0$, then the product of these three terms is equal to 0.

The period t price and quantity levels defined by the α^* solution to (A20), $P^t \equiv e^t/\alpha^*\cdot q^t$ and $Q^t \equiv \alpha^*\cdot q^t$ defined by (A25) will be denoted by $P^t(p^1,\ldots,p^T; q^1,\ldots,q^T)$ and $Q^t(p^1,\ldots,p^T; q^1,\ldots,q^T)$ at times to indicate that these levels depend on the underlying price and quantity data.

Test 1: *Quantity Levels Identity Test*: If $q^r = q^t$, then $Q^r = Q^t$.

[108] See Diewert (2022) and Diewert and Shimizu (2024) on this point.

Proof: Follows from $Q^r \equiv \alpha^*\cdot q^r$, $Q^t \equiv \alpha^*\cdot q^t$ and $q^r = q^t$.

Note that we cannot establish the counterpart test for prices; i.e., if $p^r = p^t$, then it is not necessarily true that $P^r = P^t$. If $p^r = p^t$ and $q^r = q^t$, then it is true that $P^r = P^t$.

Test 2: *Basket Test for Price Levels*: If $q^r = q^t = q$, then $P^t/P^r = p^t\cdot q/p^r\cdot q$.

Proof: If $q^r = q^t = q$, then we have $P^r = e^r/\alpha^*\cdot q^r = p^r\cdot q^r/\alpha^*\cdot q^r = p^r\cdot q/\alpha^*\cdot q$ and $P^t = e^t/\alpha^*\cdot q^t = p^t\cdot q^t/\alpha^*\cdot q^t = p^t\cdot q/\alpha^*\cdot q$. Thus $P^t/P^r = p^t\cdot q/p^r\cdot q$.

Test 3: *The Product Reversal Test* (Invariance of the price and quantity levels to changes in the ordering of the products).

Proof: The objective function in the minimization problem defined by (A20) is invariant to the ordering of the products and definitions (A25) are also invariant to the ordering of the products.

Test 4: *Linear Homogeneity Test for Quantity Levels*: If $q^t = \lambda q^r$ for some $\lambda > 0$, then $Q^t = \lambda Q^r$.

Proof: Follows from $Q^r \equiv \alpha^*\cdot q^r$, $Q^t \equiv \alpha^*\cdot q^t = \alpha^*\cdot\lambda q^r = \lambda\alpha^*\cdot q^r = \lambda Q^r$. Note that this test implies Test 1.

Test 5: *Within Period Linear Homogeneity Test for Quantity Levels*: $Q^1(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = \lambda Q^1(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, $Q^2(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = Q^2(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, …, $Q^T(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = Q^T(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$ for all $\lambda > 0$.

Proof: The initial least squares minimization problem minimizes $\{\Sigma_{t=1}^T \Sigma_{n=1}^N [s_{tn} - \alpha_n q_{tn}/\alpha\cdot q^t]^2\}$and has solution α^*. Now replace the initial period 1 quantity vector q^1 by λq^1 where $\lambda > 0$. The shares s_{tn} remain unchanged and the terms $\alpha_n q_{tn}/\alpha\cdot q^t$ also remain unchanged. Thus the initial α^* unique solution remains the solution to the new least squares minimization problem where q^1 is replaced by λq^1. Thus all of the $Q^t = \alpha^*\cdot q^t$ and $P^t = e^t/ \alpha^*\cdot q^t$ remain unchanged for $t = 2,3,\ldots,T$. However, the new $Q^1 = \alpha^*\cdot\lambda q^1 = \lambda\alpha^*\cdot q^1$ so that $Q^1(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = \lambda Q^1(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$. The new P^1 is equal to $e^1/\alpha^*\cdot\lambda q^1 = p^{1*}\cdot\lambda q^1/\alpha^*\cdot\lambda q^1 = p^{1*}\cdot q^1/\alpha^*\cdot q^1$ which is equal to the initial P^1. Thus P^1 also remains unchanged.

Corollary: $P^1(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = P^1(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, $P^2(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = P^2(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, …, $P^T(p^1,\ldots,p^T; \lambda q^1,q^2,\ldots,q^T) = P^T(p^1,\ldots,p^T; q^1,q^2\ldots,q^T)$ for all $\lambda > 0$. Thus all of the price levels are homogeneous of degree 0 in the quantity vector q^1.[109]

Note that in the above tests, q^1 was replaced by λq^1. The Product Reversal Test implies that analogous tests hold if q^t is replaced by λq^t.

Test 6: *Within Period Linear Homogeneity Test for Price Levels*: $P^1(\lambda p^1,\ldots,p^T; q^1,q^2,\ldots,q^T) = \lambda P^1(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, $P^2(\lambda p^1,\ldots,p^T; q^1,q^2,\ldots,q^T) = P^2(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$, …, $P^T(p^1,\ldots, \lambda p^T; q^1,q^2,\ldots,q^T) = P^T(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$ for all $\lambda > 0$.

[109] Note that $\lambda^0 = 1$.

Proof: The initial least squares minimization problem minimizes $\Sigma_{t=1}^{T}\,\Sigma_{n=1}^{N}\,[s_{tn} - \alpha_n q_{tn}/\alpha\cdot q^t]^2$ and has solution α^*. Now replace the initial period 1 price vector p^1 by λp^1 where $\lambda > 0$. The shares s_{tn} remain unchanged and the terms $\alpha_n q_{tn}/\alpha\cdot q^t$ also remain unchanged. Thus the initial α^* unique solution remains the solution to the new least squares minimization problem where p^1 is replaced by λp^1. Thus all of the $Q^t = \alpha^*\cdot q^t$ remain unchanged and the $P^t = e^t/\,\alpha^*\cdot q^t$ remain unchanged for t = 2,3,…,T. However, the new $P^1 = \lambda p^1\cdot q^1/\alpha^*\cdot q^1$ which is equal to λ times the initial P^1.

Corollary: $Q^1(\lambda p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T) = Q^1(p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T)$, $Q^2(\lambda p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T) = Q^2(p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T)$, …, $Q^T(\lambda p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T) = Q^T(p^1,\ldots,p^T;\, q^1,q^2,\ldots,q^T)$ for all $\lambda > 0$. Thus all of the quantity levels are homogeneous of degree 0 in the price vector p^1.

The Product Reversal Test implies that analogous tests hold if p^t is replaced by λp^t.

Test 7: *Invariance to the Ordering of Periods Test*: If we permute the ordering of the periods, then the resulting price and quantity levels are equal to the same permutation of the initial price and quantity levels.

Proof: This property follows from the additive in time nature of the objective function in the minimization problem defined by (A20).

Test 8: *Invariance of the Price and Quantity Levels to Changes in the Units of Measurement*.

Proof: Let α^* solve (A20). Define the initial price and quantity levels as $Q^t \equiv \alpha^*\cdot q^t$ and $P^t \equiv e^t/\alpha^*\cdot q^t$ for t = 1,…,T. Change the units of measurement so that q_{tn} is replaced by $q_{tn}^* \equiv \gamma_n q_{tn}$ and p_{tn} is replaced by $p_{tn}^* \equiv p_{tn}/\gamma_n$ where the γ_n are positive constants for t = 1,…,T and n = 1,…,N. Note that the expenditure shares s_{tn} and total expenditure e^t remain unchanged when we change the units of measurement. The new nonlinear least squares minimization problem is defined as follows:

(A27) $\min_{\alpha\text{'s}}\ \{\Sigma_{t=1}^{T}\,\Sigma_{n=1}^{N}\,[s_{tn} - \alpha_n q_{tn}^*/\alpha\cdot q^{t*}]^2\}$.

It can be seen that $\alpha_n q_{tn}^*/\alpha\cdot q^{t*} = \alpha_n\gamma_n q_{tn}/\Sigma_{j=1}^{N}\,\alpha_j\gamma_j q_{tj}$ and $\alpha_n^{**} \equiv \alpha_n^*/\gamma_n$ is the solution to (A27). Let $\alpha^{**} \equiv [\alpha_1^{**},\ldots,\alpha_N^{**}]$. Then we have the new Q^{t*} and P^{t*} defined as follows:

(A28) $Q^{t*} \equiv \alpha^{**}\cdot q^{t*} = \Sigma_{n=1}^{N}\,\alpha_n^{**}q_{tn}^* = \Sigma_{n=1}^{N}\,[\alpha_n^*/\gamma_n]\gamma_n q_{tn} = \Sigma_{n=1}^{N}\,\alpha_n^*\,q_{tn} = \alpha^*\cdot q^t = Q^t$; $\qquad$ t = 1,…,T;
(A29) $P^{t*} \equiv e^t/Q^{t*} = e^t/\alpha^*\cdot q^t = P^t$; $\qquad$ t = 1,…,T.

Test 9 : *Circularity or Transitivity Test* : The Price and Quantity Levels defined by definitions (A25) satisfy the following relationships:

(A30) $Q^t/Q^r = (Q^t/Q^s)(Q^s/Q^r)$; $P^t/P^r = (P^t/P^s)(P^s/P^r)$; $\qquad 1 \le r < s < t \le T$.

Proof: Since $Q^t = \alpha^*\cdot q^t$ for t = 1,…,T, the first set of equalities in (A30) follows immediately. The second set of equalities in (A30) follows from $P^t \equiv e^t/Q^t$ for t = 1,…,T and the transitivity of the e^t and Q^t.

Test 9 means that the econometric price and quantity indexes are free from chain drift.

The above Tests are modifications of standard tests for bilateral indexes to the multilateral context. The following Test 10 is not a standard test.

Test 10: *Price Responsiveness Test*: Suppose that product N is a new product that is purchased in period T. Then the price level function $P^T(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$ responds to a change in the price of this new product, p_{TN}; i.e., $P^T(p^1,\ldots,p^T; q^1,q^2,\ldots,q^T)$ is not constant with respect to variations in p_{TN}.[110]

Proof: Experiments with this method show that the LSLP price levels satisfy this test.

Zhang, Johansen and Nygaard (2019) pointed out that "standard" multilateral indexes like GEKS, CCDI, GK or the TPD multilateral indexes do not satisfy this test: the appearance of a new product does not affect the price level for the current period for these multilateral methods. Thus the LSLP Multilateral Price indexes have a possible advantage over these alternative multilateral indexes.

Even though the LSLP indexes satisfy the Responsiveness Test, they are not entirely satisfactory because *they underestimate the benefits of an increased choice set to consumers*. This problem can be explained by one of the favourite tools used by economists: the use of indifference curves to show how consumers choose between competing goods and services to maximize their utility while satisfying a budget constraint. Consider Figure 1 below.

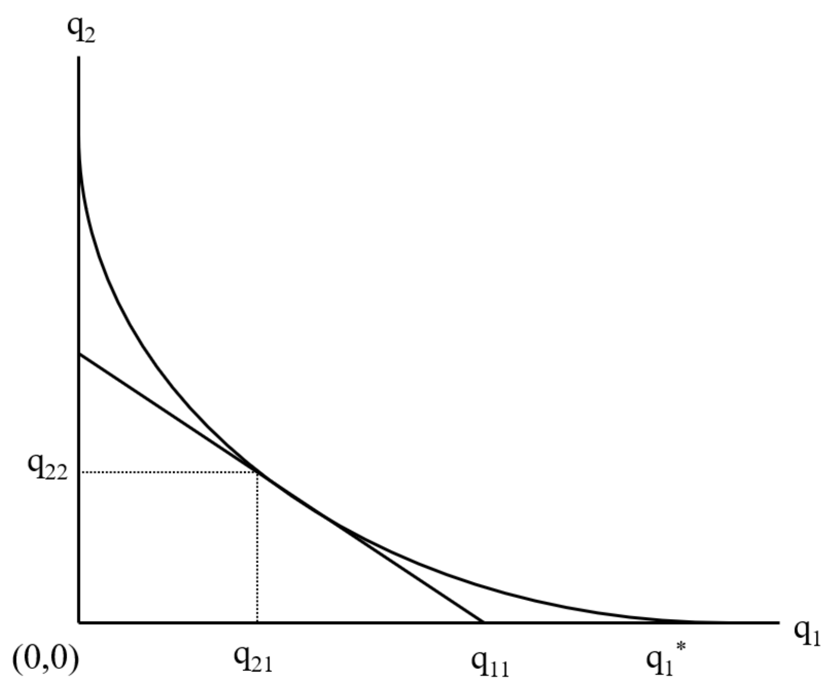

Figure 1. The Introduction of a New Product.

[110] This test is due to Zhang, Johansen and Nygaard (2019). This paper has other interesting tests for multilateral indexes.

Consider a consumer's utility maximization problem over two periods when there are only two products to choose in an elementary category. In period 1, only product 1 is available. The consumer has a budget equal to $e^1 > 0$ in period 1 to spend on products in this category but only product 1 is available. The price of product 1 in period 1 is $p_{11} > 0$. The consumer purchases $q_{11} = e^1/p_{11}$ units of product 1 in period 1. In period 2, a new product 2 appears in this category. The consumer has e^2 to spend on the two products in period 2 and faces the budget constraint, $p_{21}q_1 + p_{22}q_2 = e^2$. For simplicity, we assume that the period 2 budget e^2 is equal to e^1 and the price of product 1 in period 2, p_{21}, is equal to its price in period 1 so that $p_{21} = p_{11}$. Thus the consumer *could* choose to purchase the same amount of product 1 in period 2 as was chosen in period 1; i.e., the consumer could choose $q_1 = q_{11}$ and $q_2 = 0$ in period 2. But in period 2, the consumer is not limited to the period 1 choice: any point on the budget line that starts at q_{11} is a feasible period 2 consumption combination. The curved line that is tangent to the budget line at the point (q_{21},q_{22}) is the highest indifference curve that is also on the consumer's budget line so the point $(q_{21}.q_{22})$ in Figure 1 is the utility maximizing choice for the consumer in period 2. But if we solved the least squares minimization problem (A20) for this example, we would find that $Q^2 = Q^1$; i.e., the measured utility level in period 2, Q^2, would equal the measured utility level in period1, Q^1, and thus the measured utility ratio equals 1. But the true utility ratio is $u^2/u^1 = q_1^*/q_{11} \geq 1$. If preferences are linear, then the observed budget line would also be the highest attainable indifference curve for the consumer and q_1^* would coincide with $q_{11} = q_{21}$ and there would be no bias. But typically, preferences are not linear and so the linear utility function model leads to a quantity index that is too low and the corresponding price index has an upward new product bias in the period when the new product first appears. This is a fundamental problem that arises for all linear utility function models, including GK indexes, bilateral superlative indexes that rely on matched product prices and their multilateral extensions that use matched bilateral superlative indexes as building blocks and hedonic regression models that use time dummy variables.

Note that if we extend the above 2 period model to 3 periods where products 1 and 2 are also available in period 3 (and the relative prices of products 1 and 2 are different in periods 2 and 3 so that $p_{32}/p_{31} \neq p_{22}/p_{21}$), then we can obtain information on the *curvature* of the consumer's family of indifference curves and we can obtain an estimate for q_1^* and thus we obtain an approximation to the consumer's true nonlinear family of indifference curves. This is the approach pursued by Diewert and Feenstra (2017) (2022). We will explain a special case of their approach in the next section.

A5. The Econometric Estimation of a Rank 1 Substitution Matrix

Konüs and Byushgens (1926) introduced the following functional form for a linearly homogeneous utility function:

(A31) $f(q) \equiv (q\cdot Aq)^{1/2} = (\Sigma_{i=1}^N \Sigma_{j=1}^N a_{ij}q_iq_j)^{1/2}$; $a_{ij} = a_{ji}$; $1 \leq i \leq j \leq N$.

Thus A is an N by N symmetric matrix that contains (N+1)N/2 unknown a_{ij} parameters.[111] Using the utility maximization framework which was described in section A2 above and assuming for the moment that there are no missing products, the estimating equations (A13) become the following system of inverse demand functions:

(A32) $p^t = e^tAq^t/q^t\cdot Aq^t$; $t = 1,\ldots,T.$

[111] The matrix A satisfies certain restrictions which are spelled out in Diewert (1976).

The use of equations (A32) as a basis for a least squares estimation to determine the parameters a_{nm} will lead to aggregate price and quantity levels, $P^t \equiv e^t/Q^t$ and $Q^t \equiv f(q^t)$, that are not invariant to changes in the units of measurement. Thus we convert equations (A32) into the following system of share equations:

(A33) $s_{tn} = q_{tn}(\Sigma_{j=1}^{N} a_{nj}q_{tj})/(\Sigma_{i=1}^{N} \Sigma_{j=1}^{N} a_{ij}q_{ti}q_{tj})$; $t = 1,\ldots,T$; $n = 1,\ldots,N$.

These share equations are valid if there are missing products: if product n is not available in period t, then q_{tn} and s_{tn} both equal 0 and the resulting equation in the system of estimating equations (A33) becomes 0 = 0. Thus the system of estimating equations can accommodate missing (reservation) prices and zero quantities.

We will not attempt to estimate all (N+1)N/2 unknown parameters a_{ij} in the KBD utility function defined by (A31). In order to reduce the number of parameters in the A matrix, we define A as follows:

(A34) $A \equiv \alpha\alpha^T - \beta\beta^T$

where the transposes of the column vectors α and β are defined as $\alpha^T \equiv [\alpha_1,\ldots,\alpha_N]$ and $\beta^T \equiv [\beta_1,\ldots, \beta_N]$.[112] Thus we have reduced the number of unknown parameters in A from (N+1)N/2 to 2N.

With A defined by (A34), the system of share equations (A33) becomes the following system:

(A35) $s_{tn} = q_{tn}[\alpha_n\alpha\cdot q^t - \beta_n\beta\cdot q^t]/[(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2]$; $t = 1,\ldots,T$; $n = 1,\ldots,N$.

Equations (A35) are valid even when there are missing products because when product n is missing in period t, $s_{tn} = q_{tn} = 0$.

The new utility function, $f(q,\alpha,\beta)$ is defined as follows:

(A36) $f(q,\alpha,\beta) \equiv [q^T(\alpha\alpha^T - \beta\beta^T)q]^{1/2} = [(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2]^{1/2}$.

Note that if $\beta = 0_N$, then $f(q,\alpha, 0_N) = [(\alpha\cdot q^t)^2]^{1/2} = \alpha\cdot q^t = \Sigma_{n=1}^{N} \alpha_n q_n$; i.e., the utility function collapses down to the linear utility function that was studied in the previous section.

There are some tricky aspects to the new utility function as compared to the case of a linear utility function. We need to ensure that $(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2 > 0$ so that we can calculate the positive square root, $[(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2]^{1/2}$.[113] We also need to set $\beta_n = 0$ if product n enters the marketplace for the first time in period T. When a product becomes available in the last period of the window of observations for the first time, it

[112] Thus the matrix A will have rank at most equal to 2 with one positive eigenvalue and 1 negative or 0 eigenvalue (if $\beta = 0_N$). With this parameterization of the matrix A, f(q) defined by (A30) will satisfy the appropriate regularity conditions over the set of q such that $q \geq 0_N$ and $Aq \geq 0_N$.

[113] In practice, in solving the nonlinear least squares problem defined below by (A40), we first set $\beta = 0_N$ and solve the nonlinear least squares estimation problem defined by (A26). This gives us starting coefficients for the parameter vector α and we choose starting values for the parameter vector β that are close to 0_N.

turns out that we can estimate the α_n that corresponds to this new product, but we cannot estimate the corresponding β_n so we need to set $\beta_n = 0$ in this situation.[114]

Under the assumption that $f(q,\alpha,\beta)$ defined by (A36) is positive, we can differentiate $f(q,\alpha,\beta)$ with respect to q_n and we obtain the following expression for $f_n(q,\alpha,\beta) \equiv \partial f(q,\alpha,\beta)/\partial q_n$:

$$(A37)\ f_n(q,\alpha,\beta) = [\alpha_n\alpha\cdot q - \beta_n\beta\cdot q]/[(\alpha\cdot q)^2 - (\beta\cdot q)^2]^{1/2}\ .$$

Now assume that there are missing products and let S(t) be the set of products that are purchased in period t for t = 1,…,T. With $f(q,\alpha,\beta)$ defined by (A36), equations (A32) generalized to the case where there are missing products become the following equations:

$$(A38)\ p_{tn} = e^t f_n(q^t,\alpha,\beta)/f(q^t,\alpha,\beta)\ ; \qquad t = 1,\ldots,T;\ n\in S(t).$$

Multiplying both sides of equation t,n in (A38) by q_{tn}/e^t gives us the following system of share equations:

$$\begin{aligned}(A39)\ s_{tn} &= q_{tn}f_n(q^t,\alpha,\beta)/f(q^t,\alpha,\beta)\ ; && t = 1,\ldots,T;\ n\in S(t)\\ &= q_{tn}[\alpha_n\alpha\cdot q^t - \beta_n\beta\cdot q^t]/[(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2] &&\\ &= q_{tn}[\alpha_n\alpha\cdot q^t - \beta_n\beta\cdot q^t]/[(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2]\ ; && t = 1,\ldots,T;\ n = 1,\ldots,N.\end{aligned}$$

Line 3 in equations (A39) follows from line 2 because q_{tn} and s_{tn} equal 0 for $n\notin S(t)$. Estimates for the α_n and β_n parameters are obtained by solving the following least squares minimization problem:[115]

$$(A40)\ \min_{\alpha,\beta} \Sigma_{t=1}^{T} \Sigma_{n=1}^{N} \{s_{tn} - q_{tn}(\alpha_n\alpha\cdot q^t - \beta_n\beta\cdot q^t)/[(\alpha\cdot q^t)^2 - (\beta\cdot q^t)^2]\}^2.$$

It can be seen that a (α^*,β^*) solution to the nonlinear least squares minimization problem (A40) is not unique: if (α^*,β^*) solves (A40), then so does $(\lambda\alpha^*,\lambda\beta^*)$ where λ is a nonzero number. Thus, as was the case for the linear preferences model, we require at least one normalization (like $\alpha_1 = 1$) on the parameters to obtain a unique solution to (A40). As in section A3, we need some additional assumptions on product availability to ensure that a unique solution to (A40) exists, such as: product 1 is present in all periods and each product is present in at least one period. In addition to these restrictions on product availability, if product n is only available in one period in the window of periods, then we need to impose the restriction $\beta_n = 0$ as well. Thus when a new product becomes available in the last period T in the window of periods, it is not possible to estimate the substitution parameter β_n for this product. We need to wait for the product to be available in a subsequent period before we can obtain an estimate for β_n. This provides an argument for producing a revisable CPI when there is product churn.

Once the solution (α^*,β^*) to the least squares minimization problem (A40) has been obtained, the period t aggregate quantity and price levels, Q^t and P^t, are defined as follows:

$$(A41)\ Q^t \equiv f(q^t,\alpha^*,\beta^*) = [(\alpha^*\cdot q^t)^2 - (\beta^*\cdot q^t)^2]^{1/2}\ ;\ P^t \equiv e^t/Q^t\ ; \qquad t = 1,\ldots,T.$$

---

[114] It is not possible to estimate two new parameters, α_n and β_n, for the new product with only one equation to accomplish this estimation task. Recall Figure 1 in the previous section.

[115] The summation of terms $\Sigma_{n=1}^{N}$ can be replaced by the summation $\Sigma_{n\in S(t)}$. In constructing the data for the nonlinear regression, it is more convenient to use the summation $\Sigma_{n=1}^{N}$.

It turns out that the KBD price and quantity levels defined by (A41) satisfy the same 10 tests that the Least Squares Linear Preferences price and quantity levels defined by definitions (A26) in section A2 satisfied.

With the solution (α^*,β^*) to (A40) in hand, we can calculate Hicksian reservation prices p_{tn}^* for the products n that were *not* present in period t using the following equations:

(A42) $p_{tn}^* \equiv e^t f_n(q^t,\alpha^*,\beta^*)/f(q^t,\alpha^*,\beta^*)$; $t = 1,\ldots,T;\ n \notin S(t)$.

Note that these *imputed reservation prices* were not used in the nonlinear least squares minimization problem (A40). They can be calculated once the solution (α^*,β^*) to (A40) has been obtained.

The N by N matrix of second order partial derivatives of $f(q,\alpha^*,\beta^*)$ evaluated at $q = q^t$ is denoted by $\nabla^2 f(q^t,\alpha^*,\beta^*)$ and it is called the *period t inverse substitution matrix*. For a general linearly homogeneous and concave utility f(q), it must be a negative semidefinite matrix that satisfies the restrictions $\nabla^2 f(q^t)q^t = 0_N$. Thus the rank of $\nabla^2 f(q^t)$ is at most N−1. For our particular functional form for $f(q,\alpha^*,\beta^*)$ defined by (A36), the period t inverse substitution matrix is defined as follows:

(A43) $\nabla^2 f(q^t,\alpha^*,\beta^*) \equiv -[f(q^t,\alpha^*,\beta^*)]^{-3}[\alpha^*(\beta^*\cdot q^t)^2 - \beta^*(\alpha^*\cdot q^t)^2][\alpha^*(\beta^*\cdot q^t)^2 - \beta^*(\alpha^*\cdot q^t)^2]^T$.

If $\beta^* = 0_N$, then $\nabla^2 f(q^t,\alpha^*,\beta^*) = 0_N 0_N^T$ which is an N by N matrix of zeros. If α^* and β^* are both nonzero vectors and $\alpha^* \neq \beta^*$, then the period t substitution matrix defined by (A43) will have rank equal to one.

Diewert and Wales (1988) called a functional form for a cost function defined over N products a *semiflexible functional form of rank k* if its matrix of second order partial derivatives had rank k. Using this terminology, our $f(q,\alpha,\beta)$ defined by (A36) is a semiflexible functional form of rank 1.

How would this method of index construction work in real time? It can be seen that the estimates for the α_n and β_n parameters become more accurate as the window of observations expands.[116] Thus a possible way forward would be to proceed as Diewert and Shimizu (2024) proceeded in section 7 of their paper. In particular, their Modified Expanding Window method is recommended. To construct these indexes, start off with a window of 12 months of data and estimate the KBD model described above for this window of observations. For subsequent months, simply expand the window of observations by adding a new month of share equations to the model and get a new set of KBD quantity and price levels, Q^t and P^t. If the index can be revised, then use these new price and quantity levels to form a new set of price indexes back to period 1.[117] These new indexes should be more reliable than the previous indexes because the number of new degrees of freedom will expand faster than the number of new parameters. If the index cannot be

[116] Krsinich (2016; 401) noted that estimates for the quality adjustment parameters α_n in the Time Product Dummy method are not reliably determined until the products have been present in the marketplace for several periods.

[117] It seems "best" to link the price and quantity level estimates for the last period in the current period regression back to the price level in period 1 because the period 1 price level set equal to 1 is the "true" period 1 price level and never changes whereas our new price levels for periods 2 to t−1 are changing as we add another period of data to the regression. Thus we are following Krsinich (2016) in our choice of "best" linking period. We are also following Chessa (2016) in utilizing the idea of an expanding window of observations. The idea of using an infinitely expanding window of observations arose in Diewert (2022) in his discussion of the predicted share method that used dissimilarity measures to link the current period to past periods.

revised, then just use the current window price and quantity levels for the last month in the window as the new last period price and quantity levels.

The work of Diewert and Feenstra (2017) (2022) in estimating KBF preferences shows that the above special case of their methodology (estimating a rank 1 inverse substitution matrix instead of a rank k matrix) can work well if the number of products is relatively small. For a large number of products, the Least Squares Linear Preferences model can be estimated for a large number of products without too much difficulty but estimating the KBD model with large data sets becomes problematic. It seems that data sets with a large number of products would have to be decomposed into data sets containing a smaller number of products. Moreover, eventually, the base period would have to be changed. Finally, it is likely that for many applications, the use of linear models like LSLP or GK or Weighted Time Product Dummy models will generate indexes that are very close to the indexes generated by the KBD model. However, in situations where the number of genuinely new products with added value are appearing, the KBD model should be a useful one.

References

Abdirahman, M., D. Coyle, R. Heys and W. Stewart (2020), "A Comparison of Deflators for Telecommunications Services Output', *Economie et Statistique / Economics and Statistics* 517-518-519, 103–122.

Abdirahman, M., D. Coyle, R. Heys and W. Stewart (2022), "Telecoms Deflators: A Story of Volume and Revenue Weights", *Economie et Statistique / Economics and Statistics* 530-531, 43-59.

Abe, N. and N. Inakura (2024), "Reevaluating Household Real Consumption through Attribute-Specific Price Indices and Actual Consumption", Website: https://stats.unece.org/ottawagroup/meeting/18

Abraham, K. G. and C. Mackie (eds.) (2005), *Beyond the Market; Designing Nonmarket accounts for the United States*, Washington, DC: National Academies Press.

Afriat, S.N. (1972), "The Theory of International Comparisons of Real Income and Prices", pp. 13-69 in *International Comparisons of Prices and Outputs*, D.J. Daly (ed.), NBER, New York: Columbia University Press.

Aizcorbe, A.M. and P.C. Jackman (1993), "The Commodity Substitution Effect in CPI Data,1982-91", *Monthly Labor Review* 116: December, 25-33.

Armknecht, P.A., B.R. Moulton, K,J. Stewart (1994), "Improvements to the Food at Home, Shelter and Prescription Drug Indexes in the U.S. Consumer Price Index", Website: https://stats.unece.org/ottawagroup/meeting/1

Atkinson, T. (2005), *Atkinson Review: Final Report; Measurement of Government Output and Productivity for the National Accounts*, New York: Palgrave Macmillan.

Australian Bureau of Statistics (2016), "Making Greater Use of Transactions Data to Compile the Consumer Price Index", Information Paper 6401.0.60.003, November 29, Canberra: ABS.

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

Balk, B.M. (1981), "A Simple Method for Constructing Price Indices for Seasonal Commodities", *Statistische Hefte* 22 (1), 1–8.

Balk, B.M, (1994), "On the First Step in the Calculation of a Consumer Price Index", website: https://stats.unece.org/ottawagroup/meeting/1

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

Becker, C. (2024), "Is it Still Possible to Publish a Single Consumer price Index that Satisfies all Needs?", Website: : https://stats.unece.org/ottawagroup/meeting/18

Becker, G.S. (1965), "A Theory of the Allocation of Time", *The Economic Journal* 75, 493-517.

Belchev, P. (2024), "A Framework for the Games of Chance Index Calculation in the HICP", Website: https://stats.unece.org/ottawagroup/meeting/18

Berndt, E.R. (1991). *The Practice of Econometrics: Classic and Contemporary*, New York: Addison-Wesley Publishing Co.

Berndt, E.R. (ed.) (2001), "Research on Price Index Measurement: Agendas for the Next Twenty Years", *Journal of Economic and Social Measurement* 27, 99–130.

Boskin, M.J., E. Dulberger, R, Gordon, Z. Griliches, and D.W. Jorgenson (1996), *Toward a More Accurate Measure of the Cost of Living*, Final Report to the U.S. Senate Finance Committee, Washington, D.C. Website: https://www.ssa.gov/history/reports/boskinrpt.html

Boskin, M.J., E. Dulberger, R, Gordon, Z. Griliches, and D.W. Jorgenson (1998), "Consumer Prices, the Consumer Price Index and the Cost of Living", *Journal of Economic Perspectives* 12:1, 3-26.

Bowley, A.L. (1899), "Wages, Nominal and Real", pp. 640-661 in *Dictionary of Political Economy*, R.H.I. Palgrave (ed.), London: Macmillan.

Bowley, A.L. (1901), *Elements of Statistics*, Westminster: P.S. King and Son.

Brandt, N., P. Schreyer and V. Zipperer (2014), "Productivity Measurement with Natural Capital and Bad Outputs", OECD Economics Department Working Papers No. 1154, Paris: OECD.

Brynjolfsson, E., A. Collis, W. E. Diewert, F. Eggers and K. J. Fox (2019), "GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy." NBER Working Paper 25695, Cambridge, MA.

Burnett-Issacs, K.,N. Huang and W.E. Diewert (2020), "Developing Land and Structure Price Indexes for Ottawa Condominium Apartments", *Journal of Official Statistics* 36, 763-802.

Carli, G.-R., (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 (1994), "Recent Developments in the Retail Prices Index", website: https://stats.unece.org/ottawagroup/meeting/1 Also published in *The Statistician* 29:1 (1980), 1-32.

Cavallo, A. (2017), "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers", *American Economic Review* 107:1, 283–303.

Cavallo, A. (2020), "Inflation with Covid Consumption Baskets," NBER Working Paper 27352, Cambridge, MA.

Caves D.W., Christensen, L.R. and Diewert, W.E. (1982), "Multilateral Comparisons of Output, Input, and Productivity using Superlative Index Numbers", *Economic Journal* 92, 73-86.

Chessa, A.G. (2016), "A New Methodology for Processing Scanner Data in the Dutch CPI", *EURONA* 1, 49–69.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau (1971) "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39, 255–256.

Clements, K.W., H.Y. Izan and E.A. Selvanathan (2006), "Stochastic Index Numbers: A Review", *International Statistical Review* 74, 235-270.

Cobb, C. and P.H. Douglas (1928), "A Theory of Production", *American Economic Review* 18, 139-165.

Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 99-117 in *The Dynamics of Automobile Demand*, New York: The General Motors Corporation.

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 Indices", pp. 147-171 in International Conference on Price Indices: Papers and Final Report, First Meeting of the International Working Group on Price Indices, November, Ottawa: Statistics Canada. Website: https://stats.unece.org/ottawagroup/meeting/1

Dalén, J. (1997), "Experiments with Swedish Scanner Data", Website: https://stats.unece.org/ottawagroup/meeting/3

Davies, G.R. (1924), "The Problem of a Standard Index Number Formula", *Journal of the American Statistical Association* 27, 180-188.

de Haan, J. (2004), "Estimating Quality-Adjusted Unit Value Indices: Evidence from Scanner Data," Paper presented at the Seventh EMG Workshop, Sydney, Australia, December 12–14.

de Haan, J. (2008), “Reducing Drift in Chained Superlative Price Indexes for Highly Disaggregated Data”, paper presented at the Economic Measurement Workshop 08, University of New South Wales, December 10.

de Haan, J. (2010), “Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Re-pricing Methods”, *Jahrbücher für Nationökonomie und Statistik* 230, 772-791.

de Haan, J. (2015), “Rolling Year Time Dummy Indexes and the Choice of Splicing Method”, website: https://stats.unece.org/ottawagroup/meeting/14

de Haan, J. and F. Krsinich (2018), “Time Dummy Hedonic and Quality-Adjusted Unit Value Indexes: Do They Really Differ?”, *Review of Income and Wealth* 64:4, 757-776.

de Haan and E. Opperdoes (1997), “Estimation of the Coffee Price Index Using Scanner Data: the Choice of the Micro Index”, website: https://stats.unece.org/ottawagroup/meeting/3

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. (1974), “Applications of Duality Theory,” pp. 106-171 in M.D. Intriligator and D.A. Kendrick (ed.), *Frontiers of Quantitative Economics*, Vol. II, Amsterdam: North-Holland.

Diewert, W.E. (1976), “Exact and Superlative Index Numbers”, *Journal of Econometrics* 4, 114-145.

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

Diewert, W.E. (1983), “The Treatment of Seasonality in a Cost of Living Index”, pp. 1019-1045 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Diewert, W.E. (1987), “Index Numbers”, pp. 767-780 in *The New Palgrave: A Dictionary of Economics*, Vol. 2, J. Eatwell, M. Milgate and P. Newman (eds.), London: The Macmillan Press.

Diewert, W.E. (1992), “Fisher Ideal Output, Input and Productivity Indexes Revisited”, *Journal of Productivity Analysis* 3, 211-248.

Diewert, W.E. (1993a), “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. (1993b), “Symmetric Means and Choice under Uncertainty”', pp. 355-433 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1995a), “Axiomatic and Economic Approaches to Elementary Price Indexes”, website: https://stats.unece.org/ottawagroup/meeting/1
Also available as Working Paper 5104, National Bureau of Economic Research.

Diewert, W.E. (1995b), "Functional Form Problems in Modeling Insurance and Gambling", *The Geneva Papers on Risk and Insurance Theory* 20, 135-150.

Diewert, W.E. (1996), "Comment on CPI Biases", *Business Economics* 31:2, 30-35.

Diewert, W.E. (1997), "Commentary on Mathew D. Shapiro and David W. Wilcox: Alternative Strategies for Aggregating Prices in the CPI", *The Federal Reserve Bank of St. Louis Review,* Vol. 79:3, 127-137.

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

Diewert, W.E. (2001), "Research on Price Index Measurement: Agendas for the Next Twenty Years: Comments", *Journal of Economic and Social Measurement* 27, 105-116.

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. (2004), "On the Stochastic Approach to Linking the Regions in the ICP", Discussion Paper no. 04-16, Department of Economics, The University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2009), "Similarity Indexes and Criteria for Spatial Linking", pp. 183-216 in *Purchasing Power Parities of Currencies*: *Recent Advances in Methods and Applications*, D.S. Prasada Rao (ed.), Cheltenham, UK: Edward Elgar.

Diewert, W.E. (2010), "On the Stochastic Approach to Index Numbers", pp. 235-262 in *Price and Productivity Measurement*, with W.E. Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Trafford Press.

Diewert, W.E. (2014), "The Treatment of Financial Transactions in the SNA: A User Cost Approach", *Eurostat Review of National Accounts and Macroeconomic Indicators* 1, 73-89.

Diewert, W.E. (2018), "Productivity Measurement in the Public Sector", pp. 241-285 in *The Oxford Handbook of Productivity Analysis*, Emile Grifell-Tatje, C.A. Knox Lovell and Robin C. Sickles (eds)., Oxford: Oxford University Press.

Diewert, W.E. (2022), "Quality Adjustment Methods", Chapter 8 in *Consumer Price Index Theory*, Washington D.C.: International Monetary Fund, Washington D.C. Website: https://www.imf.org/en/Data/Statistics/cpi-manual,

Diewert, W.E. (2023), "Scanner Data, Elementary Price Indexes and the Chain Drift Problem", pp. 445-606 in *Advances in Economic Measurement*, D. Chotikapanich, A.N. Rambaldi and N. Rhode (eds.), Singapore: Palgrave Macmillan.

Diewert, W.E. and R.C. Feenstra (2017), "Estimating the Benefits and Costs of New and Disappearing Products", Discussion Paper 17-10, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and R.C. Feenstra (2022), "Estimating the Benefits of New Products", pp. 437-473 in *Big Data For the Twenty-First Century Economic Statistics*, K.G. Abraham, R.S. Jarmin, B.C. Moyer and M.D. Shapiro (eds.), Chicago: University of Chicago Press.

Diewert, W.E., D. Fixler and K. Zieschang (2016), "Problems with the Measurement of Banking Services in a National Accounting Framework", pp. 117-176 in *National Accounting and Economic Growth*, John M. Hartwick (ed.), Cheltenham UK, Edward Elgar Publishing.

Diewert, W.E. and K.J. Fox (2021) "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data," *Journal of Business and Economic Statistics* 40:1, 355-369.

Diewert, W.E. and K.J. Fox (2022a), "Measuring Inflation under Pandemic Conditions", *Journal of Official Statistics* 38:1, 255-285.

Diewert, W.E. and K.J. Fox (2022b), "Measuring Real Consumption and CPI Bias under Lockdown Conditions," *Canadian Journal of Economics* 55:S1, 480-502.

Diewert, W.E. and K.J. Fox (2024), "Alternative Output, Input and Income Concepts for the Production Accounts," *Journal of Productivity Analysis*, forthcoming.

Diewert, W.E., J. de Haan and R. Hendriks (2015), "Hedonic Regressions and the Decomposition of a House Price index into Land and Structure Components", *Econometric Reviews*, 34,1-2, 106-126.

Diewert, W.E., K.J. Fox and P. Schreyer (2017), "The Allocation and Valuation of Time", Discussion Paper 17-04, School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E., K.J. Fox and P. Schreyer (2022), "Experimental Economics and the New Commodities Problem", *Review of Income and Wealth* 68, 895-905.

Diewert, W.E. and C. Shimizu (2015), "Residential Property Price Indexes for Tokyo", *Macroeconomic Dynamics* 19, 1659-1714.

Diewert, W.E. and C. Shimizu (2017), "Hedonic Regression Models for Tokyo Condominium Sales", *Regional Science and Urban Economics* 60, 300-315.

Diewert, W.E. and C. Shimizu (2022), Residential Property Price Indexes: Spatial Coordinates versus Neighbourhood Dummy Variables", *Review of Income and Wealth* 68:3, 770-796.

Diewert, W.E. and C. Shimizu (2024), "Scanner Data, Product Churn and Quality Adjustment", 18th Ottawa Group Meeting, Ottawa, May 13.

Diewert, W.E. and T.J. Wales (1988), "A Normalized Quadratic Semiflexible Functional Form", *Journal of Econometrics* 37, 327-42.

Diewert, W.E. and V. Zelenyuk (2024), "On the Problem of the Purchasing Power of Money by A. A. Konüs and S. S. Byushgens: Translation and Commentary", *Journal of Productivity Analysis, forthcoming.*

Dulberger, E. (1989), "The Application of a Hedonic Model to a Quality Adjusted Price Index for Computer Processors", pp. 37-75 in *Technology and Capital Formation*, D.W. Jorgenson and W. Landau (eds), Cambridge MA: MIT Press.

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

Edgeworth, F.Y. (1888), "Some New Methods of Measuring Variation in General Prices", *Journal of the Royal Statistical Society* 51, 346-368.

Edgeworth, F. Y. (1923), "The Doctrine of Index Numbers According to Mr. Correa Walsh", *The Economic Journal* 11, 343-351.

Eichhorn, W. (1976), "Fisher's Tests Revisited", *Econometrica* 44, 247-256.

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

Eichhorn, W. and J. Voeller (1976), *Theory of the Price Index*, Lecture Notes in Economics and Mathematical Systems, Vol. 140, Berlin: Springer-Verlag.

Eltetö, Ö., and Köves, P. (1964), "On a Problem of Index Number Computation Relating to International Comparisons", (in Hungarian), *Statisztikai Szemle* 42, 507-518.

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*, Robert C. Feenstra and Matthew D. Shapiro (eds.), Studies in Income and Wealth Volume 64, Chicago: The University of Chicago Press.

Fisher, I. (1911), *The Purchasing Power of Money*, London: Macmillan.

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

Fixler, D., J.S. Greenlees, D. Fenwick, R. Lowe and M. Silver (2003), "Roundtable Discussion", pp. 111-120 in *Scanner Data and Price Indexes*, R.C. Feenstra and M.D. Shapiro (eds.), NBER Studies in Income and Wealth 64, Chicago: University of Chicago Press.

Frisch, R. (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica* 4, 1-39.

Fox, K.J.. P. Levell and M. O'Connell (2024), "Inflation Measurement with High Frequency Data", 18th Ottawa Group Meeting, Ottawa, May 13.

Geary, R.G. (1958), "A Note on Comparisons of Exchange Rates and Purchasing Power between Countries", *Journal of the Royal Statistical Society Series A* 121, 97-99.

Gini, C. (1931), "On the Circular Test of Index Numbers", *Metron* 9:9, 3-24.

Greenlees, J.S. (2006), "The BLS Response to the Boskin Commission Report", *International Productivity Monitor* Number 12 (Spring), 23-41.

Griliches, Z. and I. Cockburn (1994), "Generics and New Goods in Pharmaceutical Price Indexes", *American Economic Review* 84; 1213-1232.

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

Hausman, J.A. (1996), "Valuation of New Goods under Perfect and Imperfect Competition", pp. 20 -236 in *The Economics of New Goods*, T.F. Bresnahan and R.J. Gordon (eds.), Chicago: University of Chicago Press.

Hausman, J.A. (1999), "Cellular Telephone, New Products and the CPI", *Journal of Business and Economic Statistics* 17:2, 188-194.

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*, R.C. Feenstra and M.D. Shapiro (eds.), NBER Studies in Income and Wealth 64, Chicago: University of Chicago Press.

Haworth, M. (1994), "Sampling and Data Capture Issues in the CPI Construction", Website: https://stats.unece.org/ottawagroup/meeting/1

Hicks, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, 105-124.

Hicks, J.R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Hill, P. (1988), ''Recent Developments in Index Number Theory and Practice'', *OECD Economic Studies* 10, 123–148.

Hill, P. (1993), ''Price and Volume Measures'', pp. 379-406 in *System of National Accounts 1993*, Brussels/Luxembourg, New York, Paris, New York, and Washington, D.C.: Commission of the European Communities, IMF, OECD, World Bank and United Nations.

Hill, P. (1996), *Inflation Accounting: A Manual on National Accounting under Conditions of High Inflation*, Paris: OECD.

Hill, P. (2009); "Consumption of Own Production and Cost of Living Indices", pp. 429-444 in: W.E. Diewert, J. S. Greenlees and Ch. R. Hulten (editors); *Price Index Concepts and Measurement*; NBER Studies in Income and Wealth, Volume 70, Chicago; University of Chicago Press.

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

IMF/ILO/UNECE/Eurostat/OECD/The World Bank (2020), *Consumer Price Index Manual: Concepts and Methods 2020*, B. Graff(ed.), Washington D.C.: The International Monetary Fund.

Inklaar, R. and W.E. Diewert (2016), “Measuring Industry Productivity and Cross-Country Convergence”, *Journal of Econometrics* 191, 426-433.

Ioannides, C. and M. Silver (1997), “Chained, Exact and Superlative Hedonic Price Changes: Estimates from Micro Data”, Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

Ivancic, L., K.J. Fox and W.E. Diewert (2009), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, website: https://stats.unece.org/ottawagroup/meeting/11

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 (2011a), “Understanding Price Variation across Stores and Supermarket Chains: Some Implications for CPI Aggregation Methods”, Website: https://stats.unece.org/ottawagroup/meeting/12

Ivancic, L. and K.J. Fox (2011b), “Can Dissimilarity Indexes Resolve the Issue of When to Chain Price Indexes?”, website: https://stats.unece.org/ottawagroup/meeting/12

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

Ivancic, L. and K.J. Fox (2013b), “Can Dissimilarity Indexes Resolve the Issue of When to Chain Price Indexes?”, *Economic Letters* 118:1, 6-9.

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.

Jevons, W.S., (1884), “A Serious Fall in the Value of Gold Ascertained and its Social Effects Set Forth (1863)”, pp. 13-118 in *Investigations in Currency and Finance*, London: Macmillan and Co.

Keynes, J.M. (1930), *Treatise on Money*, Vol. 1, London: Macmillan.

Khamis, S.H. (1970), “Properties and Conditions for the Existence of a New Type of Index Number”, *Sankhya B* 32, 81-98.

Konüs, A.A. (1924), “The Problem of the True Index of the Cost of Living”, translated in *Econometrica* 7, (1939), 10-29.

Konüs, A.A. and S.S. Byushgens (1926), “K probleme pokupatelnoi sily deneg”, *Voprosy Konyunktury* 2(1), 151-172.

Krsinich, F. (2016), “The FEWS Index: Fixed Effects with a Window Splice”, *Journal of Official Statistics* 32, 375-404.

Laspeyres, E. (1871), “Die Berechnung einer mittleren Waarenpreissteigerung”, *Jahrbücher für Nationalökonomie und Statistik* 16, 296-314.

Lowe, J. (1823), *The Present State of England in Regard to Agriculture, Trade and Finance*, second edition, London: Longman, Hurst, Rees, Orme and Brown.

Lowe, R. (1998), “Televisions: Quality Changes and Scanner Data”, website: https://stats.unece.org/ottawagroup/meeting/4

Manser, M. and R.J. McDonald (1988), “An Analysis of the Substitution Bias in Measuring Inflation”, *Econometrica* 56, 909-930.

Marshall, A. (1887), “Remedies for Fluctuations of General Prices’, *Contemporary Review* 51, 355-375.

Moulton, B.R. (1993) “Basic Components of the CPI: Estimation of Price Changes,” *Monthly Labor Review* 116 December, 13-24.

Mudgett, B.D. (1955), “The Measurement of Seasonal Movements in Price and Quantity Indexes”, *Journal of the American Statistical Association* 50, 93-98.

Obst, C. (1999), “Draft Chapter Outline for the Revised CPI Manual”, Website: https://stats.unece.org/ottawagroup/meeting/5

Ottawa Group (1994), “Report of the Meeting”, Website: https://stats.unece.org/ottawagroup/meeting/1

Paasche, H. (1874), “Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen”, *Jahrbücher für Nationalökonomie und Statistik* 12, 168-178.

Persons, W.M. (1921), “Fisher's Formula for Index Numbers”, *Review of Economics and Stat*istics 3:5, 103-113.

Persons, W.M. (1928), “The Effect of Correlation Between Weights and Relatives in the Construction of Index Numbers”, *The Review of Economics and Statistics* 10:2, 80-107.

Pollak, R.A. (1983), “The Theory of the Cost-of-Living Index”, pp. 87-161 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada; reprinted as pp. 3-52 in R.A. Pollak (1989), *The Theory of the Cost-of-Living Index*, Oxford: Oxford University Press.

Price Statistics Review Committee (1961), *The Price Statistics of the Federal Government*, Report to the Office of Statistical Standards, Bureau of the Budget, Cambridge MA.: NBER.

Rao, D.S. Prasada (1995), “On the Equivalence of the Generalized Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Comparisons”, Working Paper No. 5, Centre for International Comparisons, University of Pennsylvania, Philadelphia.

Rao, D.S. Prasada (2005), "On the Equivalence of the Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", *Review of Income and Wealth* 51:4, 571-580.

Reinsdorf, M. (1993) "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," pp. 227-254 in M.C. Foss, M.E. Manser and A.H. Young (eds.), *Price Measurements and Their Uses*, NBER Studies in Income and Wealth 57, Chicago: University of Chicago Press.

Reinsdorf, M. (1998) "Formula Bias and Within-Stratum Substitution Bias in the US CPI", *Review of Economics and Statistics* 58:2, 175-187.

Reinsdorf, M. (1999), "Using Scanner Data to Construct CPI Basic Component Indexes", *Journal of Business & Economic Statistics* 17:2, 152-160.

Reinsdorf, M. and P. Schreyer (2020), "Measuring Consumer Inflation in a Digital Economy", pp. 369-362 in *Measuring Economic Growth and Productivity*, B.M. Fraumeni (ed.), New York: Elsevier.

Reinsdorf, M. and B.R. Moulton (1997), "The Construction of Basic Components of Costof-Living Indexes", pp. 397-436 in T.F. Bresnahan and R.J. Gordon (eds.), *The Economics of New Goods*, NBER Studies in Income and Wealth 58, (Chicago: University of Chicago Press.

Saglio, A. (1994), "Comparative Changes in Average Price and a Price Index: Two Case Studies", website: https://stats.unece.org/ottawagroup/meeting/1

Schreyer, P. (2010), "Towards Measuring the Volume Output of Education and Health Services: A Handbook", OECD Statistics Working Paper No. 31. Paris: OECD.

Schreyer, P. (2012). "Output, Outcome and Quality Adjustment in Measuring Health and Education Services", *Review of Income and Wealth* 58(2), 257–278.

Schreyer, P. and W.E. Diewert (2014), "Household Production, Leisure and Living Standards", pp. 89-114 in *Measuring Economic Sustainability and Progress*, D.W. Jorgenson, J.S. Landefeld and P. Schreyer (eds.), Chicago IL: University of Chicago Press.

Schultz (Szulc), B.J. (1994), "Choice of Price Index Formulae at the Micro-aggregation Level: The Canadian Empirical Evidence", website: https://stats.unece.org/ottawagroup/meeting/1

Selvanathan, E.A. and D.S. Prasada Rao (1994), *Index Numbers: A Stochastic Approach*, Ann Arbor: The University of Michigan Press.

Sellwood, D.J. (1994), "Constraining Macro Effects of the Use of Different Methods at the Basic Level in the Project to Harmonise Consumer Price Indices in the European Union", website: https://stats.unece.org/ottawagroup/meeting/1

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

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, 427-435.

Stiglitz, J.E., A. Sen and J.P. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Eurostat.https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report.

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland Publishing.

Triplett, J.E. (1998), "Elementary Indexes for a Consumer Price Index", Website: https://stats.unece.org/ottawagroup/meeting/4

Triplett, J.E. (1999), "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?", website: https://stats.unece.org/ottawagroup/meeting/5

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

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

Stoevska, V. (2018), "Official Consumer Price Indices—Historical Perspective", unpublished paper available at: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/ILO.pdf

Stone, R. (1956), *Quantity and Price Indexes in National Accounts*, Paris: OECD.

Szulc (Schultz), B.J. (1964), "Indices for Multiregional Comparisons", (in Polish), *Przeglad Statystyczny* 3, 239-254.

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

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland Publishing.

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

Törnqvist, L. and E. Törnqvist (1937), Vilket är förhällandet mellan finska markens och svenska kronans köpkraft?", *Ekonomiska Samfundets Tidskrift* 39, 1-39 reprinted as pp. 121-160 in *Collected Scientific Papers of Leo Törnqvist*, Helsinki: The Research Institute of the Finnish Economy, 1981.

Turvey, R. (1979), "The Treatment of Seasonal Items in Consumer Price Indices", *Bulletin of Labour Statistics*, Fourth Quarter, International Labour Office, Geneva, 13-33.

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

von Auer, L. (2019), "The Nature of Chain Drift: Implications for Scanner Data Price Indices", Website: https://stats.unece.org/ottawagroup/meeting/16

von Auer, L. (2024), "Inflation Measurement in the Presence of Stockpiling and Smoothing of Consumption", Research Papers in Economics, No. 2/24, Department of Economics, University of Trier.

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

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

Walsh, C.M. (1924), "Professor Edgeworth's Views on Index Numbers", *Quarterly Journal of Economics* 38, 500–519.

Westergaard, H. (1890), *Die Grundzüge der Theorie der Statistik*, Jena: Fischer.

Woolford K. (1994), "A Pragmatic Approach to the Selection of Appropriate Index Formulae", Website: https://stats.unece.org/ottawagroup/meeting/1

Zhang, L.-C., I. Johansen and R. Nygaard (2019), "Tests for Price Indices in a Dynamic Item Universe", *Journal of Official Statistics* 35:3, 683-697.