CONSUMER PRICE INDEX THEORY

Chapter Titles

CHAPTER 4: STOCHASTIC APPROACHES TO INDEX NUMBER THEORY[1]

Draft: April 16, 2021

Erwin Diewert University of British Columbia

Table of Contents

[1] The author thanks Carsten Boldsen, Jan de Haan, Ronald Johnson, Thomas McDowell, Jens Mehrhoff and Chihiro Shimizu for helpful comments.

1. Introduction

"In drawing our averages the independent fluctuations will more or less destroy each other; the one required variation of gold will remain undiminished." W. Stanley Jevons (1884; 26).

The stochastic approach to the determination of the price index can be traced back to the work of Jevons and Edgeworth over a hundred years ago[2]. In section 2 below, the work of these early pioneers will be explained. Basically, their approach was to take an *average of the price ratios* pertaining to two periods as their index number. However, Keynes (1930) was critical of this approach to index number theory because it did not take into account the *economic importance* of each commodity in the index. Thus in section 3, the weighted stochastic approach of Theil (1967) will be explained. This approach does take into account the economic importance of each commodity.

In section 4, an introduction to the *time product dummy stochastic approach* to index number theory will be presented. Using this approach, the focus is on providing representative *price levels* for two periods.[3] Weighted versions of this approach are described in section 5.

A weakness of the material presented in this chapter is that it is assumed that all prices are positive. In chapters 7 and 8, this assumption will be relaxed. The reason for postponing a discussion of index number theory when there are missing prices is that it is useful to develop the economic approach to index number theory before discussing the problem of missing prices. The missing price problem and the treatment of new and disappearing products will be studied in some detail in chapters 7 and 8. The economic approach to index number theory will be discussed in chapters 5 and 8.

2. Early Unweighted Stochastic Approaches to Bilateral Index Number Theory

The basic idea behind the early stochastic approaches to index number theory is that each price relative, p_n^1/p_n^0 for $n = 1,2,\ldots,N$ can be regarded as an estimate of a common inflation rate α between periods 0 and 1; i.e., it is assumed that

(1) $p_n^1/p_n^0 = \alpha + \varepsilon_n$; $n = 1,2,\ldots,N$

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

Unfortunately, P_C does not satisfy the time reversal test,[4] i.e., $P_C(p^1,p^0) \neq 1/ P_C(p^0,p^1)$.

---

[2] For references to the literature, see Diewert (1993; 37-38) (2010).

[3] The extension of the price levels approach to many periods will be undertaken in Chapter 7.

[4] 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. The upward bias of the Carli index will be illustrated in Chapter 6.

Now suppose that the stochastic specification of the error terms is changed; i.e., assume that the logarithm of each price relative, $\ln(p_n^1/p_n^0)$, is an unbiased estimate of the logarithm of the inflation rate between periods 0 and 1, β say. Thus we have:

(3) $\ln(p_n^1/p_n^0) = \beta + \varepsilon_n$; $n = 1,2,\ldots,N$

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 [5] is the *Jevons* (1865) *price index* P_J defined as:

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

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 a fatal flaw: each price relative p_n^1/p_n^0 is regarded as being *equally important* and is given an equal weight in the index number formulae (2) and (4). Keynes was particularly critical of this *unweighted stochastic approach* to index number theory. He directed the following criticism towards this approach, which was vigorously advocated by Edgeworth (1923):

"Nevertheless I venture to maintain that such ideas, which I have endeavoured to expound above as fairly and as plausibly as I can, are root-and-branch erroneous. The 'errors of observation', the 'faulty shots aimed at a single bull's eye' conception of the index number of prices, Edgeworth's 'objective mean variation of general prices', is the result of confusion of thought. There is no bull's eye. There is no moving but unique centre, to be called the general price level or the objective mean variation of general prices, round which are scattered the moving price levels of individual things. There are all the various, quite definite, conceptions of price levels of composite commodities appropriate for various purposes and inquiries which have been scheduled above, and many others too. There is nothing else. Jevons was pursuing a mirage.

What is the flaw in the argument? In the first place it assumed that the fluctuations of individual prices round the 'mean' are 'random' in the sense required by the theory of the combination of independent observations. In this theory the divergence of one 'observation' from the true position is assumed to have no influence on the divergences of other 'observations'. But in the case of prices, a movement in the price of one commodity necessarily influences the movement in the prices of other commodities, whilst the magnitudes of these compensatory movements depend on the magnitude of the change in expenditure on the first commodity as compared with the importance of the expenditure on the commodities secondarily affected. Thus, instead of 'independence', there is between the 'errors' in the successive 'observations' what some writers on probability have called 'connexity', or, as Lexis expressed it, there is 'sub-normal dispersion'.

We cannot, therefore, proceed further until we have enunciated the appropriate law of connexity. But the law of connexity cannot be enunciated without reference to the relative importance of the commodities

---

[5] Greenlees (1999) pointed out that although $(1/N)\sum_{n=1}^{N} \ln(p_n^1/p_n^0)$ is an unbiased estimator for β, the corresponding exponential of this estimator, P_J defined by (4), will generally *not* be an unbiased estimator for α under our stochastic assumptions. To see this, let $x_n = \ln(p_n^1/p_n^0)$. Taking expectations, we have: $Ex_n = \beta = \ln(\alpha)$. Thus each x_n is an unbiased estimator of overall log price change. If we wish to measure overall price change α instead of log price change β, then use $y_n \equiv \exp[x_n]$ as an estimator for α. Define the positive, convex function f of one variable x by $f(x) \equiv e^x$. By Jensen's (1906) inequality, we have $Ef(x) \geq f(Ex)$. Letting x equal the random variable x_n, this inequality becomes: $E(p_n^1/p_n^0) = Ef(x_n) \geq f(Ex_n) = f(\beta) = e^\beta = e^{\ln\alpha} = \alpha$. Thus for each n, we have $E(p_n^1/p_n^0) \geq \alpha$, and it can be seen that the Jevons price index defined by (4) will generally have a upward bias from a statistical point of view. However, if we make the measurement of *average log price change* our estimation target, then the Jevons index is no longer biased for this alternative target index.

affected—which brings us back to the problem that we have been trying to avoid, of weighting the items of a composite commodity." John Maynard Keynes (1930; 76-77).

One of the points Keynes makes in the above quotation is that prices in the economy are not independently distributed from each other and from quantities. In current macroeconomic terminology, we can interpret Keynes as saying that a macroeconomic shock will be distributed across all prices and quantities in the economy through the normal interaction between supply and demand; i.e., through the workings of the general equilibrium system. Thus Keynes seemed to be leaning towards the economic approach to index number theory (even before it was developed to any great extent), where quantity movements are functionally related to price movements. A second point that Keynes made in the above quotation is that there is no such thing as the inflation rate; there are only price changes that pertain to well specified sets of commodities or transactions; i.e., the domain of definition of the price index must be carefully specified. A final point that Keynes made is that price movements must be weighted by their *economic importance*; i.e., by quantities or expenditures.[6]

In addition to the above theoretical criticisms, Keynes also made the following strong empirical attack on Edgeworth's unweighted stochastic approach:

"The Jevons—Edgeworth "objective mean variation of general prices', or 'indefinite' standard, has generally been identified, by those who were not as alive as Edgeworth himself was to the subtleties of the case, with the purchasing power of money—if only for the excellent reason that it was difficult to visualise it as anything else. And since any respectable index number, however weighted, which covered a fairly large number of commodities could, in accordance with the argument, be regarded as a fair approximation to the indefinite standard, it seemed natural to regard any such index as a fair approximation to the purchasing power of money also.

Finally, the conclusion that all the standards 'come to much the same thing in the end' has been reinforced 'inductively' by the fact that rival index numbers (all of them, however, of the wholesale type) have shown a considerable measure of agreement with one another in spite of their different compositions. … On the contrary, the tables given above (pp. 53,55) supply strong presumptive evidence that over long periods as well as over short periods the movements of the wholesale and of the consumption standards respectively are capable of being widely divergent." John Maynard Keynes (1930; 80-81).

In the above quotation, Keynes noted that the proponents of the unweighted stochastic approach to price change measurement were comforted by the fact that all of the then existing (unweighted) indexes of wholesale prices showed broadly similar movements. However, Keynes showed empirically that these wholesale price indexes moved quite differently than his consumer price indexes.[7]

In order to overcome the Keynsian criticisms of the unweighted stochastic approach to index numbers, it is necessary to:

---

[6] An empirical example in the Appendix to chapter 6 will illustrate the importance of weighting. This example also illustrates that there can be substantial differences between the Jevons and Carli indexes.

[7] Using the OECD national accounts data for the last five decades, some broad trends in the rates of increase in prices for the various components of GDP can be observed: rates of increase for the prices of internationally traded goods have been the lowest, followed by the prices of reproducible capital goods, followed by consumer prices, followed by wage rates. From other sources, land prices have shown the highest rate of price increase over this period. Of course, if a country adjusts the price of computer related equipment for quality improvements, then the aggregate price of capital machinery and equipment tends to move *downwards* in recent years. Another source of long run differential rates of price increase is due to the fact that service prices tend to increase more rapidly than product prices. Thus there are long term systematic differences in price movements over different domains of definition.

- have a definite domain of definition for the index number and
- weight the price relatives by their economic importance.

On the second dot point above, it should be noted that the issue of weighting price ratios came up early in the history of index number theory. Young (1812) advocated some form of rough weighting of the price relatives according to their relative value over the period being considered but the precise form of the required value weighting was not indicated.[8] However, it was Walsh (1901; 83-121) (1921; 81-90) who stressed the importance of weighting the individual price ratios, where the weights are functions of the associated values for the commodities in each period and each period is to be treated symmetrically in the resulting formula:

"What we are seeking is to average the variations in the exchange value of one given total sum of money in relation to the several classes of goods, to which several variations [price ratios] must be assigned weights proportional to the relative sizes of the classes. Hence the relative sizes of the classes at both the periods must be considered." Correa Moylan Walsh (1901; 104).

"Commodities are to be weighted according to their importance, or their full values. But the problem of axiometry always involves at least two periods. There is a first period and there is a second period which is compared with it. Price variations[9] have taken place between the two, and these are to be averaged to get the amount of their variation as a whole. But the weights of the commodities at the second period are apt to be different from their weights at the first period. Which weights, then, are the right ones—those of the first period or those of the second? Or should there be a combination of the two sets? There is no reason for preferring either the first or the second. Then the combination of both would seem to be the proper answer. And this combination itself involves an averaging of the weights of the two periods." Correa Moylan Walsh (1921; 90).

In the following section, Theil's solution to the weighting problem will be described.

3. The Weighted Stochastic Approach of Theil

"It might seem at first sight as if simply every price quotation were a single item, and since every commodity (any kind of commodity) has one price-quotation attached to it, it would seem as if price-variations of every kind of commodity were the single item in question. This is the way the question struck the first inquirers into price-variations, wherefore they used simple averaging with even weighting. But a price-quotation is the quotation of the price of a generic name for many articles; and one such generic name covers a few articles, and another covers many. … A single price-quotation, therefore, may be the quotation of the price of a hundred, a thousand, or a million dollar's worths, of the articles that make up the commodity named. Its weight in the averaging, therefore, ought to be according to these money-unit's worth." Correa Moylan Walsh (1921; 82-83).

Theil (1967; 136-137) proposed a solution to the lack of weighting in the Jevons index defined by (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 $s_n^0 \equiv p_n^0 q_n^0 / p^0 \cdot q^0$, the period 0 expenditure share for commodity n. The resulting overall mean (period 0 weighted) logarithmic

[8] Walsh (1901; 84) refers to Young's contributions as follows: "Still, although few of the practical investigators have actually employed anything but even weighting, they have almost always recognized the theoretical need of allowing for the relative importance of the different classes ever since this need was first pointed out, near the commencement of the century just ended, by Arthur Young. … Arthur Young advised simply that the classes should be weighted according to their importance."

[9] A price variation is a price ratio or price relative in Walsh's terminology.

price change is $\sum_{n=1}^{N} s_n^0 \ln(p_n^1/p_n^0)$. Now repeat the above mental experiment and draw price relatives at random in such a way that each dollar of expenditure in period 1 has an equal probability of being selected. This leads to the overall mean (period 1 weighted) logarithmic price change of $\sum_{n=1}^{N} s_n^1 \ln(p_n^1/p_n^0)$. Each of these measures of overall logarithmic price change is equally valid so it is best to take a symmetric average of the two measures in order to obtain a final single measure of overall logarithmic price change[10]. Theil[11] 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) $\ln P_T(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^{N} (½)(s_n^0 + s_n^1)\ln(p_n^1/p_n^0)$.

It is possible give the following statistical interpretation of the right hand side of (5). Define the nth logarithmic price ratio r_n by:

(6) $r_n \equiv \ln(p_n^1/p_n^0)$; $n = 1,\ldots,N$.

Now define the discrete random variable, R say, as the random variable that can take on the values r_n with probabilities $\rho_n \equiv (1/2)(s_n^0 + s_n^1)$ for $n = 1,\ldots,N$. Note that since each set of expenditure shares, s_n^0 and s_n^1, sum to one, the probabilities ρ_n will also sum to one. It can be seen that the expected value of the discrete random variable R is

(7) $E[R] \equiv \sum_{n=1}^{N} \rho_n r_n = \sum_{n=1}^{N} (½)(s_n^0 + s_n^1)\ln(p_n^1/p_n^0) = \ln P_T(p^0,p^1,q^0,q^1)$

using (5) and (6). Thus the logarithm of the index P_T can be interpreted as *the expected value of the distribution of the logarithmic price ratios* in the domain of definition under consideration, where the N discrete price ratios in this domain of definition are weighted according to Theil's probability weights, $\rho_n \equiv (½)(s_n^0 + s_n^1)$ for $n = 1,\ldots,N$.

Taking antilogs of both sides of (7), we obtain the Törnqvist (1936) (1937) Theil price index, P_T. This index number formula has a number of good properties.[12] In particular, P_T satisfies the *time reversal test*:

(8) $P(p^1,p^0,q^1,q^0) = 1/P(p^0,p^1,q^0,q^1)$.

The price index P_T also satisfies the following *linear homogeneity test in current period prices*:

[10] "The [asymmetric] price index (1.6) has certain merits. It is, for example, independent of the units in which we measure the quantities of the various commodities (tons, gallons, etc.). It has the disadvantage, however, of being one sided in the sense that it is based on the distribution of expenditure in the *a*th region. We could equally well apply our random selection procedure to the *b*th region, in which case, w_{ia} is replaced by w_{ib} in (1.5) and (1.6). We must conclude that (1.6) is an asymmetric index number, which is a disadvantage because the question asked is symmetric: If the price level of the *b*th region exceeds that of the *a*th by a factor 1.2, say, we should expect that the price level of the latter region exceed that of the former by a factor 1/1.2." Henri Theil (1967; 137).

[11] "The price index number defined in (1.8) and (1.9) uses the n individual logarithmic price differences as the basic ingredients. They are combined linearly by means of a two stage random selection procedure: First, we give each region the same chance ½ of being selected, and second, we give each dollar spent in the selected region the same chance ($1/m_a$ or $1/m_b$) of being drawn." Henri Theil (1967; 138).

[12] See section 5 of Chapter 3 for a listing of the test properties of the Törnqvist Theil index.

(9) $P(p^0,\lambda p^1,q^0,q^1) = \lambda P(p^0,p^1,q^0,q^1)$

for all positive numbers λ and strictly positive vectors p^0,p^1,q^0,q^1. Thus if all period one prices increase by the same positive number λ and if the price index P satisfies the test (9), then the price index increases by this same scalar factor λ.

The time reversal test and the linearly homogeneous test can be used to justify Theil's (arithmetic) method of forming an average of the two sets of expenditure shares in order to obtain his probability weights, $\rho_n \equiv (½)[s_n^0 + s_n^1]$ for n = 1,…,N. Consider the following *symmetric mean* class of Theil type *logarithmic index number formulae*:

(10) $\ln P_{ml}(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^N m(s_n^0,s_n^1)\ln(p_n^1/p_n^0)$

where $m(s_n^0,s_n^1)$ is a homogeneous symmetric mean of the period 0 and 1 expenditure shares, s_n^0 and s_n^1 respectively. In order for P_{ml} to satisfy the time reversal test, it is necessary that the mean function m be symmetric. In order for the weights in (10) to sum to one so that the linear homogeneity test is satisfied and the weights can be interpreted as probability weights, it can be shown that the homogeneous symmetric mean function m(a,b) that appears in (10) *must* be the arithmetic mean.

The stochastic approach of Theil has another nice symmetry property. Instead of considering the distribution of the price ratios $r_n = \ln(p_n^1/p_n^0)$, we could also consider the distribution of the *reciprocals* of these price ratios, say:

(11) $$\begin{aligned} t_n &\equiv \ln(p_n^0/p_n^1); \qquad n = 1,\ldots,N \\ &= \ln(p_n^1/p_n^0)^{-1} \\ &= -\ln(p_n^1/p_n^0) \\ &= -r_n \end{aligned}$$

where the last equality follows using definitions (6). We can still associate the symmetric probability, $\rho_n \equiv (1/2)[s_n^0 + s_n^1]$, with the nth reciprocal logarithmic price ratio t_n for n = 1,…,N. Now define the discrete random variable, T say, as the random variable that can take on the values t_n with probabilities $\rho_n \equiv (1/2)(s_n^0 + s_n^1)$ for n = 1,…,N. It can be seen that the expected value of the discrete random variable T is

(12) $$\begin{aligned} E[T] &\equiv \textstyle\sum_{n=1}^N \rho_n t_n \\ &= -\textstyle\sum_{n=1}^N \rho_n r_n && \text{using (11)} \\ &= -E[R] && \text{using (7)} \\ &= -\ln P_T(p^0,p^1,q^0,q^1). \end{aligned}$$

Thus it can be seen that the distribution of the random variable T is equal to minus the distribution of the random variable R. Hence it does not matter whether we consider the distribution of the original logarithmic price ratios, $r_n \equiv \ln(p_n^1/p_n^0)$, or the distribution of their reciprocals, $t_n \equiv \ln(p_n^0/p_n^1)$: we obtain essentially the same stochastic theory.

It is possible to consider weighted stochastic approaches to index number theory where we look at the distribution of price ratios, p_n^1/p_n^0, rather than the distribution of the logarithmic price ratios, $\ln(p_n^1/p_n^0)$. Thus, again following in the footsteps of Theil, 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 $s_n^0 \equiv p_n^0 q_n^0/p^0 \cdot q^0$, the period 0 expenditure share for commodity n. Now the overall mean (period 0 weighted) price change is:

$$(13)\ P_L(p^0,p^1,q^0,q^1) = \sum_{n=1}^N s_n^0(p_n^1/p_n^0),$$

which turns out to be the Laspeyres price index, P_L defined in Chapter 2.

Now repeat the above mental experiment and draw price relatives at random in such a way that each dollar of expenditure in period 1 has an equal probability of being selected. This leads to the overall mean (period 1 weighted) price change equal to:

$$(14)\ P_{Pal}(p^0,p^1,q^0,q^1) = \sum_{n=1}^N s_n^1(p_n^1/p_n^0).$$

The right hand side of (14) is known as the Palgrave (1886) index number formula, P_{Pal}.[13]

It can be verified that neither the Laspeyres nor Palgrave price indexes satisfy the time reversal test, (8). Again following in the footsteps of Theil, we might try to obtain a formula that satisfied the time reversal test by taking a symmetric average of the two sets of shares. Consider the following class of *symmetric mean index number formulae*:

$$(15)\ P_{SM}(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^N m(s_n^0,s_n^1)(p_n^1/p_n^0)$$

where $m(s_n^0,s_n^1)$ is a homogeneous symmetric mean of the period 0 and 1 expenditure shares, s_n^0 and s_n^1 respectively. However, in order to interpret the right hand side of (15) as an expected value of the price ratios p_n^1/p_n^0, it is necessary that

$$(16)\ \sum_{n=1}^N m(s_n^0,s_n^1) = 1.$$

However, in order to satisfy (16), m cannot be a symmetric geometric or harmonic mean or any of the commonly used homogeneous symmetric means. In fact, the only simple homogeneous symmetric mean that satisfies (16) is the arithmetic mean.[14] With this choice of m, (15) becomes the following (unnamed) index number formula, P_U:

$$(17)\ P_U(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^N (\tfrac{1}{2})(s_n^0 + s_n^1)(p_n^1/p_n^0).$$

Unfortunately, the unnamed index P_U does not satisfy the time reversal test either.

The above considerations explain why Theil's stochastic index number formula P_T seems to be the preferred member of this class of index number formula.

In the following two sections, stochastic approaches to index number theory will be considered that focus on the estimation of *price levels* rather than *bilateral price indexes*. In section 4, the price level approach will be applied to the case of two time periods while in section 5, the price level approach will be applied to many periods.

4. The Time Product Dummy Approach to Bilateral Index Number Theory

---

[13] It is formula number 9 in Fisher's (1922; 466) listing of index number formulae.

[14] For a proof of this assertion, see Diewert (2000).

In this section, a stochastic model that estimates average price levels for two periods will be derived using an adaptation of the Country Product Dummy model.[15] The adaptation is to move from the context of comparing prices across two countries to the time series context where the comparison of prices is made between two time periods.

Consider the following model of price behavior for the value aggregate under consideration:

(18) $p_n^t = \pi_t\alpha_n e_{tn}$; $\quad t = 0,1; n = 1,...,N.$

The parameter π_t can be interpreted as the *time product dummy price level* for period t, α_n can be interpreted as a commodity n *quality adjustment factor* [16] and e_{tn} is a positive stochastic error term with a mean that is assumed to be 1. Define the logarithms of p_n^t and e_{tn} as $y_{tn} \equiv \ln p_n^t$ and $\varepsilon_{tn} \equiv \ln e_{tn}$ for t = 0,1; n = 1,...,N, define the logarithm of π_t as $\rho_t \equiv \ln\pi_t$ for t = 0,1 and define the logarithm of α_n as $\beta_n \equiv \ln\alpha_n$ for n = 1,...,N. Then taking logarithms of both sides of (18) leads to the following *linear regression model*:

(19) $y_{tn} = \rho_t + \beta_n + \varepsilon_{tn}$; $\quad t = 0,1; n = 1,...,N.$

It can be seen that the parameters in the linear regression model defined by (19), the ρ_t and the β_n, are not uniquely determined. If any number λ is added to each ρ_t and the same number λ is subtracted from each β_n, the right hand side of each equation in (19) will not change. Thus in order to obtain unique estimates for the ρ_t and β_n on the right hand side of equations (19), we need to impose a normalization on these parameters. Impose the following normalization:

(20) $\rho_0 \equiv 0$.

This normalization corresponds to setting the period 0 price level, $\pi_0 \equiv \exp[\rho_0]$, equal to 1. Thus $\pi_1/\pi_0 = \pi_1$ and thus the estimated $\pi_1^* \equiv \exp[\rho_1^*]$ can be interpreted as a *bilateral index number* where ρ_1^* and $\beta_1^*,...,\beta_N^*$ solve the following *least squares minimization problem*:

(21) $\min_{\rho_1,\beta_1,...,\beta_N} \ \Sigma_{n=1}^N (y_{0n} - 0 - \beta_n)^2 + \Sigma_{n=1}^N (y_{1n} - \rho_1 - \beta_n)^2$.

The first order necessary (and sufficient) conditions for solving (21) are equation (22) and the N equations (23) listed below:

(22) $N\rho_1 + \Sigma_{n=1}^N \beta_n = \Sigma_{n=1}^N y_{1n}$;
(23) $\rho_1 + 2\beta_n = y_{0n} + y_{1n}$; $\quad n = 1,...,N.$

The solution to equations (22) and (23) is given by the following estimators:

(24) $\rho_1^* = (1/N)\Sigma_{n=1}^N [y_{1n} - y_{0n}]$;

[15] See Summers (1973) who introduced the CPD model. Balk (1980) was the first to adapt the CPD method to the time series context.
[16] In the context of commodities that are close substitutes, the interpretation of the α_n as quality adjustment factors is intuitively plausible. In the context of commodities that are not close substitutes, the α_n can be interpreted as *relative utility valuation factors*; i.e., α_n represents the marginal utility value to purchasers of the product of an extra unit of q_n. This interpretation relies on the economic approach to index number theory and is pursued in more depth in Chapter 8.

(25) $\beta_n^* = (½)y_{0n} + (½)[y_{1n} - \rho_1^*]$; n = 1,...,N.

Exponentiating the estimators defined by (24) and (25) leads to the following estimators for the period 1 price level (and price index) $\pi_1^* \equiv \exp[\rho_1^*]$ and the quality adjustment factors $\alpha_n^* \equiv \exp[\beta_n^*]$:

(26) $\pi_1^* \equiv \prod_{n=1}^N (p_n^1/p_n^0)^{1/N} = P_J(p^0,p^1)$;
(27) $\alpha_n^* \equiv (p_n^0)^{1/2}(p_n^1/\pi_1^*)^{1/2}$; n = 1,...,N

where $P_J(p^0,p^1)$ is the Jevons price index defined earlier by (4). This is Summer's (1973) country product dummy multilateral method adapted to the time series context for the case of two time periods with no missing observations.

The model defined by (18) or (19) can be interpreted as a highly simplified *hedonic regression model* [17] where the α_n are interpreted as quality adjustment factors for each product n. The only characteristic of each commodity is the commodity itself. As noted above, this model is also a special case of the Country Product Dummy method for making international comparisons between the prices of different countries. A possible advantage of this regression method for constructing a price index is that a standard error for the period 1 log price level ρ_1 (and hence for π_1) can be obtained. This advantage of the stochastic approach to index number theory was stressed by Selvanathan and Rao (1994). However, suppose that the standard error (or variance) for the estimated π_1^* were zero. Then all of the error terms e_{tn} in equations (18) must be equal to 1 and under these conditions, with $\pi_0 \equiv 1$, equations (18) imply that $p^1 = \pi_1^* p^0$ so that prices move in a *proportional manner* going from period 0 to period 1. Thus a nonzero standard error simply means that prices did *not* move in a proportional manner going from period 0 to 1. This fact does not imply that a larger standard error for π_1^* means that the overall inflation rate for the commodity group is more uncertain. For example, if the quantity vector q for periods 0 and 1 were constant, then most economists would agree that the appropriate measure of overall purchaser inflation is exactly measured by the Lowe index, $p^1 \cdot q/p^0 \cdot q$. Prices need not move in a proportional manner under these conditions so the standard error for π_1^* could be large but yet a very precise exact measure of overall inflation is available. Thus it must be kept in mind that standard errors for price levels or price indexes that are generated by a stochastic approach to index number theory are measures of parameter dispersion that are *conditional on the underlying model of price formation*. If the underlying model is faulty and the error variance is high, then the parameter standard errors that are generated by the model should be viewed with some degree of caution.

5. The Weighted Time Product Dummy Approach to Bilateral Index Number Theory

There is a problem with the unweighted least squares model defined by (21): namely that the logarithm of each price quote is given exactly the *same weight* in the model no matter what the expenditure on that item was in each period. This is obviously unsatisfactory since a price that has very little economic importance (i.e., a low expenditure share in each period) is given the same weight in the regression model compared to a very important item. As was mentioned above, Walsh was the first serious index number economist to stress the importance of weighting.

[17] For an introduction to hedonic regression models, see Griliches (1971). Hedonic regression models will be studied in great detail in Chapter 8 below.

Keynes was quick to follow up on the importance of weighting[18] and Fisher emphatically endorsed weighting.[19] Griliches also endorsed weighting in the hedonic regression context.[20] Thus it is useful to consider how to introduce weights into the time product dummy model that reflect the economic importance of the various commodities into the model.

In order to take economic importance into account, replace (21) by the following *weighted least squares minimization problem*:[21]

$$(28)\ \min_{\rho_1,\beta_1,\ldots,\beta_N} \Sigma_{n=1}^{N} q_n^0[\ln p_n^0 - \beta_n]^2 + \Sigma_{n=1}^{N} q_n^1[\ln p_n^1 - \rho_1 - \beta_n]^2$$

where we have set $\rho_0 = 0$. The squared error for product n in period t is repeated q_n^t times to reflect the sales of product n in period t. Thus the new problem (28) takes into account the popularity of each product.[22]

The first order necessary conditions for the minimization problem defined by (28) are the following N + 1 equations:

$$(29)\ (q_n^0+q_n^1)\beta_n = q_n^0\ln p_n^0 + q_n^1(\ln p_n^1 - \rho_1)\ ; \qquad n = 1,\ldots,N;$$
$$(30)\ (\Sigma_{n=1}^{N} q_n^1)\rho_1 = \Sigma_{n=1}^{N} q_n^1(\ln p_n^1 - \beta_n).$$

The solution to (29) and (30) is the following one:[23]

$$(31)\ \rho_1^* \equiv \Sigma_{n=1}^{N} q_n^0 q_n^1(q_n^0 + q_n^1)^{-1} \ln(p_n^1/p_n^0)/\Sigma_{i=1}^{N} q_i^0 q_i^1(q_i^0 + q_i^1)^{-1};$$
$$(32)\ \beta_n^* \equiv q_n^0(q_n^0 + q_n^1)^{-1} \ln(p_n^0) + q_n^1(q_n^0 + q_n^1)^{-1} \ln(p_n^1/\pi_1^*)\ ; \qquad n = 1,\ldots,N$$

where $\pi_1^* \equiv \exp[\rho_1^*]$. Note that the weight for the term $\ln(p_n^1/p_n^0)$ in (31) can be written as follows:

[18] "It is also clear that the so-called unweighted index numbers, usually employed by practical statisticians, are the worst of all and are liable to large errors which could have been easily avoided." J.M. Keynes (1909; 79). This paper won the Cambridge University Adam Smith Prize for that year. Keynes (1930; 76-77) again stressed the importance of weighting in his later 1930 paper which drew heavily on his 1909 paper.

[19] "It has already been observed that the purpose of any index number is to strike a fair average of the price movements or movements of other groups of magnitudes. At first a simple average seemed fair, just because it treated all terms alike. And, in the absence of any knowledge of the relative importance of the various commodities included in the average, the simple average is fair. But it was early recognized that there are enormous differences in importance. Everyone knows that pork is more important than coffee and wheat than quinine. Thus the quest for fairness led to the introduction of weighting." Irving Fisher (1922; 43).

[20] "But even here, we should use a weighted regression approach, since we are interested in an estimate of a weighted average of the pure price change, rather than just an unweighted average over all possible models, no matter how peculiar or rare." Zvi Griliches (1971; 8).

[21] Balk (1980; 70) was the first to both apply the country product dummy model to the time series context and he was the first to introduce some form of weighting to the basic model. However, the specific forms of weighting used in this section were introduced by Diewert (2005) for the models defined by (28), (35) and (42). Rao (1995) (2005) introduced the form of weighting for the model defined by (38).

[22] One can think of repeating the term $[\ln p_n^0 - \beta_n]^2$ for each unit of product n sold in period 0. The result is the term $q_n^0[\ln p_n^0 - \beta_n]^2$. A similar justification based on repeating the price according to its sales can also be made. This repetition methodology makes the stochastic specification of the error terms somewhat complicated but the least squares minimization problem is simple enough.

[23] This solution was derived by Diewert (2005).

(33) $q_n^* \equiv \Sigma_{n=1}^N q_n^0 q_n^1 (q_n^0 + q_n^1)^{-1} / \Sigma_{i=1}^N q_i^0 q_i^1 (q_i^0 + q_i^1)^{-1}$; $n = 1,...,N$
$= h(q_n^0,q_n^1)/\Sigma_{i=1}^N h(q_i^0,q_i^1)$

where $h(a,b) \equiv 2ab/(a+b) = [\tfrac{1}{2} a^{-1} + \tfrac{1}{2} b^{-1}]^{-1}$ is the *harmonic mean* of a and b.[24]

Note that the q_n^* sum to 1 and thus ρ_1^* is a weighted average of the logarithmic price ratios $\ln(p_n^1/p_n^0)$. Using $\pi_1^* = \exp[\rho_1^*]$ and $\pi_0^* = \exp[\rho_0^*] = \exp[0] = 1$, the bilateral price index that is generated by the solution to (28) is

(34) $\pi_1^*/\pi_0^* = \exp[\rho_1^*] = \exp[\Sigma_{n=1}^N q_n^* \ln(p_n^1/p_n^0)]$.

Thus π_1^*/π_0^* is a weighted geometric mean of the price ratios p_n^1/p_n^0 with weights q_n^* defined by (33). Although this seems to be a reasonable bilateral index number formula, it must be rejected for practical use on the grounds that *the index is not invariant to changes in the units of measurement*.

Since values are invariant to changes in the units of measurement, the lack of invariance problem could be solved if we replace the quantity weights in (28) with expenditure or sales weights.[25] This leads to the following *weighted least squares minimization problem* where the weights v_n^t are defined as $p_n^t q_n^t$ for t = 0,1 and n = 1,...,N:

(35) $\min_{\rho_1,\beta_1,...,\beta_N} \Sigma_{n=1}^N v_n^0[\ln p_n^0 - \beta_n]^2 + \Sigma_{n=1}^N v_n^1[\ln p_n^1 - \rho_1 - \beta_n]^2$.

It can be seen that problem (35) has exactly the same mathematical form as problem (28) except that v_n^t has replaced q_n^t and so the solutions (31) and (32) will be valid in the present context if v_n^t replaces q_n^t in these formulae. Thus the solution to (35) is:

(36) $\rho_1^* \equiv \Sigma_{n=1}^N v_n^0 v_n^1 (v_n^0 + v_n^1)^{-1} \ln(p_n^1/p_n^0)/\Sigma_{i=1}^N v_i^0 v_i^1 (v_i^0 + v_i^1)^{-1}$;
(37) $\beta_n^* \equiv v_n^0 (v_n^0 + v_n^1)^{-1} \ln(p_n^0) + v_n^1 (v_n^0 + v_n^1)^{-1} \ln(p_n^1/\pi_1^*)$; $n = 1,...,N$

where $\pi_1^* \equiv \exp[\rho_1^*]$.

The resulting price index, $\pi_1^*/\pi_0^* = \pi_1^* = \exp[\rho_1^*]$ is indeed invariant to changes in the units of measurement. However, if we regard π_1^* as a function of the price and quantity vectors for the two periods, say $P(p^0,p^1,q^0,q^1)$, then another problem emerges for the price index defined by the solution to (35): $P(p^0,p^1,q^0,q^1)$ is not homogeneous of degree 0 in the components of q^0 or in the components of q^1. These properties are important because it is desirable that the companion implicit quantity index defined as $Q(p^0,p^1,q^0,q^1) \equiv [p^1 \cdot q^1/p^0 \cdot q^0]/P(p^0,p^1,q^0,q^1)$ be homogeneous of

[24] h(a,b) is well defined by 2ab/(a+b) if a and b are nonnegative and at least one of these numbers is positive. In order to write h(a,b) as $[\tfrac{1}{2} a^{-1} + \tfrac{1}{2} b^{-1}]^{-1}$, we require that a > 0 and b > 0.
[25] "But on what principle shall we weight the terms? Arthur Young's guess and other guesses at weighting represent, consciously or unconsciously, the idea that relative money values of the various commodities should determine their weights. A value is, of course, the product of a price per unit, multiplied by the number of units taken. Such values afford the only common measure for comparing the streams of commodities produced, exchanged, or consumed, and afford almost the only basis of weighting which has ever been seriously proposed." Irving Fisher (1922; 45).

degree 1 in the components of q^1 and homogeneous of degree minus 1 in the components of q^0.[26] We also want $P(p^0,p^1,q^0,q^1)$ to be homogeneous of degree 1 in the components of p^1 and homogeneous of degree minus 1 in the components of p^0 and these properties are also not satisfied. Thus we conclude that the solution to the weighted least squares problem defined by (35) does not generate a satisfactory price index formula.

The above deficiencies can be remedied if the *expenditure amounts* v_n^t in (35) are replaced by *expenditure shares*, s_n^t, where $v^t \equiv \Sigma_{n=1}^N v_n^t$ for t = 0,1 and $s_n^t \equiv v_n^t/v^t$ for t = 0,1 and n = 1,...,N. This replacement leads to the following *weighted least squares minimization problem*:[27]

$$(38)\ \min_{\rho_1,\beta_1,...,\beta_N} \Sigma_{n=1}^N s_n^0[\ln p_n^0 - \beta_n]^2 + \Sigma_{n=1}^N s_n^1[\ln p_n^1 - \rho_1 - \beta_n]^2.$$

Again, it can be seen that problem (38) has exactly the same mathematical form as problem (28) except that s_n^t has replaced q_n^t and so the solutions (31) and (32) will be valid in the present context if s_n^t replaces q_n^t in these formulae. Thus the solution to (38) is:

$$(39)\ \rho_1^* \equiv \Sigma_{n=1}^N s_n^0 s_n^1(s_n^0 + s_n^1)^{-1} \ln(p_n^1/p_n^0)/\Sigma_{i=1}^N s_i^0 s_i^1(s_i^0 + s_i^1)^{-1};$$
$$(40)\ \beta_n^* \equiv s_n^0(s_n^0 + s_n^1)^{-1} \ln(p_n^0) + s_n^1(s_n^0 + s_n^1)^{-1} \ln(p_n^1/\pi_1^*)\ ; \qquad n = 1,...,N$$

where $\pi_1^* \equiv \exp[\rho_1^*]$. Define the *normalized harmonic mean share weights* as $s_n^* \equiv h(s_n^0,s_n^1)/\Sigma_{i=1}^N h(s_i^0,s_i^1)$ for n = 1,...,N. Then the *weighted time product dummy bilateral price index*, $P_{WTPD}(p^0,p^1,q^0,q^1) \equiv \pi_1^*/\pi_0^* = \pi_1^*$, has the following logarithm:

$$(41)\ \ln P_{WTPD}(p^0,p^1,q^0,q^1) \equiv \Sigma_{n=1}^N s_n^* \ln(p_n^1/p_n^0).$$

Thus $P_{WTPD}(p^0,p^1,q^0,q^1)$ is equal to a share weighted geometric mean of the price ratios, p_n^1/p_n^0.[28] This index is a satisfactory one from the viewpoint of the test approach to index number theory. Of the first 16 tests listed in Chapter 3, it can be shown that $P_{WTPD}(p^0,p^1,q^0,q^1)$ satisfies all of these tests (assuming strictly positive prices and quantities) except for Test 4 (the basket test, $P(p^0,p^1,q,q) = p^1\cdot q/p^0\cdot q$), Test 12 (the quantity reversal test), Test 13 (the price reversal test), Test 15 (the mean value test for quantities) and Test 16 (the Paasche and Laspeyres bounding test). It is likely that $P_{WTPD}(p^0,p^1,q^0,q^1)$ passes the monotonicity in prices tests, T17 and T18 and it is not likely that $P_{WTPD}(p^0,p^1,q^0,q^1)$ passes the monotonicity in quantity tests, T19 and T20.[29] Moreover, Diewert (2005; 564) showed that $P_{WTPD}(p^1,p^2,q^1,q^2)$ approximated the Fisher, Walsh and

---

[26] Thus we want Q to have the following properties: $Q(p^0,p^1,q^0,\lambda q^1) = \lambda Q(p^0,p^1,q^0,q^1)$ and $Q(p^0,p^1,\lambda q^0,q^1) = \lambda^{-1}Q(p^0,p^1,q^0,q^1)$ for all $\lambda > 0$; see Chapter 3 for a list of desirable properties or tests for bilateral price indexes of the form $P(p^0,p^1,q^0,q^1)$.

[27] Note that the minimization problem defined by (38) is equivalent to the problem of minimizing $\Sigma_{n=1}^N e_{0n}^2 + \Sigma_{n=1}^N e_{1n}^2$ with respect to $\rho_1, \beta_1, ..., \beta_N$ where the error terms e_{tn} are defined by the equations $(s_n^0)^{1/2}\ln p_n^0 = (s_n^0)^{1/2}\beta_n + e_{1n}$ for n = 1,...,N and $(s_n^1)^{1/2}\ln p_n^1 = (s_n^1)^{1/2}\rho_1 + (s_n^1)^{1/2}\beta_n + e_{2n}$ for n = 1,...,N. Thus the solution to (38) can be found by running a linear regression using the above two sets of estimating equations. The numerical equivalence of the least squares estimates obtained by repeating multiple observations or by using the square root of the weight transformation was noticed long ago as the following quotation indicates: "It is evident that an observation of weight w enters into the equations exactly as if it were w separate observations each of weight unity. The best practical method of accounting for the weight is, however, to prepare the equations of condition by multiplying each equation throughout by the square root of its weight." E. T. Whittaker and G. Robinson (1940; 224).

[28] See Diewert (2005) for this formula. Note that Rao (1995) (2005) considered the extension of the model defined by (38) to T periods and so he pioneered this class of models.

[29] See Diewert (2006) who initiated an investigation of the test properties of hedonic regressions.

Törnqvist-Theil indexes to the second order around an equal price and quantity point where $p^0 = p^1$ and $q^0 = q^1$. Thus if changes in prices and quantities going from one period to the next are not too large and there are no missing products, P_{WTPD} should be close to these indexes that have emerged as being "best" in several contexts.[30]

Recall the unweighted least squares minimization problem defined by (21) that was introduced at the beginning of section 4. The solution to this unweighted bilateral time product dummy regression model generated the Jevons index as its solution. But appropriate weighting of the squared errors has changed the solution index dramatically: the index defined by (41) weights products by their economic importance and has good test properties whereas the Jevons index can generate very problematic results because of its lack of weighting according to economic importance. Note that both models have the same underlying structure; i.e., they assume that p_{tn} is approximately equal to $\pi_t\alpha_n$ for t = 0,1 and n = 1,...,N. *Thus weighting by economic importance has converted a least squares minimization problem that generates a rather poor price index into a problem that generates a rather good index.*

There is one more weighting scheme that generates an even better index in the bilateral context where we are running a time product dummy hedonic regression using the price and quantity data for only two periods. Consider the following weighted least squares minimization problem:

$$(42)\ \min_{\rho_1,\beta_1,\ldots,\beta_N} \Sigma_{n=1}^N (½)(s_n^0+s_n^1)[\ln p_n^0 - \beta_n]^2 + \Sigma_{n=1}^N (½)(s_n^0+s_n^1)[\ln p_n^1 - \rho_1 - \beta_n]^2.$$

As usual, it can be seen that problem (42) has exactly the same mathematical form as problem (28) except that $(½)(s_n^0+s_n^1)$ has replaced q_n^0 and q_n^1 and so the solutions (31) and (32) will be valid in the present context if $(½)(s_{1n}+s_{2n})$ replaces q_n^t in these formulae. Thus the solutions to (42) simplify to the following solutions:

$$(43)\ \rho_1^* \equiv \Sigma_{n=1}^N (½)(s_n^0+s_n^1)\ln(p_n^1/p_n^0);$$
$$(44)\ \beta_n^* \equiv (½)\ln(p_n^0) + (½)\ln(p_n^1/\pi_1^*)\ ; \qquad n = 1,\ldots,N$$

where $\pi_1^* \equiv \exp[\rho_1^*]$ and $\pi_0 \equiv \exp[\rho_0] = \exp[0] = 1$ since we have set $\rho_0 = 0$. Thus the bilateral index number formula that emerges from the solution to (42) is $\pi_1^*/\pi_0 = \exp[\Sigma_{n=1}^N (½)(s_n^0+s_n^1)\ln(p_n^1/p_n^0)] \equiv P_T(p^0,p^1,q^0,q^1)$, which is the Törnqvist (1936) Theil (1967; 137-138) bilateral index number formula. Thus the use of the weights in (42) has generated an even better bilateral index number formula than the formula that resulted from the use of the weights in (38).[31] This result reinforces the case for using appropriately weighted versions of the basic time product dummy hedonic regression model.[32] However, if the implied residuals in the minimization problem (42) are small (or, equivalently, if the fit in the linear regression model that

[30] However, with large changes in price and quantities going from period 0 to 1, P_{WTPD} will tend to lie below these alternative indexes. Consider a case with only 2 commodities. Let the price vectors be $p^0 \equiv [1,1]$ and $p^1 \equiv [1,0.5]$ and let the share vectors be $s^0 \equiv [0.5,0.5]$ and $s^1 \equiv [0.1,0.9]$. Thus the two products are highly substitutable and when the price of product 2 goes on sale at half price, its market share jumps from 0.5 to 0.9. The Törnqvist Theil index for this example is 0.6156 which is considerably above the Weighted Time Product Dummy index value which is 0.5767. This example is based on an example due to Diewert and Fox (2017). Missing prices can also cause substantial differences between these indexes.

[31] Diewert (1992; 223) and Balk (2008) listed the commonly used tests that $P_T(p^0,p^1,q^0,q^1)$ satisfies; see also Chapter 3.

[32] Note that the bilateral regression model defined by the minimization problem (38) is readily generalized to the case of T periods whereas the bilateral regression model defined by the minimization problem (42) cannot be generalized to the case of T periods. These facts were noted by de Haan and Krsinich (2014).

can be associated with (42) is high, so that predicted values for log prices are close to actual log prices), then *weighting will not matter very much* and the unweighted version of (42), which was (21) in the previous section, will give results that are similar to (42). This comment applies to all of the weighted hedonic regression models that are considered in this section.[33] But in most practical applications of index number theory, *prices will not move in a proportional manner* over time. Under these conditions, weighting according to the economic importance of the individual commodities will lead to more representative estimates of overall price change; i.e., the measures of price change generated by the models defined by the minimization problems (38) and (42) are to be preferred over the unweighted minimization problem defined by (21) in the previous section.

In Chapter 7, generalizations of the bilateral weighted time product dummy model defined by the weighted least squares minimization problem (38) will be generalized from 2 periods to T periods. The problems arising from missing prices will also be addressed in this chapter.

At this point, it is perhaps useful to contrast stochastic approaches to index number theory to the approaches explained in Chapter 2 (basket approaches) and in Chapter 3 (axiomatic or test approaches). These earlier approaches led to a small number of preferred indexes such as the Fisher and Walsh indexes. The *stochastic approach* or the *descriptive statistics approach* to index number theory attempts to find a single summary measure of a distribution of price changes. The practical problem associated with this method is that there are many ways of summarizing relative price distributions as was seen at the end of section 3. We chose Theil's summary measure of price change because it satisfied some key tests; i.e., we had to draw on the test approach to index number theory in order to pin down our final specific estimator of price change. Similarly, in this section, we again drew on the test approach to index number theory to determine "best" measures of price change. This is the problem with the stochastic approach to index number theory: by itself, it does not narrow down the range of possible estimators of price change. Nevertheless, in subsequent chapters, we will utilize the stochastic approach to index number theory in order to address the problems associated with measuring quality change. However, as was done in this chapter, we will draw on the other approaches to index number theory to help us to narrow down the range of possible stochastic specifications that could be used to measure quality change.

Additional material on stochastic approaches to index number theory and references to the literature can be found in Selvanathan and Rao (1994), de Haan (2004), Diewert (2004) (2010), Rao (2004), Clements, Izan and Selvanathan (2006), de Haan and Krsinich (2014) and Rao and Hajargasht (2016).

References

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

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

[33] If the residuals are small for (42), then prices will vary almost proportionally over time and all reasonable index number formulae will register price levels that are close to the estimated π_1^*; i.e., we will have $p^1 \approx \pi_1^* p^0$ and hence all reasonable bilateral index number formula will generate an estimate that is close to π_1^*.

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).

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

de Haan, J. (2004), "Hedonic Regressions: The Time Dummy Index as a Special Case of the Törnqvist Index", Paper presented at the Eighth Meeting of the International Working Group on Price Indices, 23-25th August, Helsinki: Statistics Finland.

de Haan, J. and F. Krsinich (2014), "Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes", *Journal of Business and Economic Statistics* 32:3, 341-358.

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

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

Diewert, W.E. (2000), "Notes on Producing an Annual Superlative Index Using Monthly Price Data", Discussion Paper 00-08, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1, 30 pp.

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. (2005), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", *The Review of Income and Wealth* 51:4, 561-571.

Diewert, W.E. (2006), "Adjacent Period Dummy Variable Hedonic Regressions and Bilateral Index Number Theory", *Annales d'économie et de statistique*, No. 79/80, 1-28.

Diewert, W.E. (2010), "On the Stochastic Approach to Index Numbers", pp. 235-262 in *Price and Productivity Measurement, Volume 6—Index Number Theory*, Chapter 11, W.E. Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Trafford Press.

Diewert, W.E. and K.J. Fox (2017), "Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data", Discussion Paper 17-02, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada, V6T 1L4.

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.

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

Greenlees, J. (1999), "Random Errors and Superlative Indexes", paper presented at the Annual Conference of the Western Economic Association in San Diego, California, July 8, 1999.

Griliches, Z. (1971), "Introduction: Hedonic Price Indexes Revisited", pp. 3-15 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Jensen, J.L.W.V. (1906), "Sur les fonctions convexes et les inégalités entre les valeurs moyennes", *Acta Mathematica* 30, 175-193.

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.

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. (1909), "The Method of Index Numbers with Special Reference to the Measurement of General Exchange Value", reprinted as pp. 49-156 in The Collected Writings of John Maynard Keynes (1983), Volume 11, D. Moggridge (ed.), Cambridge: Cambridge University Press.

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

Palgrave, R.H.I. (1886), "Currency and Standard of Value in England, France and India and the Rates of Exchange Between these Countries", *Memorandum submitted to the Royal Commission on Depression of Trade and Industry*, Third Report, Appendix B, pp. 312-390.

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 (2004), "The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power parities in the ICP", paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, June 30-July 3, 2004, Vancouver, Canada.

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

Rao, D.S. Prasada and D. Hajargasht (2016), "Stochastic Approach to Computation of Purchasing Power Parities in the International Comparison Program (ICP)", *Journal of Econometrics* 191:2, 414-425.

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

Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.

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.

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.

Whittaker, E. T. and G. Robinson (1940), *The Calculus of Observations*, Third Edition, London: Blackie & Sons.

Wynne, M. A. (1999), “Core Inflation: A Review of Some Conceptual Issues”, Research Department Working Paper 99-03, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX 75201-2272.

Young, A. (1812), *An Inquiry into the Progressive Value of Money in England as Marked by the Price of Agricultural Products*, London: McMillan.