

Notes on Price Measurement¹

July 22, 2014.

W. Erwin Diewert

University of British Columbia and the University of New South Wales

Email: erwin.diewert@ubc.caWebsite: <http://www.economics.ubc.ca/faculty-and-staff/w-erwin-diewert/>**List of Chapters for the Notes****Chapter 1: Early Approaches to Index Number Theory****Chapter 2: The Economic Approach to Index Number Theory****Chapter 3: An Introduction to Multilateral Index Number Theory****Chapter 4: Elementary Price Indexes****Chapter 5: Lowe Indexes and the Practical Construction of a CPI****Chapter 6: The Treatment of Seasonal Commodities****Chapter 7: Scanner Data and the Rolling Year GEKS Method for Constructing Indexes****Chapter 8: An Empirical Illustration of Index Construction using Israeli Data on Vegetables****Chapter 9: Recommendations****Chapter 1: Early Approaches to Index Number Theory****1. Introduction**

There are *three main purposes* for which it is desirable to measure the average rate of change in consumer prices going from a base period 0 to a comparison period 1 for a well defined household or group of households that pertains to a well specified value aggregate (which defines which commodity transactions are *in scope* or to be included in the value aggregate):

- As a *summary measure of the overall rate of price change* that the specified group faces over the two periods being compared for the value aggregate under consideration. Households in the specified group will generally be interested in this summary number as will governments and central bankers.
- As a *deflator* for the value aggregate under consideration. Deflating the value ratio by the price index gives us the rate of growth of the corresponding quantity aggregate and this rate of growth can often be given a welfare change interpretation. Taking the household sector as a whole, the System of National Accounts requires a deflator for household expenditures.
- As a *compensation or indexation measure*. Governments or private employers may want to index benefit or salary levels to ensure that the indexed entitlement

¹ These Notes were prepared for the Preconference Training Sessions on Price Measurement held on August 23-24, 2014, sponsored by the 33nd General Conference of the International Association for Research in Income and Wealth held in Rotterdam, Netherlands, August 24-30, 2014.

or salary level in the current period is “equivalent” in some sense to the base level of entitlement or salary. The relevant value aggregates for this purpose should be based on the expenditures of the recipients of entitlements or of the employees in the indexation group.

It can be seen that the first two uses are really complimentary aspects of the same problem, which is to decompose a ratio of value aggregates for two periods into a price change component times a quantity change component. However, constructing an index for the third purpose is more difficult, since a proportion of an entitlement or a salary can be saved rather than spent on consumer goods and services and so the determination of the relevant value aggregate is not so easy. Another significant problem is that alternative treatments of purchases of consumer durables can lead to very different entitlement indexes.

Restricting attention to uses 1 and 2 above, it can be seen that the index number problem in dealing with demographic groups (e.g., pensioners) or with a restricted commodity classification (e.g., expenditures are restricted to an agreed on universe of “essential” commodities) is more or less the same as dealing with the entire household sector but the practical problem facing index number producers is the lack of data (or more accurately, the higher costs of collecting detailed expenditure and price data for demographic groups of households).

In practice, the problems facing statistical agencies producing consumer price indexes are more complicated than indicated above. With respect to the first two uses listed above, the theoretical approaches to index number theory that will be explained in the first two Chapters of these Notes apply to situations where complete price and quantity information on the same N goods and services in the household aggregate is available for the two periods being compared. Unfortunately, this idealized situation does not apply to the real world for a number of reasons:

- The existence of *strongly seasonal commodities*; i.e., these are commodities such as Christmas trees and clothing items such as swim suits that are available for only certain months of the year. Thus the list of commodities that are comparable changes from month to month.
- The introduction of *new products* and the disappearance of older products that have been rendered *obsolete* by technical progress. Again, this means that the list of comparable commodities changes from month to month.
- *Product churn* and temporary shortages of stock. Many retailers rotate their choice of brands that they will stock on their shelves for various reasons. Once again, this leads to a lack of comparability of products in the aggregate across months.
- The *existence of durable goods*. Thus a particular household may purchase a consumer durable (such as a motor vehicle or a house) in the base period but not purchase it in the following period. However, the household will still enjoy the *services* of the previously purchased consumer durable in the comparison period.

All of the above problems lead to *a lack of matching of purchases* of products across the two periods being compared and this creates problems for all approaches to index number theory. We will attempt to deal with the above problems in these Notes.

2. Overview of the Notes

At the outset, it should be recognized that no consumer or producer price index will be conceptually perfect: data limitations and cost considerations will prevent the “perfect” index from being produced. However, it will be useful to introduce the various idealized types of index that have been suggested in the literature on index number theory over the past 200 years. These idealized indexes are called *target indexes*. There are four main approaches to the determination of the functional form for a target price index that compares the prices (and associated quantities) between two periods:

- Fixed basket and averages of fixed basket approaches;
- The test or axiomatic approach;
- The stochastic approach and
- The economic approach.

The first three approaches listed above will be covered in this chapter;² the economic approach is deferred until Chapter 2.³

Practical consumer and producer price indexes are constructed in two stages:

- A first stage at the lowest level of aggregation where price information is available but associated expenditure or quantity information is not available and
- A second stage of aggregation where expenditure information is available at a higher level of aggregation.

The aggregates that pertain to the first stage of aggregation are called *elementary aggregates*. Again, theories for “ideal” or “best” target indexes can be developed in this situation where price information is available but not quantity or expenditure information. The two approaches that have been developed in this context are:

- The test approach and
- The stochastic approach.

Thus the theories for the target index in the elementary aggregate context parallel the theories developed when both price and quantity information is available, except that the

² The first three approaches can be applied in the consumer or producer price index context. The economic approach is a bit different in the producer and consumer contexts and these differences will be explained in Chapter 2. However, in most of these Notes, we will concentrate on the consumer or household context.

³ The material in this section largely parallels the material on index number theory that is laid out in the *Consumer Price Index Manual*; see the ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004; 263-327). For brevity, in the future, we will refer to the *CPI Manual* as ILO (2004). A considerable part of the materials presented in these notes is drawn from an unpublished report, Diewert (2012).

fixed basket and economic approaches cannot be applied in the elementary index context. The remaining two approaches to the construction of elementary indexes (the test and stochastic approaches) will be discussed in Chapter 4 below.⁴

All of the above theories for target indexes apply to situations where only the prices of two months (or quarters in the case of Australia and New Zealand) are being compared and the comparison formulae do not depend on the prices and quantities of any other month. This is termed *bilateral index number theory* since only two situations are being compared. However, in comparing prices (or aggregate quantities or volumes) across countries, it is necessary to make comparisons over more than two countries. This leads us into *multilateral index number theory* and it will be considered Chapter 3. It turns out that multilateral index number theory also plays an important role when making time series comparisons of prices as we shall see in Chapters 5 and 7.

Unfortunately, the materials on bilateral and multilateral index number theory do not deal with all of the complications that are encountered when constructing a Consumer Price Index. When a practical Consumer Price Index is constructed, an *annual expenditure basket* pertaining to a past year is used at higher levels of aggregation and at the elementary level of aggregation; the prices of twelve consecutive months are compared with the corresponding prices in December or January. The resulting price index is known as a *Lowe index* in the literature and in Chapter 5, we will give an overview of this methodology.

Chapter 6 discusses the problems that are associated with the existence of *seasonal commodities*.⁵

The *Consumer Price Index Manual* recommended the use of maximum overlap superlative indexes⁶ in the context of producing useful month to month consumer price indexes with seasonal commodities. But the evidence in Feenstra and Shapiro (2003), Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011) shows that these maximum overlap indexes can be subject to a tremendous *chain drift problem*. Thus in Chapter 7, a new method for constructing month to month indexes that avoids the chain drift problem due to Balk (1981), Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011) will be discussed. This method is known as the *Rolling Year GEKS method*.⁷

In Chapter 7, a new method due to de Haan and Krsinich (2012) for constructing elementary indexes will also be explained: the *Rolling Year Time Product Dummy method* (RYTPD method). This is a stochastic approach to elementary indexes that is an adaptation of a method suggested by Summers (1973). It is similar in spirit to the Rolling

⁴ Most of the material in Chapter 4 is also presented in the ILO (2004; 355-371).

⁵ Much of the material in this section is also presented in the ILO (2004; 393-417). However, since this *Consumer Price Index Manual* material was written, some new evidence on seasonal indexes has become available due to Diewert, Finkel and Artsev (2009) and this new material is reviewed in Chapter 6.

⁶ Superlative indexes will be explained in Chapter 2.

⁷ If the window length is different from 12 or 13 months, then the method is known as the Rolling Window GEKS Method.

Year GEKS method, except that the Rolling Year TPD method uses only price information instead of both price and quantity information.

Chapter 8 illustrates the methods proposed in Chapters 6 and 7 using an Israeli data set on fresh vegetables. This Chapter shows that the various methods proposed in Chapters 6 and 7 to deal with seasonal commodities and the problem of chain drift can be implemented by statistical agencies, provided that they undertake a continuous household expenditure survey so that reasonably up to date monthly or quarterly expenditure data are available to construct families of CPIs.

Finally, Chapter 9 concludes with some recommendations for statistical agency practices based on the contents of the previous chapters.

In the following sections of this Chapter, we turn to the problem of choosing an explicit index number formula. In the following section, we will discuss three of the four main approaches that are in use today to justify various functional forms for the price index when complete price and quantity data for the value aggregate are available for a number of periods. As mentioned above, when the comparison of price changes is restricted to two periods, the price index that constructs an average measure of price change is called a *bilateral index number formula*. In the following sections of this Chapter, attention is restricted to two period comparisons but of course, if one period is held fixed (called the base period), then a bilateral price index formula can be used to make a sequence of price comparisons over a number of subsequent periods.

3. Setting the Stage

It will be useful to set the stage for the subsequent discussion of alternative approaches by defining more precisely what the index number problem is.

We specify two accounting periods, $t = 0, 1$ for which we have micro price and quantity data for N commodities pertaining to transactions by a consumer (or a well defined group of consumers). Denote the price and quantity of commodity n in period t by p_n^t and q_n^t respectively for $n = 1, 2, \dots, N$ and $t = 0, 1$. Before proceeding further, we need to discuss the exact meaning of the microeconomic prices and quantities if there are *multiple* transactions for say commodity n within period t . In this case, it is natural to interpret q_n^t as the *total* amount of commodity n transacted within period t . In order to conserve the value of transactions, it is necessary that p_n^t be defined as a *unit value*⁸; i.e., p_n^t must be equal to the value of transactions for commodity n during period t divided by the total quantity transacted, q_n^t . For $t = 0, 1$, define *the value of transactions in period t* as:

$$(1) V^t \equiv \sum_{n=1}^N p_n^t q_n^t \equiv p^t \cdot q^t$$

⁸ The early index number theorists Walsh (1901; 96), Fisher (1922; 318) and Davies (1924; 96) all suggested unit values as the prices that should be inserted into an index number formula. This advice is followed in the *Consumer Price Index Manual: Theory and Practice* with the proviso that the unit value be a narrowly defined one; see the ILO (2004; 356).

where $p^t \equiv (p_1^t, \dots, p_N^t)$ is the period t price vector, $q^t \equiv (q_1^t, \dots, q_N^t)$ is the period t quantity vector and $p^t \cdot q^t$ denotes the inner product of these two vectors.

Using the above notation, we can now state the following *levels version of the index number problem using the test or axiomatic approach*: for $t = 0, 1$, find scalar numbers P^t and Q^t such that

$$(2) V^t = P^t Q^t.$$

The number P^t is interpreted as an aggregate period t price level while the number Q^t is interpreted as an aggregate period t quantity level. The aggregate price level P^t is allowed to be a function of the period t price vector, p^t while the aggregate period t quantity level Q^t is allowed to be a function of the period t quantity vector, q^t ; i.e., we have

$$(3) P^t = c(p^t) \text{ and } Q^t = f(q^t) ; t = 0, 1.$$

However, from the viewpoint of the *test approach* to index number theory, the levels approach to finding aggregate quantities and prices comes to an abrupt halt: Eichhorn (1978; 144) showed that if the number of commodities N in the aggregate is equal to or greater than 2 and we restrict $c(p^t)$ and $f(q^t)$ to be positive if the micro prices and quantities p_n^t and q_n^t are positive, then there do not exist any functions c and f such that $c(p^t)f(q^t) = p^t \cdot q^t$ for all $p^t \gg 0_N$ and $q^t \gg 0_N$.⁹

This negative result can be reversed if we take the *economic approach* to index number theory. In this approach, we assume that the economic agent has a linearly homogeneous utility function, $f(q)$, and when facing the prices p^t chooses q^t to solve the following cost minimization problem:

$$(4) \min_q \{ p^t \cdot q : f(q) \geq u^t = f(q^t) \} = p^t \cdot q^t \equiv Y^t ; \quad t = 0, 1;$$

where period t “income” Y^t is defined as $p^t \cdot q^t$. In this setup, it turns out that $c(p)$ is the unit cost function that is dual¹⁰ to the linearly homogeneous utility function $f(q)$ and we can define P^t and Q^t as in (3) with $P^t Q^t = c(p^t)f(q^t) = p^t \cdot q^t$ for $t = 0, 1$. Why does the economic approach work in the levels version of the index number problem whereas the test approach does not? In the test approach, both p^t and q^t are regarded as completely independent variables, whereas in the economic approach, p^t can vary independently but q^t cannot vary independently; it is a solution to the period t cost minimization problem (4).

Even though the economic approach to the index number problem as formulated above “works”, it is not a *practical* solution that statistical agencies can implement and provide suitable aggregates to the public. In order to implement this solution, the statistical agency would have to hire hundreds of econometricians in order to estimate cost

⁹ Notation: $p \gg 0_N$ means all components of p are positive; $p \geq 0_N$ means all components of p are nonnegative and $p > 0_N$ means $p \geq 0_N$ but $p \neq 0_N$. Finally, $p \cdot q \equiv \sum_{n=1}^N p_n q_n$.

¹⁰ See Diewert (1974) (1993b) for materials and references to the literature on duality theory.

functions for all relevant macroeconomic aggregates and it is simply not feasible to do this. Thus we turn to our second formulation of the index number problem and it is this formulation that was initiated by Fisher (1911) (1922) in his two books on index number theory.

In the second approach to index number theory, instead of trying to decompose the value of the aggregate into price and quantity components for a single period, we instead attempt to decompose a *value ratio* for the two periods under consideration into a *price change component* P times a *quantity change component* Q .¹¹ Thus we now look for two functions of $4N$ variables, $P(p^0, p^1, q^0, q^1)$ and $Q(p^0, p^1, q^0, q^1)$ such that:¹²

$$(5) p^1 \cdot q^1 / p^0 \cdot q^0 = P(p^0, p^1, q^0, q^1) Q(p^0, p^1, q^0, q^1).$$

If we take the test approach, then we want equation (5) to hold for all positive price and quantity vectors pertaining to the two periods under consideration, p^0, p^1, q^0, q^1 . If we take the economic approach, then only the price vectors p^0 and p^1 are regarded as independent variables while the quantity vectors, q^0 and q^1 , are regarded as dependent variables.

In this second approach to index number theory, the *price index* $P(p^0, p^1, q^0, q^1)$ and the *quantity index* $Q(p^0, p^1, q^0, q^1)$ cannot be determined independently; i.e., if either one of these two functions is determined, then the remaining function is implicitly determined using equation (5). Historically, the focus has been on the determination of the price index but Fisher (1911; 388) was the first to realize that once the price index was determined, then equation (5) could be used to determine the companion quantity index.¹³ This value ratio decomposition approach to index number is called *bilateral index number theory* and its focus is the determination of “reasonable” functional forms for P and Q . Fisher’s 1911 and 1922 books address this functional form issue using the test approach.

We turn now to a discussion of the various approaches that have been used to determine the functional form for the bilateral price index, $P(p^0, p^1, q^0, q^1)$.

4. Fixed Basket Approaches to Bilateral Index Number Theory

A very simple approach to the determination of a price index over a group of commodities is the *fixed basket approach*. In this approach, we are given a basket of commodities that is represented by the positive quantity vector q . Given the price vectors

¹¹ Looking ahead to the economic approach which will be explained in more detail in Chapter 2, P will be interpreted to be the ratio of unit cost functions, $c(p^1)/c(p^0)$, and Q will be interpreted to be the utility ratio, $f(q^1)/f(q^0)$. Note that the linear homogeneity assumption on the utility function f effectively cardinalizes utility.

¹² If $N = 1$, then we define $P(p_1^0, p_1^1, q_1^0, q_1^1) \equiv p_1^1/p_1^0$ and $Q(p_1^0, p_1^1, q_1^0, q_1^1) \equiv q_1^1/q_1^0$, the single price ratio and the single quantity ratio respectively. In the case of a general N , we think of $P(p_1^0, p_1^1, q_1^0, q_1^1)$ as being a weighted average of the price ratios $p_1^1/p_1^0, p_2^1/p_2^0, \dots, p_N^1/p_N^0$. Thus we interpret $P(p_1^0, p_1^1, q_1^0, q_1^1)$ as an aggregate price ratio, P^1/P^0 , where P^t is the aggregate price level for period t for $t = 0, 1$.

¹³ This approach to index number theory is due to Fisher (1911; 418) who called the implicitly determined Q , the *correlative formula*. Frisch (1930; 399) later called (5) the *product test*.

for periods 0 and 1, p^0 and p^1 respectively, we can calculate the cost of purchasing this same basket in the two periods, $p^0 \cdot q$ and $p^1 \cdot q$. Then the ratio of these costs is a very reasonable indicator of pure price change over the two periods under consideration, provided that the basket vector q is “representative”. Thus define the *Lowe (1823) price index*, P_{Lo} , as follows:

$$(6) P_{Lo}(p^0, p^1, q) \equiv p^1 \cdot q / p^0 \cdot q .$$

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 defined by (7) and the Paasche (1874) price index P_P defined by (8):¹⁴

$$(7) P_L(p^0, p^1, q^0, q^1) \equiv p^1 \cdot q^0 / p^0 \cdot q^0 = \sum_{n=1}^N s_n^0 (p_n^1 / p_n^0) ;$$

$$(8) P_P(p^0, p^1, q^0, q^1) \equiv p^1 \cdot q^1 / p^0 \cdot q^1 = [\sum_{n=1}^N s_n^1 (p_n^1 / p_n^0)^{-1}]^{-1}$$

where the period t expenditure share on commodity n , s_n^t , is defined as $p_n^t q_n^t / p^t \cdot q^t$ for $n = 1, \dots, N$ and $t = 0, 1$. Thus the Laspeyres price index P_L can be written as a base period expenditure share weighted average of the N price ratios (or price relatives), p_n^1 / p_n^0 .¹⁵ The last equation in (8) shows that the Paasche price index P_P can be written as a period 1 (or current period) expenditure share weighted *harmonic* average of the N price ratios.¹⁶

The problem with these 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 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 Drobisch (1871) Sidgwick (1883; 68) Bowley (1901; 227)¹⁷ index, $(1/2)P_L + (1/2)P_P$, and the geometric mean, which leads to the *Fisher (1922) ideal index*, P_F , defined as

$$(9) P_F(p^0, p^1, q^0, q^1) \equiv [P_L(p^0, p^1, q^0, q^1) P_P(p^0, p^1, q^0, q^1)]^{1/2} .$$

At this point, the fixed basket approach to index number theory has to draw on the *test approach* to index number theory; i.e., in order to determine which of these fixed basket

¹⁴ Note that $P_L(p^0, p^1, q^0, q^1)$ does not actually depend on q^1 and $P_P(p^0, p^1, q^0, q^1)$ does not actually depend on q^0 . However, it does no harm to include these vectors and the notation indicates that we are in the realm of bilateral index number theory. Note also that Drobisch (1871) proposed both the Laspeyres and Paasche indexes in passing but he dismissed them; Drobisch preferred a unit value index of prices (which is not useful when aggregating over commodities that are measured in heterogeneous units).

¹⁵ This result is due to Walsh (1901; 428 and 539).

¹⁶ This expenditure share and price ratio representation of the Paasche index is described by Walsh (1901; 428) and derived explicitly by Fisher (1911; 365).

¹⁷ See Diewert (1992) (1993a) and Balk (2008) for additional references to the early history of index number theory.

indexes or which averages of them might be “best”, we need *criteria* or *tests* or *properties* that we would like our indexes to satisfy.

What is the “best” symmetric average of P_L and P_P to use as a point estimate for the theoretical cost of living index? It is very desirable for a price index formula that depends on the price and quantity vectors pertaining to the two periods under consideration to satisfy the *time reversal test*.¹⁸ We say that the index number formula $P(p^0, p^1, q^0, q^1)$ satisfies this test if

$$(10) P(p^1, p^0, q^1, q^0) = 1/P(p^0, p^1, q^0, q^1) ;$$

i.e., if we interchange the period 0 and period 1 price and quantity data and evaluate the index, then this new index $P(p^1, p^0, q^1, q^0)$ is equal to the reciprocal of the original index $P(p^0, p^1, q^0, q^1)$.

Diewert (1997; 138) showed that the Fisher ideal price index defined by (9) above is the *only* index that is a homogeneous symmetric mean of the Laspeyres and Paasche price indexes, P_L and P_P , and satisfies the time reversal test (10) above. Thus our first *symmetric basket approach* to bilateral index number theory leads to the Fisher index (9) as being “best” from the perspective of this approach.¹⁹

Instead of looking for a “best” average of the two fixed basket indexes that correspond to the baskets chosen in either of the two periods being compared, we could instead look for a “best” average basket of the two baskets represented by the vectors q^0 and q^1 and then use this average basket to compare the price levels of periods 0 and 1.²⁰ Thus we ask that the n th quantity weight, q_n , be an average or *mean* of the base period quantity q_n^0 and the period 1 quantity for commodity n q_n^1 , say $m(q_n^0, q_n^1)$, for $n = 1, 2, \dots, N$.²¹ Price statisticians refer to this type of index as a *pure price index* and it corresponds to Knibbs’ (1924; 43) *unequivocal price index*. Under these assumptions, the pure price index can be defined as a member of the following class of index numbers:

$$(11) P_K(p^0, p^1, q^0, q^1) \equiv \sum_{n=1}^N p_n^1 m(q_n^0, q_n^1) / \sum_{j=1}^N p_j^0 m(q_j^0, q_j^1).$$

¹⁸ The concept of this test is due to Pierson (1896; 128), who was so upset with the fact that many of the commonly used index number formulae did not satisfy this test (and the commensurability test to be discussed later) that he proposed that the entire concept of an index number should be abandoned. More formal statements of the test were made by Walsh (1901; 324) and Fisher (1922; 64).

¹⁹ Bowley was an early advocate of taking a symmetric average of the Paasche and Laspeyres indexes: “If [the Paasche index] and [the Laspeyres index] lie close together there is no further difficulty; if they differ by much they may be regarded as inferior and superior limits of the index number, which may be estimated as their arithmetic mean ... as a first approximation.” Arthur L. Bowley (1901; 227). Fisher (1911; 418-419) (1922) considered taking the arithmetic, geometric and harmonic averages of the Paasche and Laspeyres indexes.

²⁰ Walsh (1901) (1921a) and Fisher (1922) considered both averaging strategies in their classic studies on index numbers.

²¹ Note that we have chosen the mean function $m(q_n^0, q_n^1)$ to be the same for each commodity n . Marshall (1887) and Edgeworth (1925) recommended that m be the arithmetic mean whereas Walsh (1901) (1921a) recommended the use of the geometric mean.

In order to determine the functional form for the mean function m , it is necessary to impose some *tests* or *axioms* on the pure price index defined by (11). Suppose that we impose the time reversal test (10) and the following *invariance to proportional changes in current quantities test*:

$$(12) P(p^0, p^1, q^0, \lambda q^1) = P(p^0, p^1, q^0, q^1) \text{ for all } \lambda > 0.$$

Diewert (2001; 207) showed that these two tests determine the precise functional form for the pure price index P_K defined by (11) above: the pure price index P_K must be the *Walsh* (1901; 398) (1921a; 97) *price index*, P_W^{22} defined by (13):

$$(13) P_W(p^0, p^1, q^0, q^1) \equiv \sum_{n=1}^N p_n^1 (q_n^0 q_n^1)^{1/2} / \sum_{j=1}^N p_j^0 (q_j^0 q_j^1)^{1/2}.$$

Thus the fixed basket approach to bilateral index number theory starts out with the Laspeyres and Paasche price indexes. Some form of averaging of these two indexes is called for since both indexes are equally plausible. Averaging these two indexes directly leads to the Fisher ideal index P_F defined by (9) as being “best” while a direct averaging of the two quantity baskets q^0 and q^1 leads to the Walsh price index P_W defined by (13) as being “best”.

We turn now to another early approach to the index number problem.

5. Stochastic and Descriptive Statistics Approaches to Index Number Theory

The (unweighted) stochastic approach to the determination of the price index can be traced back to the work of Jevons (1865) (1884) and Edgeworth (1888) (1896) (1901) over a hundred years ago²³.

The basic idea behind the stochastic approach is that each price relative, p_n^1/p_n^0 for $n = 1, 2, \dots, N$, can be regarded as an estimate of a common inflation rate α between periods 0 and 1; i.e., Jevons and Edgeworth essentially assumed that

$$(14) p_n^1/p_n^0 = \alpha + \varepsilon_n ; n = 1, 2, \dots, N$$

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* (1804) *price index* P_C defined as

²² Walsh endorsed P_W as being the best index number formula: “We have seen reason to believe formula 6 better than formula 7. Perhaps formula 9 is the best of the rest, but between it and Nos. 6 and 8 it would be difficult to decide with assurance.” C.M. Walsh (1921a; 103). His formula 6 is P_W defined by (13) and his 9 is the Fisher ideal defined by (9) above. His formula 8 is the formula $p^1 \cdot q^1 / p^0 \cdot q^0 Q_W(p^0, p^1, q^0, q^1)$, which is known as the implicit Walsh price index where $Q_W(p^0, p^1, q^0, q^1)$ is the Walsh quantity index defined by (13) except the role of prices and quantities is interchanged. Thus although Walsh thought that his Walsh price index was the best functional form, his implicit Walsh price index and the “Fisher” formula were not far behind.

²³ For additional references to the early literature, see Diewert (1993a; 37-38) (1995b) and Balk (2008; 32-36).

$$(15) P_C(p^0, p^1) \equiv \sum_{n=1}^N (1/N)(p_n^1/p_n^0).$$

Unfortunately, P_C does not satisfy the time reversal test, i.e., $P_C(p^1, p^0) \neq 1/P_C(p^0, p^1)$ ²⁴.

Now assume that the logarithm of each price relative, $\ln(p_n^1/p_n^0)$, is an independent unbiased estimate of the logarithm of the inflation rate between periods 0 and 1, β say. Thus we have:

$$(16) \ln(p_n^1/p_n^0) = \beta + \varepsilon_n ; n = 1, 2, \dots, N$$

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

$$(17) P_J(p^0, p^1) \equiv \prod_{n=1}^N (p_n^1/p_n^0)^{1/N}.$$

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 (15) and (17).²⁵ Keynes (1930; 76-81) also criticized the unweighted stochastic approach to index number theory on two other grounds: (i) price relatives are not distributed independently and (ii) there is no single inflation rate that can be applied to all parts of an economy; e.g., Keynes demonstrated empirically that wage rates, wholesale prices and final consumption prices all had different rates of inflation. In order to overcome the Keynesian criticisms of the unweighted stochastic approach to index numbers, it is necessary to:

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

Theil (1967; 136-137) proposed a solution to the lack of weighting in (15). 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 n th 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 . Then the overall mean (period 0 weighted)

²⁴ 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. Walsh (1901; 327) established this inequality for the case $N = 2$. Fisher urged users to abandon the use of the Carli index but his advice was generally ignored by statistical agencies until recently: "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose." Irving Fisher (1922; 29-30).

²⁵ Walsh (1901) (1921a; 82-83), Fisher (1922; 43) and Keynes (1930; 76-77) all objected to the lack of weighting in the unweighted stochastic approach to index number theory.

logarithmic 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 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 (1967; 137) argued that a nice symmetric index number formula can be obtained if we make the probability of selection for the n th 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 is

$$(18) \ln P_T(p^0, p^1, q^0, q^1) \equiv \sum_{n=1}^N (1/2)(s_n^0 + s_n^1) \ln(p_n^1/p_n^0).$$

It is possible to give a *descriptive statistics* interpretation of the right hand side of (18). Define the n th logarithmic price ratio r_n by:

$$(19) r_n \equiv \ln(p_n^1/p_n^0) \quad \text{for } n = 1, \dots, N.$$

Now define the discrete random variable, R say, as the random variable which can take on the values r_n with probabilities $\rho_n \equiv (1/2)(s_n^0 + s_n^1)$ for $n = 1, \dots, N$. Note that since each set of expenditure shares, s_n^0 and s_n^1 , sums 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 $\ln P_T(p^0, p^1, q^0, q^1)$ as defined by the right hand side of (18). 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, ρ_n .

Taking antilogs of both sides of (18), we obtain the Theil price index; P_T .²⁶ This index number formula has a number of good properties. In particular, P_T satisfies the time reversal test (10) and the linear homogeneity test (12).²⁷

Additional material on stochastic approaches to index number theory and references to the literature can be found in Rao (1990) (1999) (2001) (2002) (2004), Selvanathan and Rao (1994), Diewert (1995b) (2004), Wynne (1997), Clements, Izan and Selvanathan (2006) and Balk (2008; 32-36)

6. Test Approaches to Index Number Theory²⁸

²⁶ This index first appeared explicitly as formula 123 in Fisher (1922; 473). P_T is generally attributed to Törnqvist (1936) but this article did not have an explicit definition for P_T ; it was defined explicitly in Törnqvist and Törnqvist (1937); see Balk (2008; 26).

²⁷ For a listing of some of the tests that P_T , P_F , and P_W satisfy, see Diewert (1992; 223). In Fisher (1922), these indexes were listed as numbers 123, 353 and 1153 respectively.

²⁸ The material in this section is based on Diewert (1992) where more detailed references to the literature on the origins of the various tests can be found.

Recall equation (5) above, which set the value ratio, V^1/V^0 , equal to the product of the price index, $P(p^0, p^1, q^0, q^1)$, and the quantity index, $Q(p^0, p^1, q^0, q^1)$. This is called the Product Test and we assume that it is satisfied. This equation means that as soon as the functional form for the price index P is determined, then (5) can be used to determine the functional form for the quantity index Q . However, a further advantage of assuming that the product test holds is that we can assume that the quantity index Q satisfies a “reasonable” property and then use (5) to translate this test on the quantity index into a corresponding test on the price index P .²⁹

If $N = 1$, so that there is only one price and quantity to be aggregated, then a natural candidate for P is p_1^1/p_1^0 , the single price ratio, and a natural candidate for Q is q_1^1/q_1^0 , the single quantity ratio. When the number of commodities or items to be aggregated is greater than 1, then what index number theorists have done over the years is propose properties or tests that the price index P should satisfy. These properties are generally multi-dimensional analogues to the one good price index formula, p_1^1/p_1^0 . Below, we list twenty-one tests that turn out to characterize the Fisher ideal price index.

We shall assume that every component of each price and quantity vector is positive; i.e., $p^t >> 0_N$ and $q^t >> 0_N$ for $t = 0, 1$. If we want to set $q^0 = q^1$, we call the common quantity vector q ; if we want to set $p^0 = p^1$, we call the common price vector p .

Our first two tests are not very controversial and so we will not discuss them.

T1: *Positivity*: $P(p^0, p^1, q^0, q^1) > 0$.

T2: *Continuity*: $P(p^0, p^1, q^0, q^1)$ is a continuous function of its arguments.

Our next two tests are somewhat more controversial.

T3: *Identity or Constant Prices Test*: $P(p, p, q^0, q^1) = 1$.

That is, if the price of every good is identical during the two periods, then the price index should equal unity, no matter what the quantity vectors are. The controversial part of this test is that the two quantity vectors are allowed to be different in the above test.³⁰

T4: *Fixed Basket or Constant Quantities Test*: $P(p^0, p^1, q, q) = \sum_{i=1}^N p_i^1 q_i / \sum_{i=1}^N p_i^0 q_i$.

That is, if quantities are constant during the two periods so that $q^0 = q^1 \equiv q$, then the price index should equal the expenditure on the constant basket in period 1, $\sum_{i=1}^N p_i^1 q_i$, divided by the expenditure on the basket in period 0, $\sum_{i=1}^N p_i^0 q_i$.

²⁹ This observation was first made by Fisher (1911; 400-406). Vogt (1980) also pursued this idea.

³⁰ Usually, economists assume that given a price vector p , the corresponding quantity vector q is uniquely determined. Here, we have the same price vector but the corresponding quantity vectors are allowed to be different.

The following four tests are *homogeneity tests* and they restrict the behavior of the price index P as the scale of any one of the four vectors p^0, p^1, q^0, q^1 changes.

T5: *Proportionality in Current Prices*: $P(p^0, \lambda p^1, q^0, q^1) = \lambda P(p^0, p^1, q^0, q^1)$ for $\lambda > 0$.

That is, if all period 1 prices are multiplied by the positive number λ , then the new price index is λ times the old price index. Put another way, the price index function $P(p^0, p^1, q^0, q^1)$ is (positively) homogeneous of degree one in the components of the period 1 price vector p^1 . Most index number theorists regard this property as a very fundamental one that the index number formula should satisfy.

Walsh (1901) and Fisher (1911; 418) (1922; 420) proposed the related proportionality test $P(p, \lambda p, q^0, q^1) = \lambda$. This last test is a combination of T3 and T5; in fact Walsh (1901, 385) noted that this last test implies the identity test, T3.

In our next test, instead of multiplying all period 1 prices by the same number, we multiply all period 0 prices by the number λ .

T6: *Inverse Proportionality in Base Period Prices*: $P(\lambda p^0, p^1, q^0, q^1) = \lambda^{-1} P(p^0, p^1, q^0, q^1)$ for $\lambda > 0$.

That is, if all period 0 prices are multiplied by the positive number λ , then the new price index is $1/\lambda$ times the old price index. Put another way, the price index function $P(p^0, p^1, q^0, q^1)$ is (positively) homogeneous of degree minus one in the components of the period 0 price vector p^0 .

The following two homogeneity tests can also be regarded as invariance tests.

T7: *Invariance to Proportional Changes in Current Quantities*: $P(p^0, p^1, q^0, \lambda q^1) = P(p^0, p^1, q^0, q^1)$ for all $\lambda > 0$.

That is, if current period quantities are all multiplied by the number λ , then the price index remains unchanged. Put another way, the price index function $P(p^0, p^1, q^0, q^1)$ is (positively) homogeneous of degree zero in the components of the period 1 quantity vector q^1 . Vogt (1980, 70) was the first to propose this test and his derivation of the test is of some interest. Suppose the quantity index Q satisfies the quantity analogue to the price test T5; i.e., suppose Q satisfies $Q(p^0, p^1, q^0, \lambda q^1) = \lambda Q(p^0, p^1, q^0, q^1)$ for $\lambda > 0$. Then using the product test (5), we see that P must satisfy T7.

T8: *Invariance to Proportional Changes in Base Quantities*: $P(p^0, p^1, \lambda q^0, q^1) = P(p^0, p^1, q^0, q^1)$ for all $\lambda > 0$.

That is, if base period quantities are all multiplied by the number λ , then the price index remains unchanged. Put another way, the price index function $P(p^0, p^1, q^0, q^1)$ is (positively) homogeneous of degree zero in the components of the period 0 quantity

vector q^0 . If the quantity index Q satisfies the following counterpart to T8: $Q(p^0, p^1, \lambda q^0, q^1) = \lambda^{-1} Q(p^0, p^1, q^0, q^1)$ for all $\lambda > 0$, then using (5), the corresponding price index P must satisfy T8. This argument provides some additional justification for assuming the validity of T8 for the price index function P .

T7 and T8 together impose the property that the price index P does not depend on the *absolute* magnitudes of the quantity vectors q^0 and q^1 .

The next five tests are *invariance* or *symmetry tests*. Fisher (1922; 62-63, 458-460) and Walsh (1921b; 542) seem to have been the first researchers to appreciate the significance of these kinds of tests. Fisher (1922, 62-63) spoke of fairness but it is clear that he had symmetry properties in mind. It is perhaps unfortunate that he did not realize that there were more symmetry and invariance properties than the ones he proposed; if he had realized this, it is likely that he would have been able to provide an axiomatic characterization for his ideal price index. Our first invariance test is that the price index should remain unchanged if the *ordering* of the commodities is changed:

T9: *Commodity Reversal Test* (or invariance to changes in the ordering of commodities):

$$P(p^{0*}, p^{1*}, q^{0*}, q^{1*}) = P(p^0, p^1, q^0, q^1)$$

where p^{t*} denotes a permutation of the components of the vector p^t and q^{t*} denotes the same permutation of the components of q^t for $t = 0, 1$. This test is due to Irving Fisher (1922), and it is one of his three famous reversal tests. The other two are the time reversal test and the factor reversal test which will be considered below.

T10: *Invariance to Changes in the Units of Measurement* (commensurability test):

$$P(\alpha_1 p_1^0, \dots, \alpha_N p_N^0; \alpha_1 p_1^1, \dots, \alpha_N p_N^1; \alpha_1^{-1} q_1^0, \dots, \alpha_N^{-1} q_N^0; \alpha_1^{-1} q_1^1, \dots, \alpha_N^{-1} q_N^1) = P(p_1^0, \dots, p_N^0; p_1^1, \dots, p_N^1; q_1^0, \dots, q_N^0; q_1^1, \dots, q_N^1) \text{ for all } \alpha_1 > 0, \dots, \alpha_N > 0.$$

That is, the price index does not change if the units of measurement for each commodity are changed. The concept of this test was due to Jevons (1884; 23) and the Dutch economist Pierson (1896; 131), who criticized several index number formula for not satisfying this fundamental test. Fisher (1911; 411) first called this test *the change of units test* and later, Fisher (1922; 420) called it the *commensurability test*.

T11: *Time Reversal Test*: $P(p^0, p^1, q^0, q^1) = 1/P(p^1, p^0, q^1, q^0)$.

That is, if the data for periods 0 and 1 are interchanged, then the resulting price index should equal the reciprocal of the original price index. Obviously, in the one good case when the price index is simply the single price ratio; this test is satisfied (as are all of the other tests listed in this section). When the number of goods is greater than one, many commonly used price indexes fail this test; e.g., the Laspeyres (1871) price index, P_L defined earlier by (7), and the Paasche (1874) price index, P_P defined earlier by (8), both *fail* this fundamental test. The concept of the test was due to Pierson (1896; 128), who was so upset with the fact that many of the commonly used index number formulae did not satisfy this test, that he proposed that the entire concept of an index number should be

abandoned. More formal statements of the test were made by Walsh (1901; 368) (1921b; 541) and Fisher (1911; 534) (1922; 64).

Our next two tests are more controversial, since they are not necessarily consistent with the economic approach to index number theory. However, these tests are quite consistent with the weighted stochastic approach to index number theory discussed earlier in section 3.3.

T12: *Quantity Reversal Test* (quantity weights symmetry test):

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

That is, if the quantity vectors for the two periods are interchanged, then the price index remains invariant. This property means that if quantities are used to weight the prices in the index number formula, then the period 0 quantities q^0 and the period 1 quantities q^1 must enter the formula in a symmetric or even handed manner. Funke and Voeller (1978; 3) introduced this test; they called it the *weight property*.

The next test is the analogue to T12 applied to quantity indexes:

T13: *Price Reversal Test* (price weights symmetry test):

$$\{\sum_{i=1}^N p_i^1 q_i^1 / \sum_{i=1}^N p_i^0 q_i^0\} / P(p^0, p^1, q^0, q^1) = \{\sum_{i=1}^N p_i^0 q_i^1 / \sum_{i=1}^N p_i^1 q_i^0\} / P(p^1, p^0, q^0, q^1).$$

Thus if we use (5) to define the quantity index Q in terms of the price index P , then it can be seen that T13 is equivalent to the following property for the associated quantity index Q : $Q(p^0, p^1, q^0, q^1) = Q(p^1, p^0, q^0, q^1)$. That is, if the price vectors for the two periods are interchanged, then the quantity index remains invariant. Thus if prices for the same good in the two periods are used to weight quantities in the construction of the quantity index, then property T13 implies that these prices enter the quantity index in a symmetric manner.

The next three tests are mean value tests.

T14: *Mean Value Test for Prices*:

$$\min_i \{p_i^1 / p_i^0 : i=1, \dots, N\} \leq P(p^0, p^1, q^0, q^1) \leq \max_i \{p_i^1 / p_i^0 : i=1, \dots, N\}.$$

That is, the price index lies between the minimum price ratio and the maximum price ratio. Since the price index is supposed to be some sort of an average of the N price ratios, p_i^1 / p_i^0 , it seems essential that the price index P satisfy this test.

The next test is the analogue to T14 applied to quantity indexes:

T15: *Mean Value Test for Quantities*:

$$\min_i \{q_i^1 / q_i^0 : i=1, \dots, N\} \leq \{V^1 / V^0\} / P(p^0, p^1, q^0, q^1) \leq \max_i \{q_i^1 / q_i^0 : i=1, \dots, N\}$$

where V^t is the period t value aggregate $V^t \equiv \sum_{n=1}^N p_n^t q_n^t$ for $t = 0, 1$. Using (5) to define the quantity index Q in terms of the price index P , we see that T15 is equivalent to the following property for the associated quantity index Q :

$$(20) \min_i \{q_i^1/q_i^0 : i=1, \dots, N\} \leq Q(p^0, p^1, q^0, q^1) \leq \max_i \{q_i^1/q_i^0 : i = 1, \dots, N\}.$$

That is, the implicit quantity index Q defined by P lies between the minimum and maximum rates of growth q_i^1/q_i^0 of the individual quantities.

In section 4, we argued that it was very reasonable to take an average of the Laspeyres and Paasche price indexes as a single “best” measure of overall price change. This point of view can be turned into a test:

T16: *Paasche and Laspeyres Bounding Test*: The price index P lies between the Laspeyres and Paasche indices, P_L and P_P , defined earlier by (7) and (8) above.

The final four tests are monotonicity tests; i.e., how should the price index $P(p^0, p^1, q^0, q^1)$ change as any component of the two price vectors p^0 and p^1 increases or as any component of the two quantity vectors q^0 and q^1 increases.

T17: *Monotonicity in Current Prices*: $P(p^0, p^1, q^0, q^1) < P(p^0, p^2, q^0, q^1)$ if $p^1 < p^2$.

That is, if some period 1 price increases, then the price index must increase, so that $P(p^0, p^1, q^0, q^1)$ is increasing in the components of p^1 . This property was proposed by Eichhorn and Voeller (1976; 23) and it is a very reasonable property for a price index to satisfy.

T18: *Monotonicity in Base Prices*: $P(p^0, p^1, q^0, q^1) > P(p^2, p^1, q^0, q^1)$ if $p^0 < p^2$.

That is, if any period 0 price increases, then the price index must decrease, so that $P(p^0, p^1, q^0, q^1)$ is decreasing in the components of p^0 . This very reasonable property was also proposed by Eichhorn and Voeller (1976; 23).

T19: *Monotonicity in Current Quantities*: if $q^1 < q^2$, then $\{\sum_{i=1}^N p_i^1 q_i^1 / \sum_{i=1}^N p_i^0 q_i^0\} / P(p^0, p^1, q^0, q^1) < \{\sum_{i=1}^N p_i^1 q_i^2 / \sum_{i=1}^N p_i^0 q_i^0\} / P(p^0, p^1, q^0, q^2)$.

T20: *Monotonicity in Base Quantities*: if $q^0 < q^2$, then $\{\sum_{i=1}^N p_i^1 q_i^1 / \sum_{i=1}^N p_i^0 q_i^0\} / P(p^0, p^1, q^0, q^1) > \{\sum_{i=1}^N p_i^1 q_i^1 / \sum_{i=1}^N p_i^0 q_i^2\} / P(p^0, p^1, q^2, q^1)$.

If we define the implicit quantity index Q that corresponds to P using (5), we find that T19 translates into the following inequality involving Q :

$$(21) Q(p^0, p^1, q^0, q^1) < Q(p^0, p^1, q^0, q^2) \text{ if } q^1 < q^2.$$

That is, if any period 1 quantity increases, then the implicit quantity index Q that corresponds to the price index P must increase. Similarly, we find that T20 translates into:

$$(22) \quad Q(p^0, p^1, q^0, q^1) > Q(p^0, p^1, q^2, q^1) \text{ if } q^0 < q^2.$$

That is, if any period 0 quantity increases, then the implicit quantity index Q must decrease. Tests T19 and T20 are due to Vogt (1980, 70).

The final test is Irving Fisher's (1921; 534) (1922; 72-81) third reversal test (the other two being T9 and T11):

T21: *Factor Reversal Test* (functional form symmetry test):

$$P(p^0, p^1, q^0, q^1) P(q^0, q^1, p^0, p^1) = \sum_{i=1}^N p_i^1 q_i^1 / \sum_{i=1}^N p_i^0 q_i^0 = V^1/V^0.$$

A justification for this test is the following one: if $P(p^0, p^1, q^0, q^1)$ is a good functional form for the price index, then if we reverse the roles of prices and quantities, $P(q^0, q^1, p^0, p^1)$ ought to be a good functional form for a quantity index (which seems to be a correct argument) and thus the product of the price index $P(p^0, p^1, q^0, q^1)$ and the quantity index $Q(p^0, p^1, q^0, q^1) = P(q^0, q^1, p^0, p^1)$ ought to equal the value ratio, V^1/V^0 . The second part of this argument does not seem to be valid and thus many researchers over the years have objected to the factor reversal test.

It is straightforward to show that the Fisher ideal price index P_F defined earlier by (9) satisfies all 21 tests. Is this the only index number formula that satisfies all of these tests? The answer is yes: Funke and Voeller (1978; 180) showed that the only index number function $P(p^0, p^1, q^0, q^1)$ which satisfies T1 (positivity), T11 (time reversal test), T12 (quantity reversal test) and T21 (factor reversal test) is the Fisher ideal index P_F defined by (9). Diewert (1992; 221) proved a similar result: namely that if P satisfied T1 and the three reversal tests T11-T13, then P must equal P_F .

Thus it seems that from the perspective of the above test approach to index number theory, the Fisher ideal index satisfies more "reasonable" tests than competing indexes and hence can be regarded as "best" from the viewpoint of this perspective.

There is another perspective to the test approach to index number theory. The above approach looked at axioms or tests that pertained to situations where the price index was a function of the two price vectors, p^0 and p^1 , and the two matching quantity vectors, q^0 and q^1 . In this framework, the two quantity vectors essentially act as weights for the prices. However, there is an alternative framework where the price index, say $P^*(p^0, p^1, e^0, e^1)$, is regarded as a function of the two price vectors, p^0 and p^1 , and the two

matching *expenditure vectors*, e^0 and e^1 .³¹ An axiomatic approach to the determination of the functional form for indexes of this type is developed in the ILO (2004; 307-309) and the Törnqvist index defined earlier by (18) emerges as “best” from the perspective of this second test approach to index number theory. Thus both the Fisher and Törnqvist indexes can be given strong axiomatic justifications.

There is one final important test that should be added to the above list of tests and that is the following *circularity test*³² which involves looking at the prices and quantities that pertain to three periods:

$$T22: \text{Circularity Test: } P(p^0, p^1, q^0, q^1) P(p^1, p^2, q^1, q^2) = P(p^0, p^2, q^0, q^2).$$

If this test is satisfied, then the rate of price change going from period 0 to 1, $P(p^0, p^1, q^0, q^1)$, times the rate of price change going from period 1 to 2, $P(p^1, p^2, q^1, q^2)$, is equal to the rate of price change going from period 0 to 2 directly, $P(p^0, p^2, q^0, q^2)$. If there is only one commodity in the aggregate, then the price index $P(p^0, p^1, q^0, q^1)$ just becomes the single price ratio, p_1^1/p_1^0 , and the circularity test T22 becomes the equation $[p_1^1/p_1^0][p_1^2/p_1^1] = [p_1^2/p_1^0]$, which is obviously satisfied. The equation in the circularity test illustrates the difference between chained index numbers and fixed base index numbers. The left hand side of T22 uses the *chain principle* to construct the overall inflation between periods 0 and 2 whereas the right hand side uses the *fixed base principle* to construct an estimate of the overall price change between periods 0 and 1.³³

It would be good if our preferred index number formulae, the Fisher, Walsh and Törnqvist indexes (P_F , P_W and P_T), satisfied the circularity test but unfortunately, none of these indexes satisfy T22. Thus if any of these indexes are used by a statistical agency, then the question arises: should the sequence of index values be computed using fixed base indexes or chained indexes? The remainder of this section will attempt to address this question.

The first point to note is that fixed base indexes cannot be used for long periods of time in today's dynamic economy where new commodities appear and older ones become obsolete. Under these conditions, it becomes increasingly difficult to match commodity prices over long periods of time and index number theory is dependent on a high degree of matching of the prices between the two periods being compared. However, this possible lack of matching does not rule out using fixed base indexes for shorter periods of time, say over a year or two.

³¹ Component n of the period t expenditure vector e^t is defined as $e_n^t \equiv p_n^t q_n^t$ for $n = 1, \dots, N$ and $t = 0, 1$. Thus if the price components p_n^t are known, then a knowledge of either the quantity components q_n^t or the expenditure components e_n^t will determine prices, quantities and expenditures in both periods.

³² The test name is due to Fisher (1922; 413) and the concept was originally due to Westergaard (1890; 218-219).

³³ Thus when the chain principle is used, the price index $P(p^t, p^{t+1}, q^t, q^{t+1})$ is used to update the period t index level to construct the period $t+1$ index level, whereas the fixed base system constructs the period $t+1$ index level relative to period 0 directly as $P(p^0, p^{t+1}, q^0, q^{t+1})$, where the period 0 level is set equal to 1. Fisher (1911; 203) introduced this fixed base and chain terminology. The concept of chaining is due to Lehr (1885) and Marshall (1887; 373).

The main advantage of using chained indexes is that if prices and quantities are trending relatively smoothly, chaining will reduce the spread between the Paasche and Laspeyres indexes.³⁴ These two indexes each provide an asymmetric perspective on the amount of price change that has occurred between the two periods under consideration and it could be expected that a single point estimate of the aggregate price change should lie between these two estimates. Thus the use of either a chained Paasche or Laspeyres index will usually lead to a smaller difference between the two and hence to estimates that are closer to the “truth”. Since annual data generally has smooth trends, the use of chained indexes is generally appropriate at this level of aggregation; see Hill (1993; 136-137).

However, the story is different at subannual levels; i.e., if the index is to be produced at monthly or quarterly frequencies. 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” to use Szulc’s (1983; 548) term. This phenomenon can occur in the context of regular seasonal fluctuations or in the context of sales. The *price bouncing problem* or the problem of *chain drift* can be illustrated if we make use of the following test due to Walsh (1901; 389), (1921b; 540) (1924; 506):³⁵

$$T23: \text{Multiperiod Identity Test: } P(p^0, p^1, q^0, q^1)P(p^1, p^2, q^1, q^2)P(p^2, p^0, q^2, q^0) = 1.$$

Thus price change is calculated over consecutive periods but an artificial final period is introduced where the prices and quantities revert back to the prices and quantities in the very first period. The Walsh test T23 asks that the product of all of these price changes should equal unity. If prices have no definite trends but are simply bouncing up and down in a range, then the above test can be used to evaluate the amount of chain drift that occurs if chained indexes are used under these conditions. *Chain drift* occurs when an index does not return to unity when prices in the current period return to their levels in the base period; see the ILO (2004; 445). Fixed base indexes operating under these conditions will not be subject to chain drift.

It is possible to be a bit more precise under what conditions one should chain or not chain. Basically, one should chain 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 indexes at each link. Of course, one needs a measure of how similar are the prices and quantities pertaining to two periods. A practical problem with this *similarity linking approach* is: exactly how should the measure of price or quantity similarity be measured?³⁶ For *annual*

³⁴ See Diewert (1978; 895) and Hill (1988) (1993; 387-388). Chaining under these conditions will also reduce the spread between fixed base and chained indexes using P_F , P_W or P_T as the basic bilateral formula.

³⁵ This is Diewert’s (1993a; 40) term for the test. Walsh did not limit himself to just three periods as in T23; he considered an indefinite number of periods. If tests T3 and T22 are satisfied, then T23 will also be satisfied.

³⁶ This similarity approach to linking bilateral comparisons into a complete set of comparisons across all observations has been pioneered by Robert Hill (1999a) (1999b) (2001) (2004) (2009). For an axiomatic approach to similarity measures, see Diewert (2009). We will address the similarity approach in more detail in Chapter 3.

time series data, it turns out that for various “reasonable” similarity measures, chained indexes are generally consistent with the similarity approach to linking observations. However, for subannual data, it is generally better to use fixed base indexes in order to eliminate the problem of chain drift.

We conclude this subsection with a discussion on how well our best indexes, P_F , P_W and P_T defined by (9), (13) and (18) above, satisfy the circularity test, T22. Fisher (1922; 277) found that for his annual data set, the Fisher ideal index P_F satisfied circularity to a reasonably high degree of approximation. It turns out that this result generally holds using annual data for P_W and P_T as well. It is possible to give a theoretical explanation for the approximate satisfaction of the circularity test for these three indexes. Alterman, Diewert and Feenstra (1999; 61) showed that if the logarithmic price ratios $\ln(p_n^t/p_n^{t-1})$ trend linearly with time t and the expenditure shares s_i^t also trend linearly with time, then the Törnqvist index P_T will satisfy the circularity test *exactly*.³⁷ Since many economic time series on prices and quantities satisfy these assumptions approximately, the above exactness result will imply that the Törnqvist index P_T will satisfy the circularity test approximately. But Diewert (1978; 888) showed that P_T , P_F and P_W numerically approximate each other to the second order around an equal price and quantity point and so these three indexes will generally be very close to each other using annual time series data. Hence since P_T will generally satisfy the circularity test to some degree of approximation, P_F and P_W will also satisfy circularity approximately in the time series context using annual data. Thus for *annual* economic time series, P_F , P_T and P_W will generally satisfy the circularity test to a high enough degree of approximation so that it will not matter whether we use the fixed base or chain principle. However, this same conclusion does *not* hold for *subannual* data that has substantial period to period fluctuations in prices. For fluctuating subannual data, chained indexes can give very unsatisfactory results; i.e., Walsh’s multiperiod identity test will be far from being satisfied. Under these conditions, fixed base indexes or multilateral methods should be used.³⁸

It should be mentioned that additional materials on the basket, stochastic and test approaches to index number theory can be found in the excellent volume by Balk (2008).

References

- Alterman, W.F., W.E. Diewert and R.C. Feenstra (1999), *International Trade Price Indexes and Seasonal Commodities*, Bureau of Labor Statistics, Washington D.C.
- Balk, B.M. (1981), “A Simple Method for Constructing Price Indices for Seasonal Commodities”, *Statistische Hefte* 22 (1), 1–8.

³⁷ This exactness result can be extended to cover the case when there are monthly proportional variations in prices and the expenditure shares have constant seasonal effects in addition to linear trends; see Alterman, Diewert and Feenstra (1999; 65).

³⁸ See Szulc (1983), Hill (1988) and Ivancic, Diewert and Fox (2011).

- Balk, B.M. (2008), *Price and Quantity Index Numbers*, New York: Cambridge University Press.
- Bowley, A.L. (1901), *Elements of Statistics*, Westminster: P.S. King and Son.
- Carli, Gian-Rinaldo, (1804), “Del valore e della proporzione de’ metalli monetati”, pp. 297-366 in *Scrittori classici italiani di economia politica*, Volume 13, Milano: G.G. Destefanis (originally published in 1764).
- Clements, K.W., H.Y. Izan and E.A. Selvanathan (2006), “Stochastic Index Numbers: A Review”, *International Statistical Review* 74, 235-270.
- Davies, G.R. (1924), “The Problem of a Standard Index Number Formula”, *Journal of the American Statistical Association* 19, 180-188.
- de Haan, J. and H.A. van der Grient (2011), “Eliminating Chain drift in Price Indexes Based on Scanner Data”, *Journal of Econometrics* 161, 36-46.
- de Haan, J. and F. Krsinich (2012), “The Treatment of Unmatched Items in Rolling Year GEKS Price Indexes: Evidence from New Zealand Scanner Data”, paper presented at the Meeting of Groups of Experts on Consumer Price Indices Organized jointly by UNECE and ILO at the United Nations Palais des Nations, Geneva Switzerland, May 30-June 1, 2012.
- 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. (1978), “Superlative Index Numbers and Consistency in Aggregation”, *Econometrica* 46, 883-900.
- 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), “Duality Approaches To Microeconomic Theory”, in *Essays in Index Number Theory*, pp. 105-175 in Volume I, Contributions to Economic Analysis 217, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North Holland.
- Diewert, W.E. (1995a), “Axiomatic and Economic Approaches to Elementary Price Indexes”, Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

- Diewert, W.E. (1995b), "On the Stochastic Approach to Index Numbers", Discussion Paper 95-31, Department of Economics, University of British Columbia, Vancouver, Canada.
- 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. (2001), "The Consumer Price Index and Index Number Purpose", *Journal of Economic and Social Measurement* 27, 167-248.
- 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, V6T 1Z1.
- Diewert, W.E. (2012), "Consumer Price Statistics in the UK", unpublished paper, June 11, 2012.
- Diewert, W.E., Y. Finkel and Y. Artsev (2009), "Empirical Evidence on the Treatment of Seasonal Products: The Israeli Experience", pp. 53-78 in *Price and Productivity Measurement: Volume 2: Seasonality*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Victoria, Canada: Trafford Press.
- Drobisch, M. W. (1871), "Ueber die Berechnung der Veränderungen der Waarenpreise und des Geldwerths", *Jahrbücher für Nationalökonomie und Statistik* 16, 143-156.
- 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. (1896), "A Defense of Index Numbers", *Economic Journal* 6, 132-142.
- Edgeworth, F.Y. (1901), "Mr. Walsh on the Measurement of General Exchange Value", *Economic Journal* 11, 404-416.
- Edgeworth, F.Y. (1925), *Papers Relating to Political Economy*, Volume 1, New York: Burt Franklin.
- 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.
- Feenstra, Robert C. and Matthew 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*, Boston: Houghton-Mifflin.
- Frisch, R. (1930), “Necessary and Sufficient Conditions Regarding the Form of an Index Number which Shall Meet Certain of Fisher’s Tests”, *American Statistical Association Journal* 25, 397-406.
- Funke, H. and J. Voeller (1978), “A Note on the Characterization of Fisher’s Ideal Index,” pp. 177-181 in *Theory and Applications of Economic Indices*, W. Eichhorn, R. Henn, O. Opitz and R.W. Shephard (eds.), Würzburg: Physica-Verlag.
- Hill, R.J. (1999a), “Comparing Price Levels across Countries Using Minimum Spanning Trees”, *The Review of Economics and Statistics* 81, 135-142.
- Hill, R.J. (1999b), “International Comparisons using Spanning Trees”, pp. 109-120 in *International and Interarea Comparisons of Income, Output and Prices*, A. Heston and R.E. Lipsey (eds.), Studies in Income and Wealth Volume 61, NBER, Chicago: The University of Chicago Press.
- Hill, R.J. (2001), “Measuring Inflation and Growth Using Spanning Trees”, *International Economic Review* 42, 167-185.
- Hill, R.J. (2004), “Constructing Price Indexes Across Space and Time: The Case of the European Union”, *American Economic Review* 94, 1379-1410.
- Hill, R.J. (2009), “Comparing Per Capita Income Levels Across Countries Using Spanning Trees: Robustness, Prior Restrictions, Hybrids and Hierarchies”, pp. 217-244 in *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, D.S. Prasada Rao (ed.), Cheltenham UK: Edward Elgar.
- Hill, T.P. (1988), “Recent Developments in Index Number Theory and Practice”, *OECD Economic Studies* 10, 123-148.
- Hill, T.P. (1993), “Price and Volume Measures”, pp. 379-406 in *System of National Accounts 1993*, Eurostat, IMF, OECD, UN and World Bank, Luxembourg, Washington, D.C., Paris, New York, and Washington, D.C.
- ILO/IMF/OECD/UNECE/Eurostat/World Bank (2004), *Consumer Price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: International Labour Office.

- 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.
- 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.
- Knibbs, Sir G.H. (1924), "The Nature of an Unequivocal Price Index and Quantity Index", *Journal of the American Statistical Association* 19, 42-60 and 196-205.
- Laspeyres, E. (1871), "Die Berechnung einer mittleren Waarenpreissteigerung", *Jahrbücher für Nationalökonomie und Statistik* 16, 296-314.
- Lehr, J. (1885), *Beiträge zur Statistik der Preise*, Frankfurt: J.D. Sauerlander.
- Lowe, J. (1823), *The Present State of England in Regard to Agriculture, Trade and Finance*, second edition, London: Longman, Hurst, Rees, Orme and Brown.
- Marshall, A. (1887), "Remedies for Fluctuations of General Prices", *Contemporary Review* 51, 355-375.
- Paasche, H. (1874), "Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen", *Jahrbücher für Nationalökonomie und Statistik* 12, 168-178.
- Pierson, N.G. (1896), "Further Considerations on Index-Numbers," *Economic Journal* 6, 127-131.
- Selvanathan, E.A. and D.S. Prasada Rao (1994), *Index Numbers: A Stochastic Approach*, Ann Arbor: The University of Michigan Press.
- Sidgwick, H. (1883), *The Principles of Political Economy*, London: Macmillan.
- Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.
- Szulc, B.J. (1983), "Linking Price Index Numbers," pp. 537-566 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

- Rao, D.S. Prasada (1990), "A System of Log-Change Index Numbers for Multilateral Comparisons", pp. 127-139 in *Comparisons of Prices and Real Products in Latin America*, J. Salazar-Carillo and D.S. Prasada Rao (eds.), New York: Elsevier Science Publishers.
- 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 (2001), "Weighted EKS and Generalized CPD Methods for Aggregation at the Basic Heading Level and Above Basic Heading Level", Paper presented at the Joint World Bank and OECD Seminar on Purchasing Power Parities, January 30-February 2, Washington D.C.
- Rao, D.S. Prasada (2002), "On the Equivalence of Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", School of Economics, University of New England, Armidale, Australia, March.
- 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.
- 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 förhållandet 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.
- van der Grient, H.A. and J. de Haan (2011), "Scanner Data Price Indexes: The 'Dutch Method' versus Rolling Year GEKS", paper presented at the 12th Meeting of the Ottawa Group, May 4-6, Wellington, New Zealand.
- Vogt, A. (1980), "Der Zeit und der Faktorumkehrtest als 'Finders of Tests'", *Statistische Hefte* 21, 66-71.
- Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.
- Walsh, C.M. (1921a), *The Problem of Estimation*, London: P.S. King & Son.

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

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.

World Bank (2014),

Wynne, M.A. (1997), "Commentary on Measuring short Run Inflation for Central Bankers", *Federal Reserve Bank of St. Louis Review* 79:3, 161-167.