

ON THE STOCHASTIC APPROACH TO INDEX NUMBERS

by W.E. Diewert*

September 10, 1995

In mathematics disputes must soon come to an end, when the one side is proved and the other disproved. And where mathematics enters into economics, it would seem that little room could be left for long-continued disputation. It is therefore somewhat surprising that one economist after another takes up the subject of index-numbers, potters over it for a while, differs from the rest if he can, and then drops it. And so nearly sixty years have gone by since Jevons first brought mathematics to bear upon this question, and still economists are at loggerheads over it. Yet index-numbers involve the use of means and averages, and these being a purely mathematical element, demonstration ought soon to be reached, and then agreement should speedily follow.

Walsh [1921; preface].

1. Introduction

The recent appearance of a book on the stochastic approach to index number theory by Selvanathan and Prasada Rao [1994] marks an appropriate occasion to provide a critical review of this approach. This is the primary purpose of the present paper.

The stochastic approach¹ to index number theory originated with Jevons [1863; 23-26] [1865; 121-122] [1869; 156-157], Edgeworth [1887; 245] [1888a] [1888b] [1889; 286-292] and Bowley [1901; 219] [1911] [1919; 346] [1926] [1928; 217]. Basically, this approach was driven by the quantity theory of money: as the quantity of gold or money increased, all prices should increase approximately proportionally. Thus a measure of the general increase in prices going from period 0 to period t could be obtained by taking an appropriate average of price relatives, p_{it}/p_{i0} , where p_{it} denotes the price of commodity i in period t . This average of the price relatives can be regarded as an index number of price change going from period 0 to t . Selvanathan and Prasada Rao [1994; 5-6] express this ancient theory in more modern language as follows:

The stochastic approach considers the index number problem as a signal extraction problem from the messages concerning price changes for different commodities. Obviously the strength of the signal extracted depends upon the messages received and the information context of the messages.

The recent resurrection of the stochastic approach to index number theory is due to Balk [1980], Clements and Izan [1981] [1987], Bryan and Cecchetti [1993] and Selvanathan and Prasada

Rao [1994]². The main attraction of the stochastic approach over competing approaches to index number theory is its ability to provide confidence intervals for the estimated inflation rates:

Accordingly, we obtain a point estimate of not only the rate of inflation, but also its sampling variance. The source of the sampling error is the dispersion of relative prices from their trend rates of change -- the sampling variance will be larger when the deviations of the relative prices from their trend rates of change are larger. This attractive result provides a formal link between the measurement of inflation and changes in relative prices.

Clements and Izan [1987; 339]

Selvanathan and Prasada Rao note the above advantage but go further and claim that the stochastic approach can be utilized to derive standard errors for many well known index number formulae:

The attraction of this approach is that it provides an alternative interpretation to some of the well known index numbers as the estimators of parameters of specific regression models. For example, the Laspeyres, Paasche, Theil-Törnqvist and other index numbers can be derived from various regression models. Further this approach provides standard errors for these index numbers.

Selvanathan and Prasada Rao [1994; 6].

At this point, it should be mentioned that the two main competing approaches to index number theory are the test approach and the economic approach.

The test approach can apply to two periods (the bilateral case) or to many periods (the multilateral case). The bilateral test approach assumes that complete price and quantity information on the relevant set of commodities is available for the two periods under consideration, say periods s and t . Denote the price and quantity vectors for these two periods by p^s, p^t and q^s, q^t , where $p^s \equiv [p_{1s}, \dots, p_{N_s}]$, etc. A bilateral price index is defined as a function P of the four sets of variables, $P(p^s, p^t, q^s, q^t)$. The bilateral test approach attempts to determine the functional form for P by assuming that P satisfies certain plausible tests, axioms or mathematical properties. In the case of only one commodity in the set of commodities to be aggregated, the imposed tests generally cause the price index $P(p_{1s}, p_{1t}, q_{1s}, q_{1t})$ to collapse down to the single price ratio, p_{1t}/p_{1s} . There is an

analogous bilateral test approach for the quantity index $Q(p^s, p^t, q^s, q^t)$. Fisher [1911; 403] observed that in the present context of complete information on prices and quantities, the price and quantity indexes, P and Q , should satisfy the following conservation of value equation:

$$(1) \quad P(p^s, p^t, q^s, q^t) Q(p^s, p^t, q^s, q^t) = p^t \cdot q^t / p^s \cdot q^s$$

where $p^t \cdot q^t \equiv \sum_{n=1}^N p_{nt} q_{nt}$. The importance of (1) is that once the functional form for P has been determined, then (1) automatically determines the functional form for Q . Moreover, tests for the quantity index Q can be translated into tests for the corresponding price index P defined via (1). Useful references for the test approach are Walsh [1901] [1921] [1924], Fisher [1911] [1921] [1922], and Diewert [1992a] [1993a; 6-10]. The early history of the test approach is reviewed by Frisch [1936; 5-7] and Diewert [1993b; 38-41].

In the test approach, the vectors of prices and quantities for the two periods are regarded as independent variables. In the economic approach, the two price vectors are regarded as independent variables but the quantity vectors are regarded as solutions to various economic maximization or minimization problems. In the consumer price context, it is assumed that the consumer has preferences over N commodities and these preferences can be represented by an aggregator or utility function $f(q_1, \dots, q_N) \equiv f(q)$. It is also assumed that in each period t , the consumer minimizes the cost $C[f(q^t), p^t]$ of achieving the utility level $f(q^t)$ when facing the period t vector of prices $p^t \equiv [p_{1t}, p_{2t}, \dots, p_{Nt}]$. The Konüs [1924] true cost of living index between periods s and t , using the reference utility level $f(q)$, is defined as the ratio of costs of achieving the reference utility level when facing the period s and t prices, $C[f(q), p^s] / C[f(q), p^t]$. If the consumer's utility function is linearly homogeneous, then the cost function $C[f(q), p]$ factors into two components, $f(q)c(p)$, where $c(p)$ is defined as the unit (utility level) cost function, $C[1, p]$. In this homogeneous case, the Konüs true cost of living index reduces to the unit cost ratio, $c(p^t) / c(p^s)$, and the corresponding quantity index is the utility ratio, $f(q^t) / f(q^s)$. Finally, consider a given formula for the price index, say

$P(p^s, p^t, q^s, q^t)$. We say that P is exact for the consumer preferences dual to the unit cost function c if under the assumption of cost minimizing behavior on the part of the consumer for periods s and t , we have

$$(2) \quad P(p^s, p^t, q^s, q^t) = c(p^t)/c(p^s).$$

Similarly, a given functional form for the quantity index, $Q(p^s, p^t, q^s, q^t)$, is exact for the linearly homogeneous utility function f , if under the assumption of cost minimizing behavior for periods s and t , we have

$$(3) \quad Q(p^s, p^t, q^s, q^t) = f(q^t)/f(q^s).$$

The economic approach to index number theory concentrates on finding functional forms for price indexes P that are exact for flexible³ unit cost functions c and on finding functional forms for quantity indexes Q that are exact for flexible linearly homogeneous utility functions f . Index number formulae that are exact for flexible functional forms are called superlative.⁴ The theory of exact index numbers was developed by Konüs and Byushgens [1926], Afriat [1972; 44-47], Samuelson and Swamy [1974] and Pollak [1989; 15-32]. The early history of exact index numbers is reviewed in Diewert [1993b; 45-50]. For examples of superlative indexes, see Diewert [1976] [1978] [1992b; 576].

As can be seen from the above brief reviews of the test and economic approaches to index number theory,⁵ these approaches are silent on the problem of providing an estimate of the reliability of the suggested bilateral index number formulae. Thus the new champions of the stochastic approach appear to have a strong a priori argument in favor of their approach.

In section 2 below, we review the original approaches of Jevons, Edgeworth and Bowley. In section 3, we review the initial new stochastic approaches of Clements and Izan [1981] and Selvanathan and Prasada Rao [1994; 51-61]. In section 4, we review the more sophisticated stochastic approaches of Balk [1980], Clements and Izan [1987] and Selvanathan and Prasada Rao

[1994; 61-110]. The stochastic specifications that are utilized in the models presented in sections 3 and 4 are easily rejected from an empirical point of view. Thus in section 5, we present a new stochastic model that seems to be in the spirit of the type of model that Edgeworth had in mind but was never able to implement. In section 6, we present some practical criticisms of the new stochastic approaches to index number theory that will make it difficult for Statistical Agencies to embrace these approaches. Section 7 concludes by reconsidering the problems involved in providing measures of reliability for index numbers based on the test or economic approaches.

2. The Early Statistical Approaches to Index Number Theory

We assume that we are given price and quantity data, p_{it} and q_{it} , for periods $t=0,1,\dots,T$ and for commodities $i=1,2,\dots,N$. The first stochastic index number model that Selvanathan and Prasada Rao [1994; 49-51] consider is the following one: for $t=1,2,\dots,T$:

$$(4) \quad p_{it}/p_{i0} = \alpha_t + \varepsilon_{it}; \quad i=1,2,\dots,N;$$

where α_t represents the systematic part of the price change going from period 0 to t and the independently distributed random variables ε_{it} satisfy the following assumptions:

$$(5) \quad E \varepsilon_{it} = 0; \quad \text{Var } \varepsilon_{it} = \sigma_t^2; \quad i=1,2,\dots,N;$$

i.e., ε_{it} has mean 0 and variance $\sigma_t^2 > 0$. The least squares and maximum likelihood estimator for α_t in Model 1 defined by (4) and (5) is the Carli [1764] price index:

$$(6) \quad \hat{\alpha}_t \equiv \sum_{i=1}^N (1/N) p_{it}/p_{i0},$$

which is the unweighted arithmetic mean of the period 0 to t price relatives, p_{it}/p_{i0} . The variance of $\hat{\alpha}_t$ is

$$(7) \quad \text{Var } \hat{\alpha}_t = (1/N)\sigma_t^2$$

and Selvanathan and Prasada Rao [1994; 51] observe that an unbiased estimator for the variance σ_t^2 is

$$(8) \quad \hat{\sigma}_t^2 = [1/(N-1)] \sum_{i=1}^N [(p_{it}/p_{i0}) - \hat{\alpha}_t]^2.$$

Using (7) and (8), a confidence interval for the Carli price index $\hat{\alpha}_t$ can be calculated under the assumption of normally distributed errors. As Selvanathan and Prasada Rao [1994; 51] note, if the dispersion of the price relatives p_{it}/p_{i0} increases, then the precision of our period t fixed base price index $\hat{\alpha}_t$ will decline.

Instead of assuming that the independent errors ε_{it} are additive, we could more plausibly assume that the errors are multiplicative.⁶ This leads to Model 2, which is defined by the following equations for $t=1, \dots, T$:

$$(9) \quad \ln [p_{it}/p_{i0}] = \pi_t + \varepsilon_{it}; \quad i=1, \dots, N;$$

$$(10) \quad E \varepsilon_{it} = 0; \quad \text{Var } \varepsilon_{it} = \sigma_t^2; \quad i=1, \dots, N.$$

The least squares and maximum likelihood estimator for π_t in Model 2 is

$$(11) \quad \hat{\pi}_t = [1/N] \sum_{i=1}^N \ln [p_{it}/p_{i0}].$$

A variance estimator for $\hat{\pi}_t$ can be constructed in a manner analogous to the use of (7) and (8) in Model 1. If we define α_t to be the exponential of π_t , we can exponentiate $\hat{\pi}_t$ to obtain the following estimator for α_t :

$$(12) \quad \exp [\hat{\pi}_t] = \prod_{i=1}^N [p_{it}/p_{i0}]^{1/N}.$$

The right hand side of (12) is the Jevons [1863; 53] geometric mean price index. Jevons [1869; 157] later applied least squares theory to equation (9) and calculated a "probable error" (or confidence interval in modern terminology) for his estimator $\hat{\pi}_t$ defined by (11). This appears to be the first relatively complete exposition of the stochastic approach to index number theory.

Jevons [1865; 120-122] also used the arithmetic mean index number (6) in his empirical work but he did not report any confidence intervals for his Carli indexes. Edgeworth [1887; 226-246] considered both arithmetic and geometric mean (unweighted) index numbers and Edgeworth [1888a] was entirely devoted to the problems involved in constructing confidence intervals for these indexes.

Bowley [1901; 203-229] [1919; 345-346] [1928; 216-222] was very much concerned with the problems involved in determining the precision of index numbers.⁷ Bowley [1911] was concerned with the precision of weighted index numbers while Bowley [1926] extended his earlier work to cover the case of correlated price relatives. Finally, Bowley was aware that precision in official indexes was rather important, since so many government expenditures were indexed to official price indexes. The following quotation refers to a potential upward bias of 18 percentage points in the Ministry of Labour index numbers for the UK over the years 1914-1918:

Every 4 points cost over a million pounds in the annual railway wage bill.
Bowley [1919; 348].

We turn now to an exposition of the new stochastic models.

3. The New Stochastic Approach to Index Numbers

Model 3 consists of equations (4) again but our old assumptions (5) on the independently distributed errors ε_{it} are now replaced by the following assumptions:

$$(13) \quad E \varepsilon_{it} = 0; \text{Var } \varepsilon_{it} = \sigma_i^2/w_i; i=1, \dots, N$$

where the w_i are nonrandom fixed shares to be determined later; i.e. the w_i satisfy

$$(14) \quad w_i > 0 \text{ for } i=1, 2, \dots, N \quad \text{and} \quad \sum_{i=1}^N w_i = 1.$$

Since the w_i are positive, we can multiply both sides of equation i in (4) by the square root of w_i , $w_i^{1/2}$, in order to obtain homoscedastic errors. The resulting least squares and maximum likelihood estimator for the period 0 to t inflation rate α_t is

$$(15) \quad \hat{\alpha}_t \equiv \sum_{i=1}^N w_i [p_{it}/p_{i0}] / \sum_{i=1}^N w_i = \sum_{i=1}^N w_i [p_{it}/p_{i0}]$$

where the second equality follows using (14). Using (13), it can be seen that $\hat{\alpha}_t$ is an unbiased estimator for α_t and its variance is

$$(16) \quad \text{Var } \hat{\alpha}_t = \sum_{i=1}^N w_i^2 [\sigma_i^2/w_i] = \sigma_t^2$$

where the second equality follows using (14). An unbiased estimator for σ_t^2 is

$$(17) \quad \hat{\sigma}_t^2 \equiv [1/(N-1)] \sum_{i=1}^N w_i [(p_{it}/p_{i0}) - \hat{\alpha}_t]^2.$$

Under the additional assumption that the residuals ε_{it} are normally distributed, (16) and (17) may be used to obtain confidence intervals for the share weighted index numbers $\hat{\alpha}_t$ defined by (15).

Selvanathan and Prasada Rao [1994; 51-55] consider the following special cases for the w_i :

$$(18) \quad w_i \equiv p_{i0}q_{i0} / \sum_{i=1}^N p_{n0}q_{n0}; \quad i=1, \dots, N;$$

$$(19) \quad w_i \equiv p_{i0}q_{it} / \sum_{i=1}^N p_{n0}q_{nt}; \quad i=1, \dots, N.$$

In order to make the w_i fixed variables, we need to assume that base period prices and quantities, p_{i0} and q_{i0} , and current period quantities, q_{it} , are fixed. Thus in equations (4), the only random variables are the current period prices p_{it} .

Substituting (18) into (15) causes $\hat{\alpha}_t$ to become the fixed base Laspeyres price index, $p^t \cdot q^0/p^0 \cdot q^0$, and substituting (19) into (15) leads to the Paasche price index, $p^t \cdot q^t/p^0 \cdot q^t$. Furthermore, substitution of (18) and (19) into (15)-(17) yields estimators for the variances of the fixed base Laspeyres and Paasche price indexes. Thus the new stochastic approach of Selvanathan and Prasada Rao does lead to estimates of the precision of these well known indexes (provided that their stochastic assumptions (13) are correct).

We turn now to the new stochastic approach of Clements and Izan [1981]. Consider two distinct periods s and t where $0 \leq s < t \leq T$. Let π_{st} be the logarithm of the price change going from period s to t . The equations which define Model 4 are:

$$(20) \quad \ln [p_{it}/p_{is}] = \pi_{st} + \varepsilon_{ist}; \quad i=1, \dots, N;$$

$$(21) \quad E \varepsilon_{ist} = 0; \quad \text{Var } \varepsilon_{ist} = \sigma_{st}^2/w_i; \quad i=1, \dots, N$$

where the weights w_i again satisfy (14). Multiplying both sides of (20) through by $(w_i)^{1/2}$ leads to homoscedastic variances. The least squares and maximum likelihood estimator for π_{st} in this transformed model is

$$(22) \quad \hat{\pi}_{st} = \sum_{i=1}^N w_i \ln [p_{it}/p_{is}].$$

Using (21), the variance of $\hat{\pi}_{st}$ is σ_{st}^2 . An unbiased estimator for σ_{st}^2 is

$$(23) \quad \hat{\sigma}_{st}^2 \equiv [1/(N-1)] \sum_{i=1}^N w_i [\ln(p_{it}/p_{is}) - \hat{\pi}_{st}]^2.$$

Let $w_{it} \equiv p_{it}q_{it}/\sum_{i=1}^N p_{it}q_{it}$ be the expenditure share of commodity i in period t . Clements and Izan [1981; 745-746] and Selvanathan and Prasada Rao [1994; 76-77] choose the weights w_i which appear in (21) as follows:⁸

$$(24) \quad w_i \equiv (1/2)w_{is} + (1/2)w_{it}; \quad i=1, \dots, N;$$

i.e., w_i is chosen to be the average expenditure share on commodity i over periods s and t .

Substituting (24) into (22) yields

$$(25) \quad \exp [\hat{\pi}_{st}] = \prod_{i=1}^N [p_{it} / p_{is}]^{(1/2)w_{is} + (1/2)w_{it}}.$$

The right hand side of (25) is known as the Törnqvist [1936] price index.⁹

Under the assumption of normally distributed errors, (23) can be used to form confidence intervals for $\hat{\pi}_{st}$, the logarithm of the Persons-Törnqvist price index. However, since the weights w_i defined by (24) depend on p_{is} and p_{it} , it will be necessary to assume that the conditional (on w_i) distribution of $\ln(p_{it}/p_{is})$ is normal and satisfies assumptions (21). Thus the stochastic assumptions justifying Model 4 are more tenuous than those for Model 3 above.

The variance assumptions (13) and (21), $\text{Var } \varepsilon_{it} = \sigma_i^2/w_i$ and $\text{Var } \varepsilon_{ist} = \sigma_{st}^2/w_i$, require some justification.¹⁰ The following quotation indicates how Clements and Izan justify their assumptions on the variances of the log price relatives:

If all goods were equally important, then the assumption that $\text{var } \varepsilon_i$ is the same for all i would be acceptable. However, this is not the case, since the budget share \bar{w}_i varies with i . If we think in terms of sampling the individual prices to form Dp_i for each commodity group, then it seems reasonable to postulate that the collection agency invests more resources in sampling the prices of those goods more important in the budget. This implies that $\text{var } \varepsilon_i$ is inversely proportional to \bar{w}_i .

Clements and Izan [1981; 745].

In contrast to the explicit sampling approach of Clements and Izan [1981], Selvanathan and Prasada Rao [1994] (with the exception of their section 7.4) regarded their prices as being accurately known, or in any case, they wanted their analysis to apply to this case.¹¹ They justify their variance assumptions in (13) and (18) as follows:¹²

Under this assumption we have that the variance of the price relative of i is λ_i^2/w_{i0} and is inversely proportional to w_{i0} . This means that the variability of a price relative falls as the commodity becomes more important in the consumer's budget.

Selvanathan and Prasada Rao [1994; 52].

In their more sophisticated stochastic model to be discussed in the next section, Clements and Izan [1987] no longer relied on their earlier sampling theory justification for their variance assumptions of the form (21). Instead, they provided the following justification:

As ε_{it} is the change in the i th relative price, specification (7) implies that the variability of a relative price falls as the commodity becomes more important in the consumer's budget. Thus the variability of a relative price of a good having a large budget share, such as food, will be lower than that of a commodity with a smaller share, such as cigarettes. This is a plausible specification, since there is less scope for a relative price to change as the commodity in question grows in importance in the budget.

Clements and Izan [1987; 341].

As can be seen from the above quotations, the justifications presented for the variance assumptions in the new stochastic approaches are rather weak.¹³ We will return to this point in section 5 below.

Clements and Izan [1981; 747] and Selvanathan and Prasada Rao [1994; 89] point out a positive feature of the new stochastic models such as Model 3 or 4: the resulting index numbers such as (15) or (22) are invariant to the level of commodity aggregation, provided that the same shares w_i that appear in the variance specifications (13) or (21) are used to do the aggregation. To see this, consider Model 3 represented by (4) and (13) and suppose that commodities 1 and 2 are aggregated together. Let p_{At} be the price of the aggregate commodity in period t . The weights w_1 and w_2 are used to define the following aggregate period 0 to t price relative:

$$(26) \quad p_{At}/p_{A0} \equiv [w_1/(w_1+w_2)][p_{1t}/p_{01}] + [w_2/(w_1+w_2)][p_{2t}/p_{20}].$$

Replace the first two equations in (4) by the new aggregated equation $p_{At}/p_{A0} = \alpha_t + \varepsilon_{At}$. Using the first two equations in (4) as well as (26), it can be seen that the new aggregate error is equal to

$$(27) \quad \varepsilon_{At} = [w_1/(w_1+w_2)] \varepsilon_{1t} + [w_2/(w_1+w_2)] \varepsilon_{2t}.$$

Using (13) and (27), the expectation of p_{At}/p_{A0} is equal to α_t , the expectation of ε_{At} is 0 and the variance of ε_{At} is

$$(28) \quad \text{Var } \varepsilon_{At} = [w_1/(w_1+w_2)]^2 [\sigma_{1t}^2/w_1] + [w_2/(w_1+w_2)]^2 [\sigma_{2t}^2/w_2] = \sigma_t^2/(w_1+w_2).$$

Thus the mean and variance of the aggregated error are of the same form as the means and variances of the original errors, ε_{1t} and ε_{2t} ; see (13). It is straightforward to show that the maximum likelihood estimator $\hat{\alpha}_t^*$ for α_t in the aggregated model is equal to the disaggregated estimator $\hat{\alpha}_t$ defined by (15).

We turn now to more sophisticated new stochastic approaches to price indexes.

4. A Specific Price Trends Stochastic Approach

The models presented in the previous section are similar to the classical stochastic models presented in section 2, except that the variance assumptions were different. These simple signal extraction models were effectively criticized by Keynes [1930; 58-84]. Clements and Izan summarize this Keynesian criticism as follows:

Thus the rate of inflation can be estimated by averaging over these n observations. This approach was correctly criticized by Keynes (1930, pp. 85-88) on the basis that it requires the systematic component of each price change to be identical. In other words, all prices must change equiproportionally so that there can be no changes in relative prices. The objective of this article is to rehabilitate the stochastic approach by answering Keynes's criticism by allowing for systematic changes in relative prices.

Clements and Izan [1987; 339]

Selvanathan and Prasada Rao [1994; 61] also acknowledge that the Keynesian criticism applies to their Laspeyres and Paasche models (Model 3 with the w_i defined by (18) and (19)

respectively). In order to rectify this deficiency in their Laspeyres model, Selvanathan and Prasada Rao [1994; 61-73] generalize their model as follows: assume that the period t over period 0 price ratios satisfy

$$(29) \quad p_{it}/p_{i0} = \alpha_t + \beta_i + \varepsilon_{it}; \quad i=1, \dots, N; \quad t=1, \dots, T$$

where the independently distributed residuals ε_{it} satisfy the following assumptions:

$$(30) \quad E \varepsilon_{it} = 0; \quad \text{Var } \varepsilon_{it} = \sigma_t^2/w_i; \quad i=1, \dots, N; \quad t=1, \dots, T.$$

As usual, the positive variance weights w_i are assumed to be shares; i.e., the w_i satisfy (14). Selvanathan and Prasada Rao [1994; 62] interpret β_i as the expectation of the change in the i th relative price in addition to general inflation; i.e., it is the systematic part of commodity i price change in addition to the overall period 0 to t price change α_t .¹⁴ Selvanathan and Prasada Rao [1994; 62] note that the parameters α_t and β_i are not identified. Thus they add an identifying restriction of the following form:

$$(31) \quad \sum_{i=1}^N w_i \beta_i = 0.$$

The restriction (31) says that a share weighted average of the specific commodity price trends β_i sums to zero, a very reasonable assumption since the parameter α_t contains the general period t trend. What is not so reasonable, however, is the assumption that the w_i which appear in (31) are the same as the w_i which appear in (30).

Let us call the model which consists of (14) and (29)-(31) Model 5. Maximum likelihood estimators, $\hat{\alpha}_t$, $\hat{\beta}_i$, and $\hat{\sigma}_t^2$, for the parameters which appears in this Model can be obtained in a manner analogous to the way Selvanathan and Prasada Rao [1994; 63-66] derived estimators for their specific version of this Model. Define the maximum likelihood residuals $\hat{\varepsilon}_{it}$ by:

$$(32) \quad \hat{\varepsilon}_{it} \equiv (p_{it}/p_{i0}) - \hat{\alpha}_t - \hat{\beta}_i; \quad i=1, \dots, N; \quad t=1, \dots, T.$$

The maximum likelihood estimators for the parameters of Model 5 can be obtained by solving the following system of equations, along with equations (32):

$$(33) \quad \hat{\alpha}_t = \sum_{i=1}^N w_i p_{it}/p_{i0} ; \quad t=1, \dots, T;$$

$$(34) \quad \hat{\sigma}_t^2 = (1/N) \sum_{i=1}^N w_i \hat{\varepsilon}_{it}^2 ; \quad t=1, \dots, T;$$

$$(35) \quad \hat{\beta}_i = \sum_{t=1}^T [1/\hat{\sigma}_t^2] [(p_{it}/p_{i0}) - \hat{\alpha}_t] / \sum_{t=1}^T [1/\hat{\sigma}_t^2] ; \quad i=1, \dots, N.$$

Substitution of equations (33) into (35) shows that the $\hat{\beta}_i$ satisfy the restriction (31). Equations (34) show that the period t variance estimator $\hat{\sigma}_t^2$ is a weighted sum of the squares of the period t maximum likelihood residuals, $\hat{\varepsilon}_{it}^2$. Equations (35) show that the i th commodity effect $\hat{\beta}_i$ is a weighted average over T periods of the deviations of the period 0 to t price relatives p_{it}/p_{i0} from the period t general inflation rates $\hat{\alpha}_t$, where the weights are inversely proportional to the period t variance estimates, $\hat{\sigma}_t^2$. Equations (33) show that the estimator for the period 0 to t general inflation rate α_t is a simple weighted average of the period 0 to t price relatives, p_{it}/p_{i0} -- an amazingly simple result!

If we let the weights w_i equal the base period expenditure shares w_{i0} , we obtain the specific price trends stochastic model of Selvanathan and Prasada Rao [1994; 61-67] and the period 0 to t inflation estimate $\hat{\alpha}_t$ defined by (33) collapses down to the fixed base Laspeyres price index, $p^t \cdot q^0/p^0 \cdot q^0$. It is easy to show that $\hat{\alpha}_t$ is an unbiased estimator for α_t with the variance σ_t^2 . Thus Selvanathan and Prasada Rao feel that they have justified the use of the fixed base Laspeyres price index (and provided measures of its variability) from the viewpoint of a sophisticated stochastic approach that blunts the force of the Keynesian objection to stochastic index number models.

However, a problem with Model 5 is that its specification for the specific commodity effects β_i in equations (29) is not very compelling. A more credible specific price trends stochastic model was developed by Clements and Izan [1987; 341-345] and repeated by Selvanathan and Prasada Rao [1994; 78-87]. The equations which characterize the model of these authors are:

$$(36) \quad \ln [p_{it}/p_{i,t-1}] = \pi_t + \beta_i + \varepsilon_{it} ; \quad i=1, \dots, N ; \quad t=1, \dots, T;$$

$$(37) \quad E \varepsilon_{it} = 0 ; \quad \text{Var } \varepsilon_{it} = \sigma_t^2/w_i ; \quad i=1, \dots, N ; \quad t=1, \dots, T;$$

As usual, the variance weights w_i which appear in (37) are assumed known and assumed to satisfy (14). As in the previous model, the π_t and β_i are not identified. Hence Clements and Izan [1987: 342] assume that the β_i satisfy the following restriction:

$$(38) \quad \sum_{i=1}^N w_i \beta_i = 0$$

where the w_i weights which appear in (38) are the same as those appearing in (37). It is this coincidence which leads to the following elegant formulae for the maximum likelihood estimators for the parameters of Model 6, consisting of (14) and (36)-(38):

$$(39) \quad \hat{e}_{it} \equiv \ln [p_{it}/p_{it-1}] - \hat{\pi}_t - \hat{\beta}_i; \quad i=1, \dots, N; \quad t=1, \dots, T;$$

$$(40) \quad \hat{\pi}_t = \sum_{i=1}^N w_i \ln [p_{it}/p_{it-1}]; \quad t=1, \dots, T;$$

$$(41) \quad \hat{\sigma}_t^2 = (1/N) \sum_{i=1}^N w_i \hat{e}_{it}^2; \quad t=1, \dots, T;$$

$$(42) \quad \hat{\beta}_i = \sum_{t=1}^T [1/\hat{\sigma}_t^2] [\ln(p_{it}/p_{it-1}) - \hat{\pi}_t] / \sum_{s=1}^T [1/\hat{\sigma}_s^2]; \quad i=1, \dots, N.$$

The interpretation of (40) to (42) is analogous to the earlier interpretation of (33)-(35). However, the interpretation of the specific commodity price trend parameters β_i is much more reasonable for Model 6 than for Model 5: the β_i in the i th equation of (36) can be thought of as an average (multiplicative) price trend in the commodity i chain price relatives p_{it}/p_{it-1} around the general period $t-1$ to t inflation rates, $\exp[\pi_t]$, over all T periods in the sample; i.e., exponentiating both sides on the equation in (36) that corresponds to commodity i and period t and dropping the error term yields p_{it}/p_{it-1} , approximately equal to $\exp[\pi_t]$ times $\exp[\beta_i]$. Thus the specification (36) will capture constant commodity specific growth rates over the sample period in prices (in addition to the general growth in prices).

Note that the logarithm of the period $t-1$ to t inflation rate, π_t , is estimated by the right hand side of (40), which is identical to the right hand side of (22) if we set $s=t-1$ and use the same weights w_i in each formula.

Recall that $w_{it} \equiv p_{it}q_{it}/\sum_{n=1}^N p_{nt}q_{nt}$ is the i th expenditure share in period t . Clements and Izan

[1987; 342] make the following specification for the w_i which appear in (37) and (38):

$$(43) \quad w_i \equiv \sum_{t=0}^T w_{it} / (T+1); \quad i=1, \dots, N;$$

i.e., the w_i are the mean expenditure shares over the entire sample period.

Of course, since the w_i defined by (43)¹⁵ are not generally equal to the w_i defined by (24) when $s=t-1$, the Model 6 period $t-1$ to t inflation estimates $\hat{\pi}_t$ defined by (40) will not coincide precisely with the Model 4 estimates $\hat{\pi}_{t-1,t}$ defined by (22) when $s=t-1$. Thus Model 6 does not lead to a precise justification for the Törnqvist price index of Model 4, but Clements and Izan [1987; 343] argue that since the shares defined by (43) will not differ much from the shares defined by (24) when $s=t-1$, their specific price trends model provides an approximate justification for the use of the Persons-Törnqvist price index.

Clements and Izan [1987; 344-350] go on to show how variance estimates for the price indexes $\hat{\pi}_t$ defined by (40) can be derived. However, as in Model 4, the w_i defined by (43) depend on the prices p_{it} and hence the "fixed" weights w_i which appear in (37) and (38) are not really independent of the price relatives p_{it}/p_{it-1} . Hence the applicability of Model 6 when the w_i are defined by (43) is in doubt.¹⁶

This completes our review of the new stochastic approaches to index number theory. In the following two sections, we subject these approaches to a critical appraisal.

5. A Formulation of Edgeworth's Stochastic Approach to Index Numbers

The new stochastic models presented in the previous two sections suffer from a rather major defect: the variance assumptions of the type $\text{Var } \varepsilon_{it} = \sigma_i^2/w_i$ where w_i is an observed expenditure share of some sort are simply not supported empirically. Clements and Izan [1987; 345] note explicitly that their variance assumptions (37) and (43) are not supported by their empirical example.¹⁷ However, formal statistical tests are not required to support the common observation that

the food and energy components of the consumer price index are more volatile than many of the remaining components. Food has a big share while energy has a small share -- volatility of price components is simply not highly correlated with the corresponding expenditure shares.

The observation that different price components have widely differing degrees of volatility dates back to the origins of index number theory. For example, Edgeworth [1887; 244] observed:

Cotton and Iron, for example, fluctuate in this sense much more than Pepper and Cloves.

Later, Edgeworth [1918; 186] commenting on Mitchell's work observed:

...that the fluctuation in price from year to year is much greater for some kinds of commodities than for others... Thus manufactured goods are steadier than raw materials. There are characteristic differences among the price fluctuations of the groups consisting of mineral products, forest products, animal products, and farm crops. Again, consumers' goods are steadier in price than producers' goods, the demand for the farmer being less influenced by vicissitudes in business conditions.

For a summary of Mitchell's evidence on the variability of different components of US wholesale prices over the years 1890-1913, see Mitchell [1921; 40-43]. Finally, Mills [1927; 46] summarizes his evidence on the monthly variability of commodity prices as follows:

It is clear from Table 4 that individual commodities differ materially in the matter of price variability and, also, that the variability of specific commodities has changed from period to period.

In the light of the above criticism of Models 3 through 6, let us reconsider the classical stochastic models presented in section 2. However, instead of assuming that the period 5 residuals have a common variance, we now assume that the log of each chain commodity price relative, $\ln[p_{it}/p_{it-1}]$, after adjusting for a common period t inflation factor π_t , has its own commodity specific variance σ_i^2 . Thus Model 7 is defined by the following equations:

$$(44) \quad \ln[p_{it}/p_{it-1}] = \pi_t + \varepsilon_{it}; \quad i=1, \dots, N; t=1, \dots, T;$$

$$(45) \quad E \varepsilon_{it} = 0; \text{Var } \varepsilon_{it} = \sigma_i^2; \quad i=1, \dots, N; t=1, \dots, T.$$

The parameter π_t is the logarithm of the period $t-1$ to t price index for $t=1, \dots, T$ and for $i=1, \dots, N$.

The parameter σ_i^2 is the variance of the inflation adjusted logarithmic price ratios $\ln[p_{it}/p_{i,t-1}] - \pi_t$ for $t=1, \dots, T$.

It is interesting to note that a model similar to that defined by (44) and (45) was first vaguely suggested by Edgeworth as the following quotations indicate:

A third principle is that less weight should be attached to observations belonging to a class which are subject to a wider deviation from the mean.

Edgeworth [1887; 224].

Or, if more weight attaches to a change of price in one article rather than another, it is not on account of the importance of that article to the consumer or to the shopkeeper, but on account of its importance to the calculator of probabilities, as affording an observation which is peculiarly likely to be correct...

Edgeworth [1889; 287].

In combination of these values derived from observation, less weight should be attached to one belonging to a class which is subject to a wider deviation from the mean, for which the mean square of deviation is greater.

Edgeworth [1923; 574].

The term may include weighting according to 'precision' in the sense in which that term is attributed to errors of observation; a sense in which the price of pepper might deserve more weight than that of cotton, as M. Lucien March has the courage to maintain.

Edgeworth [1925; 383].

In the last quotation, Edgeworth is referring to March [1921; 81] who endorsed Edgeworth.¹⁸

Irving Fisher summarized Edgeworth's rather vague suggestions efficiently as follows:

Professor Edgeworth has made somewhat analogous, though less definite, proposals. He suggests that any commodity belonging to a class which is subject to wide scattering is a less reliable indicator than one belonging to a class not so subject. To take account of such differences in reliability he suggests that weights be assigned to each commodity in inverse proportion to the square of some variability-measure of the class to which it belongs.

This idea is scarcely capable of specific application, partly because the classification of commodities is so arbitrary and multiform, partly because of the difficulty of calculating any useful variability-measure for each class when determined. I wish Professor Edgeworth would take my 36 commodities, assign each to what he believes is its proper class, estimate each class-variability-measure, and calculate an index number accordingly.

Fisher [1922; 380].

We now show how estimators for our neo-Edgeworthian model defined by (44) and (45) can be obtained. The log of the likelihood function corresponding to Model 7 is, apart from inessential constants,

$$(46) \quad L(\pi_1, \dots, \pi_T; \sigma_1^2, \dots, \sigma_N^2) \equiv -\sum_{i=1}^N T \ln \sigma_i^2 - \sum_{i=1}^N \sum_{t=1}^T \sigma_i^2 [\ln(p_{it}/p_{it-1}) - \pi_t]^2.$$

Differentiating (46) with respect to the parameters and setting the resulting partial derivatives equal to 0 leads to the following system of $T+N$ simultaneous non linear equations to determine the maximum likelihood estimators for Model 7 (assuming that the $\hat{\sigma}_i^2$ are all strictly positive):

$$(47) \quad \hat{\pi}_t = \sum_{i=1}^N [1/\hat{\sigma}_i^2] \ln(p_{it}/p_{it-1}) / \sum_{i=1}^N [1/\hat{\sigma}_i^2]; \quad t=1, \dots, T;$$

$$(48) \quad \hat{\sigma}_i^2 = [1/T] \sum_{t=1}^T [\ln(p_{it}/p_{it-1}) - \hat{\pi}_t]^2 \quad i=1, \dots, N.$$

The interpretation of the specific commodity price variance estimators $\hat{\sigma}_i^2$ defined by (48) is straightforward. Equation t in (47) says that the estimator for the logarithm of the period $t-1$ to t inflation rate, π_t , is a weighted average of the individual period $t-1$ to t log price changes, $\ln[p_{it}/p_{it-1}]$, with the weight for the i th log price change being inversely proportional to its estimated variance, $\hat{\sigma}_i^2$. Thus Model 7 seems to capture the essence of Edgeworth's suggested stochastic approach to index number theory.

There can be at most one finite solution to equations (47) and (48) that has all $\hat{\sigma}_i^2$ strictly positive. A suggested algorithm for finding this solution if it exists is the following one. Begin iteration 1 by estimating $\hat{\pi}_t$ as the mean of the unweighted log price changes:

$$(49) \quad \hat{\pi}_t^{(1)} \equiv \sum_{i=1}^N (1/N) \ln(p_{it}/p_{it-1}); \quad t=1, \dots, T.$$

Thus $\exp[\hat{\pi}_t^{(1)}]$ is the Jevons geometric mean price index for the $t-1$ to t price change. Once the $\hat{\pi}_t^{(1)}$ have been defined, define the iteration 1 variances $[\hat{\sigma}_i^{(1)}]^2$ by (48) replacing $\hat{\pi}_t$ by $\hat{\pi}_t^{(1)}$. At the first stage of iteration 2, define the $\hat{\pi}_t^{(2)}$ by (47) using the iteration 1 $[\hat{\sigma}_i^{(1)}]^2$ in the right hand sides of (47). At the second stage of iteration 2, define the $[\hat{\sigma}_i^{(2)}]^2$ by (48) using the $\hat{\pi}_t^{(1)}$ in the right hand sides of (48). Now carry on repeating these first and second stage iterations until the estimates converge.

It can be shown that if the $\hat{\sigma}_i^{(k)}$ remain positive, then each stage of each iteration will lead to a strict increase in the log likelihood function (46) until convergence has been achieved.

Unfortunately, the above algorithm may not always work in degenerate cases. For example, consider the case where the period t prices are proportional to the base period prices for all t . In this case, the $\hat{\pi}_t$ are explicit functions of the proportionality factors and all of the commodity variances defined by (48) will be 0. There are other problems as well: if we pick any i and define $\hat{\pi}_t \equiv \ln[p_{it}/p_{it-1}]$ for $t=1, \dots, T$ and let σ_i^2 tend to 0 (with the other σ_j^2 positive and finite), we find that the log likelihood function approaches plus infinity. To rule out degenerate solutions of this type, it may be necessary to add a positive lower bound to the admissible variances in our model; i.e., we may need to add to (44) and (45) the following restrictions:

$$(50) \quad \sigma_i^2 \geq \sigma^2 > 0; i=1, \dots, N$$

for some σ^2 chosen a priori.

In any case, we now turn to a critical evaluation of these new stochastic models for price indexes.

6. A Critical Evaluation of the New Stochastic Approaches

Our first criticism of the new stochastic models presented in sections 3 and 4 has already been made: the variance assumptions made in these models are not consistent with the observed behavior of prices. This is a very fundamental criticism which has not been addressed by the proponents of these new models. Hence the assertion of Selvanathan and Prasada Rao [1994; 6] that their stochastic approach has provided standard errors for several well known index number formulae is correct only if their stochastic assumptions are correct, which seems very unlikely!

Our second criticism is directed towards the specific price trend models of Balk [1980], Clements and Izan [1987] and Selvanathan and Prasada Rao [1994; 63-66]: these models force the

same weights w_i to serve two distinct purposes and it is unlikely that their choice of weights could be correct for both purposes. In particular, their expenditure based weights are unlikely to be correct for the first purpose (which is criticism 1 again).

Our third criticism of the new stochastic approaches presented in sections 4 and 5 is that the resulting price indexes are not invariant to the number of periods T in the sample. Balk [1980; 72-73] was very concerned with this problem (since he works in a Statistical Agency and hence must suggest "practical" solutions to problems) and he presented some evidence on the stability of his estimated index numbers as T was increased. His evidence indicates that our third criticism is empirically important. Due to the fact that variances of price relatives can change considerably over time (recall Mills [1927; 46]), our neo-Edgeworthian Model 7 presented in the previous section will be particularly subject to this instability criticism.

The above invariance problem also occurs in the multilateral context and in the multiperiod time series context when we want our estimated index numbers to satisfy the circularity test; i.e., to be transitive. Walsh, after noting how multilateral transitivity can be achieved by using weights that pertain to all of the periods in the sample (e.g., recall equations (43) in Model 6), draws attention to the above invariance problem and also notes why the multilateral case is more difficult than the bilateral case:

In no case is this remedy satisfactory, for two principle reasons: (1) Because the present epoch is extending every year, requiring recalculations; and it does not appear that a later recalculation will be more correct than an earlier. Besides, how is a past variation between two years several years ago to be affected by present variations? (2) Because we really do not know how to calculate weights, or to determine equivalence of mass-units, or to average mass-quantities, over more than two periods, since the geometric average loses its virtue when applied to more than two figures.

Walsh [1901; 399].

Our fourth criticism of the new stochastic approaches is simple a restatement of the fundamental objection of Keynes:

The hypothetical change in the price level, which would have occurred if there had been no changes in relative prices, is no longer relevant if relative prices have in fact changed -- for the change in relative prices has in itself affected the price level.

I conclude, therefore, that the unweighted (or rather the randomly weighted) index number of prices -- Edgeworth's 'indefinite' index number -- ...has no place whatever in a rightly conceived discussion of the problems of price levels.

Keynes [1930; 78].

Thus if price relatives are different, then an appropriate definition of average price change cannot be determined independently of the economic importance of the corresponding goods. What is an appropriate definition of aggregate price change? Earlier in his book, Keynes [1930; 59-61] indicated that the price relatives in a producer or consumer price index should be weighted according to their relative importance as indicated by a census of production or by a consumer budget study. Thus the best index number formula according to Keynes is an expenditure weighted sum of relative prices; i.e., the price relatives must be weighted according to their economic importance, not according to their statistical importance, *à la* Edgeworth.¹⁹ Of course, in the approach advocated by Keynes, there is still the problem of choosing the "best" economic weights (base or current period expenditure shares or a mixture of them), but precise answers to this question simply lead back to the test or economic approaches to index number theory.

Criticism four can be restated as follows. The early statistical approaches of Jevons and Edgeworth (see section 2 above) treated each price relative as an equally valid signal of the general inflation rate: the price relative for pepper is given the same weight as the price relative for bread. This does not seem reasonable to "Keynesians" if the quantity of pepper consumed is negligible.

Another more technical way of restating the Keynesian objection to stochastic approaches can be accomplished by drawing on the models presented in section 5: if we make more reasonable variance assumptions, models of the form (36)-(38) are reasonable, except that the constant β_i 's should be replaced by sets of period specific β_{it} 's. But then the resulting model has too many

parameters to be identified.

Our conclusion at this stage is: in the present context where all prices and quantities are known without sampling error, signal extraction approaches to index number theory should be approached with some degree of caution.²⁰

Of course, there is a huge role for statistical approaches to index numbers when we change our terms of reference and assume that the given price and quantity data are only samples. The founders of the test approach, Walsh [1924; 516-517] and Fisher [1922; 336-340], did not deny a strong role for statistical techniques in the sampling context. In addition to the work of Bowley [1901] [1911] [1919] [1926] [1928], more recent references on the sampling aspects of price indexes include Mudgett [1951; 51-54], Adelman [1958], McCarthy [1961], Kott [1984] and the BLS [1988].

7. Other Approaches to the Determination of the Precision of an Index

Having rejected the new stochastic approaches to index number theory (when all prices and quantities are known with certainty over the sample period), we have to admit that the proponents of these new approaches have a point: if all of the price relatives pertaining to two periods are identical, it must be the case that the "precision" of the index number computation for those two periods is greater than when the price relatives are widely dispersed. On the other hand, the proponents of the test and economic approaches to index number theory use their favorite index number formula and thus provide a precise answer whether the price relatives are widely dispersed or not. Thus the test and economic approaches give a false sense of precision.

The early pioneers of the test approach addressed the above criticism. Their method works as follows: (i) decide on a list of desirable tests that an index number formula should satisfy; (ii) find some specific formulae that satisfy these tests (if possible); (iii) evaluate the chosen formulae with the data on hand and (iv) table some measure of the dispersion of the resulting index number

computations (usually the range or standard deviation was chosen). The resulting measure of dispersion can be regarded as a measure of functional form error.

Fisher [1922; 226-229] applied the above method to address the charge that the test approach gave a false precision to index numbers. He found 13 index number formulae (including the ideal) that satisfied the commodity, time and factor reversal tests and were not "freakish"; i.e., descended from modes or medians (and hence discontinuous). Fisher [1922; 227] found that the standard deviations between his 13 best fixed base indexes increased as the two periods being compared grew further apart; his "probable error" reached a maximum of about .1% when his 13 indexes were compared between 1913 and 1918. Fisher called this functional form error, instrumental error. In response to outraged criticisms from Bowley, Fisher later summarized his results as follows:

What I do claim to have demonstrated is something quite different, namely, that the 'instrumental' error, i.e., that part of the total error which may be ascribed to any inaccuracy in the mathematical formula used, is, in the case of the ideal formula (and, in fact, in the case of a score of other formulae as well), usually less than one part in 1000.

Fisher [1923; 248].

Warren Persons [1928; 19-23] also implemented the above test approach to the determination of functional form error. Persons looked for index number formulae which satisfied the time reversal test and his new test, the absence of weight correlation bias test. He found nine admissible index number formulae (including the Persons-Törnqvist and the Fisher ideal) and used Fisher's [1922] data to numerically evaluate these nine formulae. Finally, Persons [1928; 23] tabled the range of the resulting indexes over the sample period; he found that the range was a maximum in 1917 when it slightly exceeded 1%. It turned out that indexes satisfying Fisher's tests had a narrower range of dispersion than the indexes satisfying Persons' tests for the same data set.

Walsh [1921; 97-107] almost recommended the above approach to functional form error. He chose six index number formulae on the basis of how close they came empirically to satisfying

his multiperiod identity test.²¹ Walsh [1921; 106] used a small but somewhat extreme data set from Bowley [1901; 226] to evaluate his six index number formulae; he found that their range was about 2%. However, Walsh did not stop at this point; he went on to choose a single best index number formula:

To return to theory: would anything be gained by drawing an average of the results yielded by several methods? Hardly, as they have different merits. All that we can do is choose the best, after testing all the candidates; for to average the others with the best, would only vitiate the result.

Walsh [1921; 106-107].

What was Walsh's [1921; 102] theoretically best index number formula? None other than Irving Fisher's [1922] ideal index!²²

It is clear that there are some problems in implementing the above test approach to the determination of functional form error; i.e., what tests should we use and how many index number formulae should be evaluated in order to calculate the measure of dispersion? However, it is interesting to note that virtually all of the above index number formulae suggested by Fisher, Persons and Walsh approximate each other to the second order around an equal price and quantity point.²³

The above approach may be used to estimate the functional form error that arises from choosing an index number formula that is based on the economic approach. The economic approach recommends the use of a superlative index number formula, such as the Fisher-Walsh ideal formula²⁴ or the Persons-Törnqvist formula²⁵ or the direct and implicit quadratic mean of order r families of price indexes which include two indexes recommended by Walsh [1901; 105].²⁶ Many of these superlative indexes appear in the list of best test approach index number formulae recommended by Fisher, Persons and Walsh.²⁷ As was done for the test approach, the functional form error involved in using any specific superlative index could be approximated by evaluating a number of superlative indexes and then tabling a measure of their dispersion.

A specific proposal to measure the dispersion of superlative indexes is the following one.

Choose the following members of Diewert's [1976; 131] quadratic mean of order r price indexes P_r : P_2 (the Fisher-Walsh ideal price index), P_1 (Walsh), P_0 (Persons-Törnqvist), and P_{-2} . Choose the following members from Diewert's [1976; 132] implicit quadratic mean of order r price indexes \bar{P}_r : \bar{P}_1 (implicit Walsh), \bar{P}_0 (implicit Törnqvist) and \bar{P}_{-2} .²⁸ These formulae include the most frequently used superlative indexes. To measure the dispersion of these indexes, consider the following dispersion measure D , which is the range of the seven indexes divided by the minimum index:

$$(51) \quad D(p^s, p^t, q^s, q^t) \equiv [\max\{P_2, P_1, P_0, P_{-2}, \bar{P}_1, \bar{P}_0, \bar{P}_{-2}\} / \min\{P_2, P_1, P_0, P_{-2}, \bar{P}_1, \bar{P}_0, \bar{P}_{-2}\}] - 1$$

where $P_i \equiv P_i(p^s, p^t, q^s, q^t)$ and $\bar{P}_j \equiv \bar{P}_j(p^s, p^t, q^s, q^t)$. D can be interpreted as the percentage difference between the highest and lowest price indexes in the set of admissible indexes.

Note that $D(p^s, p^t, q^s, q^t) \geq 0$. Moreover, since each of the seven indexes that appear on the right hand side of (1) satisfy the Fisher [1911; 534] [1922; 64]-Walsh [1901; 368] time reversal test:

$$(52) \quad P(p^s, p^t, q^s, q^t) = 1/P(p^t, p^s, q^s, q^t),$$

it can be verified that the dispersion measures defined by (51) will satisfy the following base period invariance property:

$$(53) \quad D(p^s, p^t, q^s, q^t) = D(p^t, p^s, q^s, q^t);$$

i.e., if we interchange periods, the dispersion remains unchanged.

The dispersion measure defined by (51) can be adapted to the test approach: the set of index number formulae that would appear in (51) would be restricted to formulae that satisfied the appropriate set of tests. In particular, assume that the admissible P satisfy the time reversal test (52) and Walsh's [1901; 385] strong proportionality test:

$$(54) \quad P(p^s, \lambda p^s, q^s, q^t) = \lambda \text{ for } \lambda > 0;$$

i.e., if the period t price vector p^t is proportional to the period s price vector p^s , then the price index equals the common proportional factor. Under these hypotheses on the class of admissible price indexes in (51), the dispersion measure defined by the appropriate version of (51) would satisfy the

base period invariance test (53) and it would equal 0 if all of the price relatives were identical.

Returning to the economic approach to index numbers and the specific measure of formula error defined by (51), it can be verified that if both prices and quantities are proportional during the two periods under consideration, so that $p^t = \alpha p^s$ and $q^t = \beta q^s$ for some $\alpha > 0$, $\beta > 0$, then each of the seven indexes which appears in the right hand side of (51) is equal to α and hence the dispersion measure $D(p^s, \alpha p^s, q^s, \beta q^s)$ will attain its lower bound of 0. However, if only prices are proportional, then $D(p^s, \alpha p^s, q^s, q^t)$ will not necessarily equal 0. If we want a measure of dispersion that will equal zero when only prices are proportional, a different approach is required, which we now turn to.

A more direct approach to the reliability of a price index, $P(p^s, p^t, q^s, q^t)$, is to simply look at the variability of the individual price relatives, p_{it}/p_{is} , around the index number "average" value, $P(p^s, p^t, q^s, q^t)$. In order to implement this approach, define the i th absolute deviation by

$$(55) \quad d_i(p^s, p^t, q^s, q^t) \equiv |(p_{it}/p_{is}) - P(p^s, p^t, q^s, q^t)|, \quad i=1, \dots, N.$$

A measure of relative price variability, V , could be defined as an appropriate function of the deviations d_i defined by (55), say:

$$(56) \quad V(p^s, p^t, q^s, q^t) \equiv M[d_1(p^s, p^t, q^s, q^t), \dots, d_N(p^s, p^t, q^s, q^t)]$$

where M is a linearly homogeneous symmetric mean.²⁹

A desirable property for a price variability measure V is that it satisfy the base period invariance property (53) (where we replace D by V). Unfortunately, the V defined by (56) and (55) will not generally have this property.

In order to obtain a base period invariant measure of price variability between two periods, define the i th absolute logarithmic deviation e_i by

$$(57) \quad e_i(p^s, p^t, q^s, q^t) \equiv |\ln(p_{it}/p_{is}) - \ln P(p^s, p^t, q^s, q^t)|, \quad i=1, \dots, N.$$

Define a logarithmic price variability measure V by

$$(58) \quad V(p^s, p^t, q^s, q^t) \equiv M[e_1(p^s, p^t, q^s, q^t), \dots, e_N(p^s, p^t, q^s, q^t)]$$

where again M is a homogeneous symmetric mean. If the index number formula P satisfies the time reversal test (52), then it can be verified that $e_i(p^s, p^t, q^s, q^t) = e_i(p^t, p^s, q^t, q^s)$ and hence the V defined by (58) satisfies the base period invariance property (53) (with V replacing D).

Define the mean of order r of N positive numbers x_1, \dots, x_N for $r \neq 0$ by³⁰

$$(59) \quad M_r(x_1, \dots, x_N) \equiv [\sum_{i=1}^N (1/N)x_i^r]^{1/r}.$$

The means of order r , M_r , are homogeneous symmetric means and hence can be used as M 's in (58). For example, if we choose $r=2$ and substitute M_2 into (58), we obtain the following logarithmic price variability measure:

$$(60) \quad V_2(p^s, p^t, q^s, q^t) \equiv [(1/N)\sum_{i=1}^N \{\ln(p_{it}/p_{is}) - \ln P(p^s, p^t, q^s, q^t)\}^2]^{1/2}.$$

Note that (60) bears some resemblance to the earlier stochastic measure of reliability, $\hat{\sigma}_{st}$ defined by the square root of (23). It should also be noted that a monotonic transformation of the measure of relative price variability defined by (60), $N[V_2(p^s, p^t, q^s, q^t)]^2$, was suggested as a measure of the nonproportionality of prices by Allen and Diewert [1981; 433]: the price index P that they used in (60) was the Jevons equally weighted geometric mean defined by the right hand side of (12) (with p^s replacing p^0).

Unfortunately, the measures of price variability defined by (58) or (60) are still not satisfactory in the present context. The problem is that some price relatives are completely unimportant and hence should not be given the same weight as items that are important in the budgets of the consumer or producer for the two periods under consideration: recall Edgeworth and March's discussion about the relative importance of pepper versus wheat or cotton. We could use the budget shares of period s , w_{is} , or the budget shares of period t , w_{it} , as weights, but it seems less arbitrary to use an even handed average of these two sets of weights.³¹ Thus we will weight the i th absolute logarithmic price deviation e_i defined by (57) by $m(w_{is}, w_{it})$, where m is a linearly homogeneous symmetric mean of two variables. Note that the symmetry property of m implies that

$$(61) \quad m(w_{is}, w_{it}) = m(w_{it}, w_{is}), \quad i=1, \dots, N.$$

Thus our final class of price variability measures is defined as follows:

$$(62) \quad V(p^s, p^t, q^s, q^t) \equiv M[m(w_{1s}, w_{1t})e_1(p^s, p^t, q^s, q^t), \dots, m(w_{Ns}, w_{Nt})e_N(p^s, p^t, q^s, q^t)]$$

where the e_i are defined by (57) and M is again a homogeneous symmetric mean. If the price index P satisfies the time reversal test (52) and the share aggregator function m satisfies (61), then it can be verified that the V defined by (62) satisfies the base period invariance test (53). The V defined by (62) will also be nonnegative.³² Furthermore, if the price index P satisfies the strong proportionality test (54), then V will equal 0 if prices are proportional; i.e., $V(p^s, \lambda p^t, q^s, q^t) = 0$.

The most straightforward special case of (62) is obtained if we let M and m be means of order 1; i.e., arithmetic means. In this case, V becomes

$$(63) \quad V_1(p^s, p^t, q^s, q^t) \equiv \sum_{i=1}^N (1/N) \bar{w}_{ist} |\ln(p_{it}/p_{is}) - \ln P(p^s, p^t, q^s, q^t)|$$

where $\bar{w}_{ist} \equiv (1/2)(w_{is} + w_{it})$ is the average expenditure share on commodity i during periods s and t . The measure (63) is simply the arithmetic average of the weighted absolute logarithmic deviations, $\bar{w}_{ist} e_i(p^s, p^t, q^s, q^t)$. The only disadvantage of this measure is that it is not differentiable. A differentiable special case of (62) is obtained if we set $M = M_2$ and still let m be the arithmetic mean:

$$(64) \quad V_2(p^s, p^t, q^s, q^t) \equiv \left\{ \sum_{i=1}^N (1/N) \bar{w}_{ist} [\ln(p_{it}/p_{is}) - \ln P(p^s, p^t, q^s, q^t)]^2 \right\}^{1/2}.$$

Note the resemblance of (64) to the square root of (23). Comparing (64) to (63), V_2 gives larger weight to the larger weighted absolute logarithmic deviations, $\bar{w}_{ist} e_i(p^s, p^t, q^s, q^t)$. Both of the measures V_1 and V_2 will serve as satisfactory measures of variability or degree of nonproportionality of relative prices relative to the index number formula $P(p^s, p^t, q^s, q^t)$.

There is another approach to the measurement of relative price variability that has the advantage that it is simultaneously a measure of relative quantity variability. Consider the following variability measure due to Robert Hill [1995; 81]³³:

$$(65) \quad V_H(p^s, p^t, q^s, q^t) \equiv |\ln(p^s \cdot q^t \cdot p^t \cdot q^s / p^t \cdot q^t \cdot p^s \cdot q^s)| \geq 0$$

$$(66) \quad \equiv |\ell_{\ln}(P_L/P_P)|$$

$$(67) \quad \equiv |\ell_{\ln}(Q_L/Q_P)|$$

where $P_L \equiv p^t \cdot q^s/p^s \cdot q^t$ and $P_P \equiv p^t \cdot q^t/p^s \cdot q^t$ are the Laspeyres and Paasche price indexes and $Q_L \equiv p^s \cdot q^t/p^s \cdot q^s$ and $Q_P \equiv p^t \cdot q^t/p^t \cdot q^s$ are the Laspeyres and Paasche quantity indexes. Equation (66) shows that the variability measure defined by (65) can be written as the absolute value of the log of the ratio of the Laspeyres and Paasche price indexes while (67) shows a similar equality involving the ratio of the Laspeyres and Paasche quantity indexes. Thus if prices in the two periods are proportional (so that $p^t = \alpha p^s$), then $P_L = P_P = \alpha$ and using (66), $V_H = 0$. Similarly, if quantities in the two periods are proportional (so that $q^t = \beta q^s$), then $Q_L = Q_P = \beta$ and using (67), $V_H = 0$. Hence as Hill [1995; 81] observed, if either prices are proportional (recall Hicks' [1946; 312-313] Aggregation Theorem) or quantities are proportional (recall Leontief's [1936; 54-57] Aggregation Theorem), then the variability measure V_H defined by (65) attains its lower bound of 0. Note also that if we interchange periods, V_H remains unchanged; i.e., it satisfies the base period invariance property (53).

If x is close to 1, then $\ln x$ can be closely approximated by the first order approximation, $x-1$.

Hence the Hill variability measure V_H can be approximated by the following variability measure:

$$(68) \quad \bar{V}(p^s, p^t, q^s, q^t) \equiv |(P_L/P_P) - 1| = |(Q_L/Q_P) - 1|.$$

The variability measure \bar{V} has the same mathematical properties that were noted for V_H . Both measures are base period invariant measures of the spread between the Paasche and Laspeyres price (or quantity) indexes; both measures are approximately equal to the absolute value of the percentage difference between the Paasche and Laspeyres indexes. From the viewpoint of the test approach to index numbers, Bowley [1901; 227], Fisher [1922; 403] and Diewert [1992a; 219-220] proposed that the price index P should be between the Paasche and Laspeyres price indexes. These bounds are also valid from the economic point of view if we have a homothetic or linearly homogeneous

aggregator function. Thus the variability measures defined by (65) and (68) provide convenient methods of describing the width of these index number bounds.

Note that the variability or nonproportionality measures V_H and V do not depend on a particular index number formula P . However, if the index number formula P is a symmetric mean of the Paasche and Laspeyres indexes (e.g., $P = (P_L P_P)^{1/2}$, the Fisher Walsh ideal index), then P will lie between P_L and P_P and V_H or \tilde{V} may be used as reliability measures for P .

We have presented three classes of dispersion measures (see (51), (62) and (65) or (68) above) that could be used to measure the reliability of an index number formula. The use of (62), (65) or (68) as measures of dispersion would meet some of the criticisms of the test and economic approaches that have been made by the proponents of the stochastic approach. If all of the relative prices were identical, the above dispersion measures would attain their lower bounds of zero, but if the price relatives were dispersed, nonzero measures of dispersion or variability would be obtained if (51) or (62) were used.

It is now almost 75 years after Walsh [1921] made his comments on the diversity of approaches to index number theory and economists are still "at loggerheads." However, perhaps this diversity is a good thing. The new stochastic approach to index numbers has at least caused this proponent of the test and economic approaches to think more deeply about the foundations of the subject.

FOOTNOTES

- * Professor of Economics at the University of British Columbia and Research Associate at the National Bureau of Economic Research. This research was supported by a Strategic Grant from the Social Sciences and Humanities Research Council of Canada. Thanks are due to Keltie Stearman for typing a difficult manuscript.
1. This term is due to Frisch [1936; 3-4].
 2. See Selvanathan and Prasada Rao [1994; 6] for an extensive list of their recent contributions.
 3. A flexible functional form is one that has a second order approximation property; see Diewert [1974; 115].
 4. See Diewert [1976; 117].
 5. Selvanathan and Prasada Rao [1994; 15-44] provide a rather inadequate review of the test and economic approaches. For example on page 17, they attribute Walsh's [1901] [1921; 97] price index to Drobisch, they misspell Marshall and they cite an incorrect reference to Marshall [1887], the cofounder of the Edgeworth-Marshall index.
 6. Edgeworth [1887; 237-243] argued on empirical and logical grounds that Model 2 was more plausible than Model 1, assuming normally distributed errors. His logical argument was based on the positivity of prices; hence a price relative could have any upper bound but had a definite lower bound of zero, leading to an asymmetric distribution of price relatives. However, the logarithm of a price relative could be symmetrically distributed.
 7. Mills [1927; 240-247] succinctly reviewed the above literature and also computed standard errors for various index number formulae using BLS data on US wholesale prices.
 8. These authors choose period s to be period $t-1$ but this choice is not essential to their argument.
 9. This index first appeared as formula 123 in Fisher [1922; 473]. Fisher [1922; 265] listed

it as number 15 in his list of the 29 best formulae, but he did not otherwise distinguish it. Walsh [1921; 97] almost recommended (25), but he used the geometric average of the weight, $(w_{is}w_{it})^{1/2}$, in place of the arithmetic average. Finally, Persons [1928; 21-22] recommended (25), the Fisher ideal index, $(p^t \cdot q^t p^s / p^s \cdot q^t p^s \cdot q^s)^{1/2}$, and seven other indexes as being the best from the viewpoint of his test approach. Thus (25) should perhaps be known as the Persons-Törnqvist formula.

10. The first person to make a variance specification of this form appears to have been Edgeworth [1887; 247] as the following quotation indicates: "Each price which enters into our formula is to be regarded as the mean of several prices, which vary with the differences of time, of place, and of quality; by the mere friction of the market, and, in the case of 'declared values', through errors of estimation, it is reasonable to support that this heterogeneity is greater the larger the volume of transactions." Edgeworth did not make any formal use of these observations.
11. "Even in the case where prices of all the commodities of relevance are measured, and measured without any errors, the question of reliability of a given index arises." (Selvanathan and Prasada Rao [1994; 4]).
12. The reader will deduce that, in the interests of a homogeneous presentation, I have modified the original notation of Clements and Izan and Selvanathan and Prasada Rao.
13. In his new stochastic model, Balk [1980; 72] simply assumed a variance specification analogous to (13) or (21) without any justification other than mathematical convenience.
14. It is immediately evident that the specification (29) is not very satisfactory. As we go from period 0 to 1, it is reasonable to postulate that β_1 is the systematic part of the commodity 1 price change p_{11}/p_{10} in addition to the general period 0 to 1 price change α_1 but it is not reasonable to assume that this same β_1 will characterize the systematic part of the commodity

1 relative price changes p_{it}/p_{i0} for later periods, $t=2,3,\dots,T$, since as t increases, these fixed base systematic trends will tend to increase in magnitude.

15. It is interesting to note that Walsh [1901; 398] almost derived the transitive multilateral system of index numbers defined by (40) and (43): in place of the arithmetic means of the sample expenditure shares defined by (43), Walsh recommended the use of the corresponding geometric means. It should also be noted that Balk's [1980; 71] specialization of his seasonal model is a special case of Model 6 with w_i defined as $\sum_{t=0}^T P_{it} q_{it} / \sum_{s=0}^T \sum_{j=1}^N P_{js} q_{js}$.
16. Note that Model 5 when $w_i \equiv w_{i0}$ does not suffer from this difficulty. However, the interpretation of the β_i in Model 5 is more problematic.
17. "As can be seen, the variances are not inversely proportional to the budget shares as required by (16')." (Clements and Izan [1987; 345]).
18. March [1921; 81] observed that if the price of paper varied less than the price of wheat, then the former price should be given more weight in the index number formula.
19. Keynes' belief in the importance of economic weighting (as opposed to Edgeworth [1901; 410] and Bowley [1901; 219] who at times believed that weighting was unimportant) dates back at least to Keynes [1911; 46].
20. The dynamic factor index approach of Bryan and Cecchetti [1993; 19] is an example of a signal extraction approach to index numbers that we did not cover due to its complexity. Their approach is only subject to our criticisms 3 and 4. Their approach is also subject to a criticism that can be levelled against the specific price trend models of section 4: the nonstationary components of their specific price trends (their counterparts to the β_i which appear in Models 5 and 6 above) are assumed to be constant over the sample period.
21. Walsh [1921; 104] called his test the circular test but it is slightly different from the Westergaard-Fisher [1922; 413] circular test; see Diewert [1993b; 39].

22. Walsh [1901] [1921] was an originator of the test approach to index number theory and he also proposed the use of the ideal index either before Fisher [1921] or coincidentally. Perhaps the reason why Walsh has been forgotten but Fisher lives on is due to the rather opaque writing style of Walsh whereas Fisher wrote in a very clear style.
23. Thus these indexes are either superlative or pseudo-superlative; i.e., they approximate superlative indexes to the second order around an equal price and quantity point; see Diewert [1978; 896-898].
24. See Diewert [1976; 134].
25. See Diewert [1976; 121].
26. See Diewert [1976; 134-135]. The two Walsh indexes are obtained when we set $r=1$. Walsh [1921; 97] listed his two recommended indexes as formulae (5) and (6). The right hand side of (5) needs to be multiplied by the expenditure ratio for the two periods under consideration, since on the previous page, Walsh [1921; 96] assumed that these expenditures were equal.
27. On the basis of its consistency with revealed preference theory and its consistency with linear and Leontief aggregator functions, Diewert [1976; 137-138] recommended the Fisher-Walsh ideal index as the best superlative index number formula. Allen and Diewert [1981; 435] also endorse this index number formula as being the best superlative one since it is consistent with both Hicks' [1946; 312-313] Composite Commodity Theorem and Leontief's [1936; 54-57] Aggregation Theorem.
28. Fisher's [1922; 461-487] identification numbers for these formulae are: 353, 1153, 123, the geometric mean of 13 and 19, 1154, 124, and the geometric mean of 14 and 20.
29. A symmetric mean $M(x_1, \dots, x_N)$ is defined to be a continuous, symmetric increasing function of N real variables that has the mean value property, $M(\lambda, \dots, \lambda) = \lambda$. $M(x_1, \dots, x_N)$ will also

satisfy the following min-max property: $\min_i \{x_i\} \leq M(x_1, \dots, x_N) \leq \max_i \{x_i\}$. This last property and (55) imply that V will be nonnegative.

30. See Hardy, Littlewood and Polya [1934; 12].
31. Our reasoning is similar to that of Walsh [1921; 90], who made the case for the use of average weights in a price index as follows: "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 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."
32. If P satisfies the usual homogeneity properties with respect to prices and quantities (e.g., see tests PT5-PT8 in Diewert [1992a; 215-216]), then it can be shown that $V(p^s, p^t, q^s, q^t)$ will be homogeneous of degree zero in each of its four sets of variables.
33. Hill defined $V_H \equiv \ln[\max\{Q_L, Q_P\}/\min\{Q_L, Q_P\}]$. It can be shown that this definition is equivalent to (67).

REFERENCES

- Adelman, I. [1958], "A New Approach to the Construction of Index Numbers." Review of Economics and Statistics 40, 240-249.
- Afriat, S.N. [1972], "The Theory of International Comparisons of Real Income and Prices," pp. 13-69 in D.J. Daly (ed.), International Comparisons of Prices and Outputs, New York: Columbia University Press.
- Allen, R.C. and W.E. Diewert [1981], "Direct versus Implicit Superlative Index Number Formulae," The Review of Economics and Statistics 63, 430-435.
- Balk, B.M. [1980], "A Method for Constructing Price Indices for Seasonal Commodities," Journal of the Royal Statistical Society A, 143, 68-75.
- Bowley, A.L. [1901], Elements of Statistics, Westminster England: P.S. King and Son.
- Bowley, A.L. [1911], "The Measurement of the Accuracy of an Average," Journal of the Royal Statistical Society 75, 77-88.
- Bowley, A.L. [1919], "The Measurement of Changes in the Cost of Living," Journal of the Royal Statistical Society 82, 343-372.
- Bowley, A.L. [1926], "The Influence on the Precision of Index-Numbers of Correlation between the Prices of Commodities," Journal of the Royal Statistical Society 89, 300-319.
- Bowley, A.L. [1928], "Notes on Index Numbers," The Economic Journal 38, 216-237.
- Bryan, M.F. and S.G. Cecchetti [1993], "The Consumer Price Index as a Measure of Inflation," Economic Review, Federal Reserve Bank of Cleveland 29, 15-24.
- Bureau of Labor Statistics [1988], BLS Handbook of Methods, Bulletin 2285, Washington: U.S. Government Printing Office.

- Carli, Gian-Rinaldo [1764], "Del valore e della proporzione de' metalli monetati," reprinted in pp. 297-336 in Scrittori classici italiani di economia politica, Volume 13, Milano: G.G. Destefanis, 1804.
- Clements, K.W. and H.Y. Izan [1981], "A Note on Estimating Divisia Index Numbers," International Economics Review 22, 745-747.
- Clements, K.W. and H.Y. Izan [1987], "The Measurement of Inflation: A Stochastic Approach," Journal of Business and Economic Statistics 5, 339-350.
- Diewert, W.E. [1974], "Applications of Duality Theory," pp. 106-171 in M.D. Intriligator and D.A. Kendrick (eds.), Frontiers of Quantitative Economics, Vol. II, Amsterdam: North-Holland.
- Diewert, W.E. [1976], "Exact and Superlative Index Numbers," Journal of Econometrics 4, 115-145.
- Diewert, W.E. [1978], "Superlative Index Numbers and Consistency in Aggregation," Econometrica 46, 883-900.
- Diewert, W.E. [1992a], "Fisher Ideal Output, Input and Productivity Indexes Revisited," The Journal of Productivity Analysis 3, 211-248.
- Diewert, W.E. [1992b], "Exact and Superlative Welfare Change Indicators," Economic Inquiry 30, 565-582.
- Diewert, W.E. [1993a], "Overview of Volume 1," pp. 1-31 in Essays in Index Number Theory, Vol. 1, W.E. Diewert and A.O. Nakamura (eds.). Amsterdam: North-Holland.
- Diewert, W.E. [1993b], "The Early History of Price Index Research," pp. 33-65 in Essays in Index Number Theory, Vol. 1, W.E. Diewert and A.O. Nakamura (eds.). Amsterdam: North-Holland.

- Diewert, W.E. [1993c], "Symmetric Means and Choice Under Uncertainty," pp. 355-433 in Essays in Index Number Theory, Vol. 1, W.E. Diewert and A.O. Nakamura (eds.). Amsterdam: North-Holland.
- Edgeworth, F.Y. [1887], Measurement of Change in Value of Money I, the first Memorandum presented to the British Association for the Advancement of Science; reprinted as pp. 198-259 in Papers Relating to Political Economy, Vol. 1, New York: Burt Franklin, 1925.
- Edgeworth, F.Y. [1888a], Tests of Accurate Measurement, the second Memorandum presented to the British Association for the Advancement of Science; reprinted as pp. 304-343 in Papers Relating to Political Economy, Vol. 1, New York: Burt Franklin, 1925.
- Edgeworth, F.Y. [1888b], "Some New Methods of Measuring Variation in General Prices," Journal of the Royal Statistical Society 51, 346-368.
- Edgeworth, F.Y. [1889], Measurement of Change in Value of Money, the third Memorandum presented to the British Association for the Advancement of Science; reprinted as pp. 259-297 in Papers Relating to Political Economy, Vol. 1, New York: Burt Franklin, 1925.
- Edgeworth, F.Y. [1901], "Mr. Walsh on the Measurement of General Exchange Value," The Economic Journal 11, 404-416.
- Edgeworth, F.Y. [1918], "The Doctrine of Index-Numbers according to Professor Wesley Mitchell," The Economic Journal 28, 176-197.
- Edgeworth, F.Y. [1923], "Mr. Correa Walsh on the Calculation of Index-Numbers," Journal of the Royal Statistical Society 86, 570-590.
- Edgeworth, F.Y. [1925], "The Plurality of Index-Numbers," The Economic Journal 35, 379-388.
- Eichhorn, W. [1978], Functional Equations in Economics, Reading, Massachusetts: Addison-Wesley.
- Eichhorn, W. and J. Voeller [1976], Theory of the Price Index, Berlin: Springer-Verlag.

- Fisher, I. [1911], The Purchasing Power of Money, London: Macmillan.
- Fisher, I. [1921], "The Best Form of Index Number," Journal of the American Statistical Association 17, 533-537.
- Fisher, I. [1922], The Making of Index Numbers, Boston: Houghton Mifflin.
- Fisher, I. [1923], "Professor Bowley on Index-Numbers" (with a comment by Bowley), The Economic Journal 33, 246-252.
- Frisch, R. [1936], "Annual Survey of General Economic Theory: The Problem of Index Numbers," Econometrica 4, 1-39.
- Hardy, G.H., J.E. Littlewood and E. Polya [1934], Inequalities, London: Cambridge University Press.
- Hicks, J.R. [1946], Value and Capital, Second Edition, Oxford: Clarendon Press.
- Hill, R.J. [1995], "Purchasing Power Parity Methods of Making International Comparisons," Ph.D. Thesis, Economics Department, University of British Columbia, Vancouver, Canada.
- Jevons, W.S. [1863], "A Serious Fall in the Value of Gold Ascertained and its Social Effects Set Forth," reprinted as pp. 13-118 in Investigations in Currency and Finance, London: Macmillan and Co., 1884.
- 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 as pp. 119-150 in Investigations in Currency and Finance, London: Macmillan and Co., 1884.
- Jevons, W.S. [1869], "The Depreciation of Gold," Journal of the Statistical Society of London 32, 445-449; reprinted as pp. 151-159 in Investigations in Currency and Finance, London: Macmillan and Co., 1884.
- Konüs, A.A. [1924], "The Problem of the True Index of the Cost of Living," translated in Econometrica 7, (1939), 10-29.

- Konüs, A.A. and S.S. Byushgens [1926], "K. Probleme Pokupatelnoi Cili Deneg," Voprosi Konyunkturi II:1, 151-172.
- Keynes, J.M. [1911], "Comment on the Course of Prices at Home and Abroad, 1890-1910," Journal of the Royal Statistical Society 75, 45-47.
- Keynes, J.M. [1930], A Treatise on Money Vol. 1, New York: Harcourt, Brace and Company.
- Kott, P.S. [1984], "A Superpopulation Theory Approach to the Design of Price Index Estimators with Small Sampling Biases," Journal of Business and Economic Statistics 2, 83-90.
- Leontief, W. [1936], "Composite Commodities and the Problem of Index Numbers," Econometrica 4, 39-59.
- March, L. [1921], "Les modes de mesure du mouvement général des prix," Metron 1, 57-91.
- Marshall, A. [1887], "Remedies for Fluctuations of General Prices," Contemporary Review 51, 355-375.
- McCarthy, P.J. [1961], "Sampling Considerations in the Construction of Price Indexes with particular Reference to the United States Consumer Price Index," pp. 197-232 in The Price Statistics of the Federal Government, George J. Stigler, chairman of commission and editor, New York: National Bureau of Economic Research.
- Mills, F.C. [1927], The Behavior of Prices, New York: The National Bureau of Economic Research.
- Mitchell, W.C. [1921], "The Making and Using of Index Numbers," pp. 7-114 in Index Numbers of Wholesale Prices in the United States and Foreign Countries, Revision of Bulletin 173, Bureau of Labor Statistics, Washington D.C.: Government Printing Office.
- Mudgett, B.D. [1951], Index Numbers, New York: John Wiley and Sons.
- Persons, W.M. [1928], The Construction of Index Numbers, Cambridge, Massachusetts: The Riverside Press.

Pollak, R.A. [1989], The Theory of the Cost-of-Living Index, Oxford: Oxford University Press.

Samuelson, P.A. and S. Swamy [1974], "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," American Economic Review, 24, 566-597.

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

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

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 and Son.

Walsh, C.M. [1924], "Professor Edgeworth's Views on Index-Numbers," Quarterly Journal of Economics 38, 500-519.

Westergaard, H. [1890], Die Grundzüge der Theorie der Statistik, Jena: Gustav Fischer.