Harmonized Indexes of Consumer Prices: Their Conceptual Foundations

W. ERWIN DIEWERT*

JEL Classification: C43, C81, D11, D91, E31, E52, E58, O47.
Keywords: Inflation target, Consumer Price Index, Harmonized Index of Consumer Prices, the System of National Accounts, superlative indexes, fixed basket indexes, stochastic indexes, index number tests, cost of living indexes, tourist expenditures, elementary indexes, quality change, substitution bias, representativity bias, chained indexes, fixed base indexes, owner occupied housing, seasonal commodities, user cost, rental equivalence, acquisitions approach, unit values, intertemporal cost of living indexes, the temporary equilibrium, asset prices.

NONTECHNICAL SUMMARY

Many central banks have explicit inflation targets that they attempt to achieve. Section 2 looks at possible criteria for choosing an inflation index for purposes of monetary targeting.

Section 3 argues that the most useful target indexes will be associated with broad output flows that appear in the system of national accounts. Cases could be made for target price indexes that correspond to $C+I+G+X, C+I+G, C+I$ or C (but *not* $C+I+G+X-M$). In addition to a *primary target price index* like the Consumer Price Index, it will be useful for the central bank to have available (for monitoring purposes) *output* and *input price indexes* that apply to the private production sector. The output price index will be a comprehensive producer price index (at basic prices), which will weight (gross) output prices positively and domestic intermediate prices negatively. The input price index will be an aggregate of import prices, wage rates, user costs of reproducible capital and land and resource user costs. These two indexes are required to

* Department of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1., e-mail: diewert@econ.ubc.ca, website: http://www.econ.ubc.ca/diewert/hmpgdie.htm.
Paper presented at the Joint Meeting of the Centre for Economic Policy Research and the European Central Bank: *Issues in the Measurement of Price Indices*, Frankfurt, November 16–17, 2001. The author would like to thank the Social Sciences and Humanities Research Council of Canada for financial support and John Astin, Ludwig von Auer, Bert Balk, Carsten Boldsen Hansen, Steven Cecchetti, Yoel Finkel, Vitor Gaspar, Rósmundur Gudnason, Peter Hill, Johannes Hoffmann, Leendert Hoven, Pat Jackman, Jan Karlsson, Ulrich Kohli, Alice Nakamura, Masato Okamoto, Mick Silver, Jack Triplett, Mark Wynne, Sylvester Young and other participants at the Workshop for helpful comments and discussions. None of the above are to be held responsible for any opinions expressed in the paper.

deflate the value of outputs and the value of inputs respectively into measures of real output and real input and these latter measures in turn can be used to form productivity measures.

Section 4 looks at a price index for the components of household wealth as a possible target inflation index. However, household wealth does not seem to be as fundamental as consumption. Wealth is the nominal constraint in the consumer's intertemporal budget constraint and as such, the "price" of wealth has no real economic meaning. If we try to think about wealth in real terms, then the most natural thing to do is to deflate it by the consumer price index. However, since ALCHIAN and KLEIN made a theoretical argument advocating that central banks target a price index for wealth, some space is devoted to the limitations of their argument.

Once we have decided on a transactions domain of definition for the inflation index that the central bank is to target, it is still necessary to decide on an index number concept; i. e., how exactly are the many thousands of individual prices to be combined into an overall index? Thus in section 5, four alternative approaches to index number theory are considered. These approaches are: (i) fixed basket approaches; (ii) the test approach; (iii) the stochastic approach and (iv) the economic approach. It turns out that these four approaches to index number theory generate three index number formulae that turn out to be "best". Fortunately, these three formulae turn out to approximate each other very closely so that for many purposes, it is not necessary to pick any one of the four broad approaches to index number theory as being "best".

With all of the above introductory material in hand, section 6 looks at the properties of the Harmonized Index of Consumer Prices. A few methodological difficulties are found with the HICP. However, every Consumer Price Index faces certain methodological difficulties.

Section 7 takes a look at some of the more difficult measurement problems that arise when a statistical agency attempts to construct a Consumer Price Index. These problems include the treatment of quality change, substitution or representativity bias, chained versus fixed base indexes, the choice of formula at the lowest level of aggregation and the treatment of owner occupied housing and seasonal commodities. Tentative "solutions" to many of these problems are presented.

Section 8 concludes.

TABLE OF CONTENTS

1. INTRODUCTION

"The answer to the question what is the *Mean* of a given set of magnitudes cannot in general be found, unless there is given also the object for the sake of which a mean value is required." Francis. Y. Edgeworth (1888, p. 347)

"Thus, there are many cost of living index answers because there are many cost of living questions. Each question can be thought of in terms as a compensation for inflation, and each cost of living index is an answer that provides an appropriate measurement for some purpose." Jack E. Triplett (1983, p. 460).

The *Harmonized Index of Consumer Prices* (HICP) published by Eurostat, restricted to the 12 European Union countries who have adopted the Euro as their currency[1], is the index used in the European Central Bank's definition of price stability.[2] Loosely speaking, it is the *target price index* for the European Central Bank. But what is a "ideal" price index to target for inflation monitoring purposes? And how does the HICP stack up as an index that the central bank should use for inflation targeting purposes? We will attempt to provide tentative answers to these two questions.

What exactly is a central bank *target inflation index*? We discuss this question in more detail in section 2 below. It must be a *broad* or *general measure of price change* occurring between two periods. But what exactly is the *domain of definition* of an "inflation" index; i.e., over what set of economic agents or institutional units and over what set of commodities and transactions will the index be defined? We discuss possible answers to this question in sections 3 and 4 below where we use the structure of the System of National Accounts to provide possible transaction domains of definition. Once the domain of definition problem has been settled, there is also the problem of choosing an appropriate *index number concept* in order to implement the target index. This problem is discussed in section 5.

Having looked at the problem of choosing an inflation measure to target from an a priori theoretical point of view in earlier sections, in section 6 we look at the properties of Eurostat's Harmonized Index of Consumer Prices and compare these properties with the "ideal" index. In section 7, we look at some of the difficult measurement problems that both the HICP and an ideal inflation measure must deal with in practice.

Section 8 concludes.

1. This collection of countries is referred to as the Euro area. Eurostat also calculates HICP's for the three EU countries outside the monetary union (the UK, Denmark and Sweden) as well as for Iceland and Norway (who are not formal members of the EU).
2. "Upon announcement of the monetary policy strategy, the Governing Council decided to provide a quantitative definition of price stability in the Euro area: 'a year-on-year increase in the Harmonized Index of Consumer Prices of below 2 percent'. The use of the word 'increase' makes it clear that year-on-year falls in the HICP are inconsistent with the definition of price stability. The definition is therefore symmetric, in the sense that it excludes both negative and significantly positive rates of change in the price index." Issing (2001; p. 194–195).

2. CRITERIA FOR CHOOSING AN INFLATION INDEX FOR MONETARY POLICY PURPOSES

When we choose a price index for inflation targeting purposes, two important decisions have to be made:

- What is the set of *transactions* in the economy under consideration that should be in the target index domain of definition?
- Once the transactions domain of definition has been chosen, what *index number concept* should we choose for the price index that pertains to the chosen transactions?

Early "inflation" theories for the price index specified that the set of transactions that the price index should encompass is *the set of all monetary transactions that occurred in the economy in the two periods being compared.* This domain of definition for an "inflation" index dates back to IRVING FISHER[3] at least:

"Without attempting to construct index numbers which particular persons and classes might sometimes wish to take as standard, we shall merely inquire regarding the formation of such a general index number. It must, as has been pointed out, include all goods and services. But in what proportion shall these be weighted? How shall we decide how much weight should be given, in forming the index, to the *stock* of durable capital and how much weight to the *flow* of goods and services through a period of time, the flow to individuals, which mirrors consumption? The two things are incommensurable. Shall we count the railways of the country as equally important with a month's consumption of sugar, or with a year's?

To cut these Gordian knots, perhaps the best and most practical scheme is that which has been used in the explanation of the P in our equation of exchange [$MV = PT$], an index number in which every article and service is weighted according to the value of it *exchanged* at base prices in the year whose level of prices it is desired to find. By this means, goods bought for immediate consumption are included in the weighting, as are also all durable capital goods exchanged during the period covered by the index number. What is repaid in contracts so measured is the same *general purchasing power.* This includes purchasing power over everything purchased and purchasable, including real estate, securities, labor, other services, such as the services rendered by corporations, and commodities." IRVING FISHER (1911, p. 217–218).

Thus FISHER suggested that the inflation index should be a price index that applies *to all monetary transactions* that take place in the two periods under consideration.[4] However, under present economic conditions, this extremely broad definition of an "inflation" index is of limited use and rarely implemented due to the preponderance of trans-

3. FISHER'S (1911; p. 201) original choice of functional form for the price index in his equation of general purchasing power was the Paasche index.
4. When we speak of "monetary transactions", do we mean to include only transactions that are conducted using currency and checking accounts or do we want to include transactions conducted through broader monetary instruments such as credit cards?

actions in currency and stock market trading, which totally overwhelm other more interesting transactions.[5] Thus it is necessary to narrow the scope of "all monetary transactions" to a smaller domain of definition that encompasses transactions over a specified set of commodities and a specified set of transactors. Choosing the set of transactions to be covered by "the" inflation price index we term *the domain of definition (or scope of the index) problem*.[6]

In addition to choosing a domain of definition for the target inflation index, we also need to choose an *index number concept*. In order to do this, it will be useful to specify the nature of a price index a bit more formally as follows. First, as noted above, the domain of definition for a value aggregate V must be chosen. It refers to a certain set of transactions pertaining to a time period. We now consider the same value aggregate for two time periods, periods 0 and 1. For the sake of definiteness, we call period 0 *the base period* and period 1 *the current period* and we assume that we have collected observations on the base period price and quantity vectors, $p_0 \equiv [p_1^0, \dots, p_N^0]$ and $q^0 \equiv [q_1^0, \dots, q_N^0]$ respectively, as well as on the current period price and quantity vectors, $p^1 \equiv [p_1^1, \dots, p_N^1]$ and $q^1 \equiv [q_1^1, \dots, q_N^1]$ respectively.[7] The value aggregates in the two periods are defined in the obvious way as:

$$V^0 \equiv \sum_{i=1}^{N} p_i^0 q_i^0, \quad V^1 \equiv \sum_{i=1}^{N} p_i^1 q_i^1. \tag{1}$$

A *price index* pertaining to the specified value aggregate is defined as a function or measure which summarizes the *change* in the prices of the N commodities in the value aggregate from situation 0 to situation 1. More specifically, a *price index* $P(p^0, p^1, q^0, q^1)$ along

5. FISHER (1911; p. 225–226) noted that it would be difficult to obtain data for all transactions: "It is, of course, utterly impossible to secure data for all exchanges, nor would this be advisable. Only articles which are standardized, and only those the use of which remains through many years, are available and important enough to include. These specifications exclude real estate, and to some extent wages, retail prices, and securities, thus leaving practically nothing but wholesale prices of commodities to be included in the list of goods, the prices of which are to be compounded into an index number." FISHER (1911; p. 226–227) went on to argue that for the United States in the early years of the century, real estate transactions amounted "only to a fraction of 1 per cent of the total transactions", security transactions amounted to "about 8 per cent of the total transactions of the country", wages "amount to about 3 per cent and retail prices could be omitted "because wholesale and retail prices roughly correspond in their movements". Obviously, these rough approximations are no longer relevant; currency transactions alone account for about $1.4 trillion US dollars per trading day.
6. As we have seen FISHER (1911; p. 204–230) provided an extensive discussion of the domain of definition problem as did Knibbs (1924; p. 47–49), Pollak (1989), DIEWERT (1997) and TRIPLETT (2001; F320-F322). DIEWERT (1997; p. 134–136) discussed whether seasonal goods should be excluded from the household domain of definition, whether consumer durables should be excluded, whether future goods or savings should be included, whether leisure should be included, whether commodity taxes should be included, and whether commodities that have highly variable prices should be excluded.
7. Note that we are assuming that there are no new or disappearing commodities in the value aggregates.

with the corresponding *quantity index* (or *volume index*) $Q(p^0, p^1, q^0, q^1)$ are defined to be two functions of the 4N variables p^0, p^1, q^0, q^1 (these variables describe the prices and quantities pertaining to the value aggregate for periods 0 and 1) where these two functions satisfy the following equation:[8]

$$V^1/V^0 = P(p^0, p^1, q^0, q^1)Q(p^0, p^1, q^0, q^1). \tag{2}$$

If there is only one item in the value aggregate, then the price index P should collapse down to the single price ratio, p_1^1/p_1^0 and the quantity index Q should collapse down to the single quantity ratio, q_1^1/q_1^0. In the case of many items, the price index P is to be interpreted as some sort of weighted average of the individual price ratios, $p_1^1/p_1^0, \ldots, p_N^1/p_N^0$.

The above approach to index number theory shows that the index number problem can be regarded as the problem of *decomposing* the change in a value aggregate, V^1/V^0, into the product of a part that is due to *price change*, $P(p^0, p^1, q^0, q^1)$, and a part that is due to *quantity change*, $Q(p^0, p^1, q^0, q^1)$. This approach to the determination of the price index is the approach that is taken in the national accounts, where a price index is used to *deflate* a value ratio in order to obtain an estimate of quantity change. Thus in this approach to index number theory, the primary use for the price index is as a *deflator*. Note that once the functional form for the price index $P(p^0, p^1, q^0, q^1)$ is known, then the corresponding quantity or volume index $Q(p^0, p^1, q^0, q^1)$ is completely determined by $P(p^0, p^1, q^0, q^1)$. In section 5 below, we will consider in more detail the problems involved in choosing a specific functional form for the price index $P(p^0, p^1, q^0, q^1)$.

Obviously, we need some *criteria* for choosing both an appropriate domain of definition and index number concept for the central bank target index of inflation. Some possible criteria are:

- the index should have *broad coverage* rather than be narrow in scope;
- the index should be *comparable across countries or regions*;
- the index should be *timely*; i.e., appear frequently rather than infrequently;
- the index should be *reliable*; i.e., it should not exhibit random fluctuations from period to period;
- the index should be *objective* and *reproducible*; i.e., different statisticians given the specifications for the index and the same basic data should produce the same index number value;
- the index should be as *simple* as possible so that it can be explained to the public;
- the index should be *theoretically consistent* if possible; and finally,
- the index should be *inexpensive to produce.*

Some discussion of the above criteria is called for. In particular, not all of the criteria need be mutually consistent. For example, the criteria of timeliness and reliability can

8. The first person to suggest that the price and quantity indices should be jointly determined in order to satisfy equation (2) was FISHER (1911; p. 418). FRISCH (1930; p. 399) called (2) the product test.

be in conflict: the shorter that we make the time period, then typically, the amount of random noise in the index will increase.[9] Usually, the reliability and cheapness criteria will be in conflict (it costs more to make the index more reliable due to increased sampling costs, etc.) as will be the broadness and cheapness criteria. Objectiveness and broadness of coverage may also be in conflict: objectiveness could be interpreted to mean an absence of imputations (or at least imputations that are not reproducible) and this can lead to problems in imputing rents for owner occupied dwellings. But if the index excludes imputations for owner occupied rent, then the broadness requirement of the index may be impaired. Moreover, excluding imputed rents for owner occupied dwellings may lead to a lack of comparability in the index across regions.[10] The requirement that the index be simple would seem to rule out indexes that make use of rather complicated economic theories. However, some commodities are inherently difficult to price (e.g., the treatment of bank services, insurance and gambling expenditures) and it may be necessary to rely on economic theory to give possible approaches to constructing prices for these difficult to measure services. Leaving out these difficult to price services would conflict with the broadness criterion.

Obviously, the fact that we cannot come up with a definitive, mutually consistent list of criteria that the "ideal" inflation index should satisfy means that our discussion of possible alternative domains of definition for the index cannot be definitive; our discussion can only suggest some "reasonable" possibilities that should be considered.

In order to choose "the" appropriate inflation index, it is first necessary to ask about the *purpose* that the index would be used for. Our primary purpose for the inflation index is as the formal *target index of inflation* that central bankers could target or are obliged to target in order to fulfill their mandates to maintain price stability.[11] However, the inflation index could also serve to fulfill other purposes. HILL notes that there are several uses for such an index in the *national accounts*:

> "A general index of inflation is needed for a variety of purposes. In the SNA, it is used to calculate the following: neutral and real holding gains and losses, internal and external trading gains and losses, real national and disposable income, real interest and constant intra-period price level (CPL) accounts. In business accounting it may be used for similar purposes, such as Current Purchasing Power accounting. A general price index is needed for policy purposes to monitor the general rate of inflation and to set inflation targets. It may also be used to implement indexation agreements under conditions of high or chronic inflation ... In general, the most suitable multipurpose general price indices seem to be those for total final uses or for total domestic final uses. Whatever index is preferred, however, it must be stressed that there remains a need for a range of

9. From Figure 1 in FIELDING and MIZEN (2001; p. 28), it appears that the variability of monthly inflation rates for seven EU countries was greater than the corresponding quarterly rates. The data used were from the Eurostat monthly price database over the period 1983(1) to 1994(12) for nine product categories.
10. From HOFFMANN and KURZ (2001:p. 3), only 40 % of Germans live in owner occupied dwellings while from BOVER and VELILLA (2001; p. 4), about 85 % of Spaniards live in owner occupied dwellings. Thus dropping imputed rents from the index domain of definition would make the German and Spanish inflation indexes somewhat incomparable.
11. This literature is reviewed by MEYER (2001) and ZIESCHANG, BLOEM and ARMKNECHT (2001).

other price indices to meet more specific analytic and policy purposes. A general index of inflation should not drive out other indices." PETER HILL (1996, p. 16–17).

Thus both the national accounting literature and the recent literature on inflation targeting suggest that a general measure of inflation is needed for various purposes. The question is: what exactly should this general measure be; i. e., what are possible sets of transactions that could be used as the domain of definition for this general purpose price index?

Some possible domains of definition are:

- The set of all monetary transactions that take place in the economy.
- Flow aggregates taken from the national accounts.
- Asset or wealth price indexes.
- Flow aggregates taken from economic theory.

We have already discussed the first possibility and noted that due to the preponderance of transactions that are involved in currency trading, enthusiasm for the set of all monetary transactions as the appropriate domain of definition for the inflation index seems to be limited.

The system of national accounts decomposes the economy into flows that are classified by commodity and by institutional sector (mainly households, domestic producers, governments and the rest of the world). Many of the value cells in the national accounts classification give rise to price indexes or deflators which could be candidates for the inflation index. We will consider several possible flow account transactions domains in the following section and then consider stock account domains in section 4 below.

3. THE TARGET INDEX AND THE SNA FLOW ACCOUNTS

Let us partition the economy into *two sets of institutional sectors*: (1) households, governments and the rest of the world (ROW); (2) private and public domestic producers. We also *distinguish two classes of goods*: (1) *outputs produced by domestic producers*: consumption C, government final demands G, gross investment expenditures I and exports X; (2) *primary inputs utilized by domestic producers*: imports M, labour inputs L, capital inputs K and natural resource and land inputs R. Putting the two classifications together, the economy can be approximately[12] described as follows:

	Households, Governments, ROW	*Domestic Producers*
Output Markets	(i) $C + G + I + X$	(ii) Outputs less domestic intermediates
Inputs	(iv) $M + L + K + R$	(iii) Primary inputs plus imports

The above fourfold classification of the value cells for the national accounts is almost the traditional one: the only difference is that we have moved imports down one line and

12. We have neglected the treatment of taxes and many other complications.

classified them as primary inputs. The transactions in cell (i) represent total final expenditures, $C + G + I + X$. By the usual national income accounting arguments (i.e., demand equals supply and each sector or economic agent has a balanced budget during the accounting period), the value flows in (i) can be estimated by the value flows in (ii), (iii) or (iv). Each of the value flows (i) to (iv) considered as ratios over two periods could be decomposed into broad measures of price and quantity change and the associated price index could be used as an inflation index.

Before we examine each of these four alternatives in detail, it is necessary to explain why we classified imports as a primary input rather than as a negative export as is traditional in the national accounts.[13] There are two reasons for this reclassification. The first reason (and the most important reason) is that the traditional final expenditures deflator (or GDP deflator) behaves rather perversely if import prices *rise* because the immediate effect of this rise is to *reduce* the deflator.[14] KOHLI (1982, p. 211) (1983) noticed this problem with the GDP deflator many years ago:

> "Actually, it can easily be seen that any terms of trade change away from the base period price ratio results in a fall in real national product. This clearly reveals the weakness of this measure of real value added, the drawbacks of direct index numbers, and the dangers of aggregating positive with negative quantities." ULRICH KOHLI (1983, p. 142).

Indeed, it is the fact that import quantities have *negative weights* in the GDP deflator that causes it to be unsuitable as a measure of general inflation.

The second reason why we reclassified imports as a primary input is that we want our broad measure of price change to be as stable or smooth as possible over time but if we treat imports as negative exports, the resulting price index will tend to fluctuate more violently than an index which excludes these negative exports. Why is this? The negative weights for import prices will cause the expenditure share weights for the positive components to be *greater than one* and this will magnify the effects of price changes for the nonimport components.[15]

13. Our treatment of imports as a primary input follows that of KOHLI (1978) (1991).
14. An example of this anomalous behavior of the GDP deflator just occurred in the advance release of gross domestic product for the third quarter of 2001 for the US national income and product accounts: the chain type price indexes for C, I, X and M decreased (at annual rates) over the previous quarter by 0.4 %, 0.2 %, 1.4 % and 17.4 % respectively but yet the overall GDP deflator increased by 2.1 %. Thus there was general deflation in all sectors of the economy but yet the overall GDP deflator increased. This is difficult to explain to the public! See Table 4 in the Bureau of Economic Analysis (2001).
15. HILL (1971; p. 17) noted that value added and GDP deflators had weights that summed to something greater than one if we took the absolute values of these weights: "Considered as a weighted average, the unusual feature is that the input index carries a negative weight while the output index carries a weight correspondingly greater than unity." HILL (1996; p. 95) later noted that these magnified weights would tend to make the resulting index more sensitive to random fluctuations in the prices: "The indices are therefore sensitive to errors in both the output and input indices." The problem will not be a significant one in a carefully constructed national value added deflator (where imports are treated as primary inputs) because domestic intermediate inputs will cancel out in a national index. However, imports do not cancel out in the same way.

It should be noted that most of the empirical literature on monetary targeting that is centered around Taylor rules[16] routinely assumes that the quarterly GDP deflator is the "right" price index for inflation targeting purposes.[17] In view of the negative weights for imports problem discussed above, *I would not recommend this index for inflation targeting purposes.*[18]

We now consider the deflation of the flows represented by (i) to (iv) above from the viewpoint of their suitability as domains of definition for the inflation index.[19]

Consider first the household income flows in the domain of definition (iv). It is possible to develop a household sources of income price index that is analogous to the KONÜS (1924) cost of living index that is used to deflate consumer expenditures; see the constant utility income deflator concept in DIEWERT and FOX (1998). However, since this concept is unfamiliar to the public and to price statisticians, we can dismiss it as a practical alternative at this time.

Now consider the primary input expenditure flows in (iii) and restrict attention to the private sector part of these expenditures (or exclude the primary input expenditures of general government). The economic theory of the input price index is well developed and could be applied to these expenditures.[20] However, although the theory of the input price index is well developed, statistical agencies have not yet been able to supply these indexes (with a few exceptions) since they face a number of problems:

- It is difficult to construct price indexes for the thousands of types of labour that exist in modern economies.
- Price statisticians have been reluctant to construct rental prices or user costs of capital[21], which are the appropriate prices that should be associated with the use of capital inputs in the economic approach to the input price index.
- Accurate estimates for resource depletion effects have not been a high priority for most statistical agencies.

16. See for example TAYLOR (1993) and the contributions in TAYLOR (1999).
17. Probably it is the widespread availability of quarterly data both on output and the GDP deflator that has led empirical researchers to use the GDP deflator as their target price index. But since the components of GDP including imports are readily available on a quarterly basis as well, it would not be difficult to compute a quarterly $C + I + G + X$ price deflator.
18. Recall that HILL noted that an index of general inflation was needed to implement Current Purchasing Power accounting. The Accounting Research Division of the American Institute of Certified Public Accountants recommended that the GNP implicit price deflator be used as the measure of general price level change in business accounting because its universe encompasses the entire economy; see TIERNEY (1963; p. 112). In our view, a more appropriate index for this purpose is the CPI.
19. For similar discussions, see HILL (1996; p. 94–97), DIEWERT (1996; p. 29–31), WOOLFORD (1999) and ZIESCHANG, BLOEM and ARMKNECHT (2001).
20. See the references in DIEWERT (1980; p. 455–467) and CAVES, CHRISTENSEN and DIEWERT (1982; p. 1395–1399).
21. Basically, it is not easy to construct objective user costs; i.e., many somewhat arbitrary judgements must be made in order to construct rental prices; see DIEWERT (1980; p. 470–486) and section 7.6 below for some of these difficulties.

Thus although the deflator for the domain of definition (iii) could be used as a broad index of inflation, it is only available in countries that attempt to construct estimates of the total factor productivity of the economy.[22]

Now consider the expenditure flows in the domain of definition (ii) but exclude the provision of the services rendered by general government. The economic theory of the output price index is well developed and could in theory be applied to these flows.[23] However, again economic statisticians have not been able to provide economy wide value added deflators. In particular, accurate output price indexes for the service industries are generally lacking as are intermediate input price deflators for service inputs into the goods producing industries.[24]

Finally, we consider the flows in the domain of definition (i), deliveries to final demand, $C+I+G+X$.[25] Any subcomponent of this aggregate that includes C is probably broad enough to be suitable as a domain of definition for an inflation price index. WOOLFORD (1999) argues, like ZIESCHANG, BLOEM and ARMKNECHT (2001), that C is too narrow for inflation monitoring purposes and suggests the Domestic Final Purchases (DFP) consisting of $C+I+G$ is the appropriate domain of definition for the central bank target inflation index. The DFP price index drops exports from its domain of definition presumably because price movements in exports do not directly affect the inflation faced by domestic final demanders.[26] HILL recommended both the $C+I+G+X$ and $C+I+G$ domains of definition as being suitable for a general index of inflation:

22. Total factor productivity growth can be defined as an output quantity index divided by an input quantity index which in turn (under the assumption that the value of inputs equals the value of outputs) is equal to an input price index divided by an output price index; see JORGENSON and GRILICHES (1967).
23. See FISHER and SHELL (1972; p. 53), SAMUELSON and SWAMY (1974; p. 588), ARCHIBALD (1977), DIEWERT (1980; p. 460–464) and CAVES, CHRISTENSEN and DIEWERT (1982; p. 1399–1401). Note that the prices that are required in order to implement a producer output price index are the prices that producers actually face in their output markets and these prices correspond to what are called basic prices in the latest version of the national accounting framework. ZIESCHANG, BLOEM and ARMKNECHT (2001; p. 6) characterize the difference between basic prices (the prices producers face) and purchasers' prices (the prices final demanders face) as follows: "The 1993 SNA (paragraphs 6.204–6.207) considers two main valuation principles as appropriate depending on whether a transaction is viewed from the buyer's or seller's point of view, namely, purchasers' prices and basic prices. Purchasers' prices refer to the amount paid by the purchaser per unit of a good or service, including taxes on products and charges for transportation, distribution, and insurance invoiced by other providers in the same transaction, and excluding subsidies on products. Basic prices refer to the amount received per unit of a good or service by the seller, excluding taxes on products and charges for transportation, distribution, and insurance invoiced by other providers in the same transaction, and including subsidies on products."
24. "The basic problem with measuring productivity change in the service sector is the unavailability of suitable price index numbers." BERT M. BALK (2001; p. 37). For additional material on the difficulty of economic measurement in the service sector, see DIEWERT and FOX (1999) (2001).
25. HILL (1996; p. 94) terms the price index defined over this domain of definition the "price index for total final expenditures".
26. HILL (1996; p. 96) terms the price index defined over this domain of definition the "total gross domestic final expenditures price index".

"For most purposes, the price indices for total final expenditures and gross domestic expenditures seem to provide more suitable indicators of the rate of general inflation than the indices for GDP or total supplies and uses. The choice between them must be governed to some extent by the use for which they are intended and it is difficult to argue that one measure is inherently superior to the others for all purposes. In general, it would appear desirable to make both of the indices for final expenditure and also that for GDP available to analysts and policy makers." PETER HILL (1996, p. 97).

Thus HILL endorsed both the $C + I + G + X$ and the $C + I + G$ domains of definition for a general inflation measure. However, DIEWERT argued for a smaller domain of definition:

"However, from the viewpoint of measuring the impact of inflation on domestic final demanders, we should exclude exports which belong to the rest of the world. This leaves us with $C + G + I$, which HILL (1996) calls total gross domestic final expenditures. However, the prices of investment goods are not relevant to the deflation of current period household expenditures on goods and services which they consume in the current period and hence gross investment expenditures can be deleted. This leaves us with $C + G$. We cannot readily justify the deletion of government final expenditures on goods and services since many government goods (i.e., subsidized housing and transportation) are direct substitutes for privately provided consumer goods and other government outputs (i.e., garbage collection, road maintenance and protection services) certainly provide a flow of current period services to consumers. The problem with including government expenditures in the index number formula is that it is usually difficult to obtain meaningful prices for deflating these expenditures. Thus in the end, for a variety of reasons, we end up with the consumer price index as being perhaps the best indicator of short run inflation in the economy." W. ERWIN DIEWERT (1996, p. 31).

Summarizing the above discussion, it can be seen that a case can be made for $C + I + G + X$, $C + I + G$, $C + I$ or C as the appropriate domain of definition for a central bank target price index but probably, the strongest case can be made for C. A central bank can aim to control only one target price index and the question is: what should that target be? Stabilizing the price of consumption seems more fundamental than stabilizing the price of an aggregate that large numbers of the public will have difficulty identifying with. On the other hand, stabilizing the price of consumption is a target that can easily be explained to the public. Moreover, consumption is what we ultimately care about so stabilizing its price seems reasonable but certainly a case can be made for the other three possible domains of definition.

If we do choose to stabilize the price of household consumption expenditures, then as a side benefit for this choice, we can draw on the well established economic theories for the cost of living index to help us implement the index.[27] Unfortunately, there is more than one economic theory that we can draw on. In particular, we have the usual *plutocratic* approach to the cost of living index where households are weighted according to their expenditures in each period versus the *democratic* approach to the cost of living index where each household is given equal weight.[28] On social welfare grounds, a strong

27. The economic theory of the cost of living index dates back to KONÜS(1924). For extensions of this theory to many consumers, see POLLAK (1980) (1981) (1983) and DIEWERT (1983a) (2001).
28. The terms democratic and plutocratic are due to Prais (1959). For material on cost of living indexes for groups of households, see POLLAK (1980) (1981) and DIEWERT (1983a; 190) (2001a).

case can be made for the latter concept but the data requirements for implementing this concept are much more demanding.[29] Hence, at the present time, we are probably left with the option of implementing a plutocratic type consumption price index.[30]

The conclusions that can be drawn from this section are the following ones:

- It is not a good idea to use the GDP deflator as an inflation target since all prices could fall and yet the index could rise.
- Within the components of final demand, our preferred domain of definition for the target inflation index is just consumption expenditures, C, but cases could be made for $C + G, C + I + G$ and $C + I + G + X$ as possible domains of definition.
- After choosing C as the target domain of definition, one must still choose between a plutocratic or democratic consumer price index. On practical grounds, our preference is for the plutocratic concept.
- Along with the primary target price index, it will be useful for the central bank to have available (for monitoring purposes) output and input price indexes that apply to the private production sector. The output price index will be a comprehensive producer price index (at basic prices), which will weight (gross) output prices positively and domestic intermediate prices negatively. The input price index will be an aggregate of import prices, wage rates, user costs of reproducible capital and land and resource user costs. These two indexes are required to deflate the value of outputs and the value of inputs respectively into measures of real output and real input and these latter measures in turn can be used to form productivity measures. In fact, the ratio of the input price index to the output price index is itself a measure of total factor productivity growth in the economy.[31]

We turn now to the System of National Accounts balance sheet accounts for possible inflation targets.

4. THE TARGET INDEX AND THE SNA STOCK ACCOUNTS

There is at least one other broad price index that could be used as an inflation target index. This is a price deflator for the components of household wealth. Note that this price deflator can be associated with the *balance sheet accounts* in the system of national accounts instead of the income and expenditure flow accounts as in the previous section.

However, household wealth does not seem to be as fundamental as consumption. Wealth is the nominal constraint in the consumer's intertemporal budget constraint and as such, the "price" of wealth has no real economic meaning. If we try to think about wealth in real terms, then the most natural thing to do is to deflate it by the consumer price index!

29. It is difficult to get accurate data on household consumption expenditures by type of household.
30. The plutocratic index is also suitable for national accounts deflation purposes.
31. See JORGENSON and GRILICHES (1967) or DIEWERT (1992a; 168).

Although we do not think that the price of household wealth is a useful primary index for a central bank to target, some recent papers by GOODHART (1995, 2001) have resurrected a paper by ALCHIAN and KLEIN (1973), which made the following claims:

> "Two commonly cited and newsworthy price indices are the Bureau of Labor Statistics' Consumer Price Index and the Commerce Department's GNP deflator. These indices have become an important part of our economic intelligence and are frequently considered to be the operational counterparts of what economists call "the price level". They, therefore, often are used as measures of inflation and often are targets or indicators of monetary and fiscal policy. Nevertheless, these price indices, which represent measures of current consumption service prices and current output prices, are theoretically inappropriate for the purpose to which they are generally put. The analysis in this paper bases a price index on the Fisherian tradition of a proper definition of intertemporal consumption and leads to the conclusion that a price index used to measure inflation must include asset prices. A correct measure of changes in the nominal money cost of a given utility level is a price index for wealth." ARMEN A. ALCHIAN and BENJAMIN KLEIN (1973, p. 173).

GOODHART has recently argued that ALCHIAN and KLEIN'S theoretical argument advocating that central banks target a price index for wealth has never been refuted:

> "The argument that an analytically correct measure of inflation should take account of asset price changes was made most forcefully by ALCHIAN and KLEIN in 1973, and has never, in my view, been successfully refuted on a *theoretical* plane, though, as we shall see, in Section 2.1 their *particular* proposals have several, perhaps incapacitating, *practical* deficiencies." CHARLES GOODHART (2001, F335).

Since the above quotations directly contradict our view that the price of wealth is not a useful inflation target, it will be useful to spell out the ALCHIAN and KLEIN arguments and our reservations about their analysis.

Consider a single consumer who is maximizing utility over some planning horizon of say $T+1$ periods. Suppose that there are N commodities of interest to this consumer in each period and let $q_t \equiv [q_{t1}, \ldots, q_{tN}]$ denote a period t consumption vector for $t = 0, 1, \ldots, T$. Let the preferences of the consumer be represented by the intertemporal utility function, $f(q_0, q_1, \ldots, q_T)$. Let $p_t \equiv [p_{t1}, \ldots, p_{tN}]$ be a vector of period t prices that the consumer faces at the beginning of period 0 for $t = 0, 1, \ldots, T$. We will give more precise interpretations for these prices below. Following HICKS (1946, p. 130 and 305), define the consumer's *intertemporal expenditure function* E as follows:

$$E(u, p_0, p_1, \ldots, p_T) \equiv \min_{q's} \left\{ \sum_{t=0}^{T} p_t q_t : f(q_0, q_1, \ldots, q_T) = u \right\} \quad (3)$$

where u is a reference utility level, $p \equiv [p_0, p_1, \ldots, p_T]$ is a vector of intertemporal prices that the consumer might face at the beginning of period 0 and $p_t q_t \equiv \sum_{n=1}^{N} p_{tn} q_{tn}$ denotes the inner product between the vectors p_t and q_t. Following POLLAK (1975, p. 181), define the consumer's *intertemporal cost of living index* for two intertemporal price vectors, $p^a \equiv [p_0^a, p_1^a, \ldots, p_T^a]$ and $p^b \equiv [p_0^b, p_1^b, \ldots, p_T^b]$, and for a reference level of intertemporal utility u as follows:

$$P(p^a, p^b, u) \equiv E(u, p^b)/E(u, p^a). \tag{4}$$

The above price index corresponds to ALCHIAN and KLEIN'S (1973, p. 175) *iso-utility price index* defined by their equation (3). Note that on the right hand side of our equation (4), the same reference level of utility is used in the numerator and denominator. Thus only prices change in the numerator and denominator of (4); all other variables (tastes and the reference utility level) are held constant. This is the defining structure of an economic price index.

As POLLAK (1975, p. 181) notes, there are two different interpretations that can be put on the intertemporal price vectors p^a and p^b that appear in (4). Consider the price vector $p^0 \equiv [p_0^0, p_1^0, \ldots, p_T^0]$ that corresponds to the price vector that the consumer actually faces at the beginning of period 0. In POLLAK's "futures prices" interpretation of his model, the n-th component of the period 0 price vector p_t^0, p_{tn}^0, is the price which must be paid at the beginning of period 0 for a contract promising to deliver one unit of commodity n in period t.[32] POLLAK's second interpretation is his "spot price" interpretation:

> "The 'spot price' interpretation gives a different gloss to the same model. Instead of futures markets, we assume perfect foresight and let p_{tn}^{0*} denote the 'spot' price of q_{tn}; that is, p_{tn}^{0*} is the amount which must be paid in period t for the delivery of one unit of good n in period t. We also assume perfect capital markets, so that individuals can borrow or lend without limit at the market rate of interest, and we let r_t denote the interest rate connecting period t with period $t+1$. There is no period 0, but by convention, we let $r_0 \equiv 0$." ROBERT A. POLLAK (1975, p. 181).

POLLAK goes on to show that the period t "present value price" for commodity n, p_{tn}^0, is defined as (using our notation):

$$p_{tn}^0 \equiv p_{tn}^{0*}/(1+r_1)(1+r_2)\ldots(1+r_t); \quad t = 1, 2, \ldots, T; = 1, \ldots, N. \tag{5}$$

The present value prices defined by (5) can replace the earlier futures prices in the intertemporal index defined by (4). POLLAK indicates that the two versions of his model for the intertemporal cost of living index are essentially the same:

> "The difference between the 'spot' and the 'futures price' versions of the intertemporal cost of living index is one of notation rather than of substance. The 'spot' version explicitly identifies the role of interest rates, while their role remains implicit in the 'futures price' version. ROBERT A. POLLAK (1975, p. 182).

The above theory for the intertemporal cost of living index is fine as far as it goes but it can be seen that *it does not lead to a useful inflation index that could be targeted by a central bank*. The Hicks Arrow Debreu model of a futures economy leads to a single equili-

32. Thus p^0 is the equilibrium vector of prices in what is known as the ARROW (1953) DEBREU (1959; p. 101) futures economy. It is interesting to note that HICKS (1946; p. 136) anticipated this model: "It is possible, at the other extreme, to conceive of an economy in which, for a considerable period ahead, everything was fixed up in advance. If all goods were bought and sold forward, not only would current demands and supplies be matched, but also planned demands and supplies. In such a 'Futures Economy', the first two kinds of disequilibrium would be absent."

brium price vector p^0 when markets open and that is the end of the story: prices are completely determined for all subsequent periods! Hence there are no price vectors p^a and p^b to be compared; there is only the single equilibrium price vector p^0.[33]

In order to obtain a useful intertemporal cost of living index of the type defined by (4) above, it will be necessary to use the temporary equilibrium model which was developed by HICKS where the assumptions of perfect foresight or the existence of a complete set of futures markets are replaced by the assumption that economic agents form expectations about future spot prices (which may not be correct) and only current period prices are determined:[34]

> "On the basis of these inherited resources, entrepreneurs (and even private individuals as well) may be supposed to draw up plans, which determine their current conduct and their intended conduct in future weeks. An entrepreneur's plan includes decisions about the quantities of products he will sell in the current week and in future weeks, and about the quantities of inputs (services, materials, perhaps even new acquisitions of plant), which he will purchase or hire in current and future weeks. A private person's plan includes decisions about the quantities of commodities he will buy (and perhaps also the quantities of services he will supply) in current and future weeks. Thus, as part of the plans, the current demands and supplies of all goods and services are determined; though they are determined jointly with people's intentions to demand and supply at future dates." JOHN R. HICKS (1946, p. 130).

> "In determining the system of prices established on the first Monday, we shall also have determined with it the system of plans which will govern the distribution of resources during the following week. If we suppose these plans to be carried out, then they determine the quantity of resources which will be left over at the end of the week, to serve as the basis for the decisions which have to be taken on the second Monday. On that second Monday a new system of prices has to be set up, which may differ more or less from the system of prices which was established on the first. The wider sense of EquilibriumEquilibrium over Time, as we may call it, to distinguish it from the Temporary Equilibrium which must rule within any current week suggests itself when we start to compare the price situations at any two dates. A stationary state is in full equilibrium, not merely when demands equal supplies at the currently established prices, but also when the same prices continue to rule at all dates when prices are constant over time." JOHN R. HICKS (1946, p. 131–132).

We can now use the Hicksian temporary equilibrium idea to give a third interpretation for the intertemporal cost of living index defined by (4) above. Let $p^a = p^0 \equiv [p_0^0, p_1^0, \ldots, p_T^0]$ where $p_0^0 \equiv [p_{01}^0, p_{02}^0, \ldots, p_{0N}^0]$ is the set of period 0 spot prices that the consumer faces in period 0 and the vectors p_t^0 for $t = 1, 2, \ldots, T$ are discounted spot prices that the consumer *expects* to face in future periods where the expectations are formed in period 0.

33. ALCHIAN and KLEIN (1973; p. 175) hint that they recognize this problem with their model but they chose to ignore it: "This model, like the standard microeconomic model under which the usual price indices are derived, assumes the absence of all information or transactions costs and therefore lacks a theoretical justification for the value of a price index. (Introduction of uncertainty by the use of costlessly made contingency contracts (e.g., ARROW, 1953), where all transactors know the true state of the world when it occurs, is also economically equivalent to a world of perfect information with no rationale for a price index.) We will here ignore this fundamental question . . ."
34. WALRAS (1954; Part V) set up the first temporary equilibrium model where equilibrium spot prices were determined along with the interest rate. FISHER (1930) also set up a special case of the more general temporary equilibrium model due to Hicks.

Let $p^b = p^1 \equiv [p_1^1, p_2^1, \ldots, p_{T+1}^1]$ where $p_1^1 \equiv [p_{11}^1, p_{12}^1, \ldots, p_{1N}^1]$ is the set of period 1 spot prices that the consumer faces in period 1 and the vectors p_t^1 for $t = 2, 3, \ldots, T+1$ are discounted spot prices that the consumer *expects* to face in future periods where the expectations are formed in period 1. With these definitions for p^a and p^b, the intertemporal cost of living defined by (4) becomes:

$$P(p^0, p^1, u) \equiv E(u, p^1)/E(u, p^0). \tag{6}$$

However, there are two major problems with this temporary equilibrium interpretation for the intertemporal cost of living index.

The first problem has to do with aging. The expenditure function E defined by (3) above is the right one for an individual who expects to live $T+1$ periods at the beginning of period 0. When we get to period 1, this expenditure function is no longer relevant for this individual: he or she is one period older and now (in general) expects to live for only T periods, not $T+1$ periods. Hence the numerator in the right hand side of (6) is not the right expenditure function to apply to this individual and the entire intertemporal cost of living model breaks down.[35]

The second problem is also a fundamental one. In the atemporal theory of the cost of living index, we can (in principle) observe the prices that consumers face in periods 0 and 1 (i.e., the price vectors p_0^0 and p_1^1 using our earlier notation) and observe their quantity choices (say q_0^0 and q_1^1) for the two periods being compared. This means that we can (with a lag) readily calculate Laspeyres and Paasche price indexes, P_L and P_P, as follows:

$$P_L(p_0^0, p_1^1, q_0^0) \equiv p_1^1 q_0^0 / p_0^0 q_0^0, \; P_P(p_0^0, p_1^1, q_1^1) \equiv p_1^1 q_1^1 / p_0^0 q_1^1. \tag{7}$$

Once we have calculated the above indexes, we can form bounds to various cost of living indexes and by taking the geometric mean of P_L and P_P, we can obtain a fairly good approximation to a true cost of living index.[36] However, in the context of the intertemporal cost of living index defined by (6) above, we cannot observe either the consumer's expected future discounted prices or the corresponding expected future consumption vectors. Hence, it will be impossible to form empirical approximations to (6) and *we are left with a concept that we simply cannot implement*. This negative conclusion applies to all three interpretations for the intertemporal cost of living index defined by (4) or (6) and in particular, it applies to ALCHIAN and KLEIN's (1973) iso-utility price index defined on page 175 of their paper.

35. DIEWERT (2001a) observed that a somewhat similar problem occurs in the usual plutocratic and democratic cost of living indexes: the list of individuals being compared in the two periods under consideration is not constant over the two periods due to births, deaths, immigration and emigration. Also as individuals age, their capabilities change due to aging (and learning) and hence the assumption of constant current period preferences is suspect.
36. See KONÜS (1924), POLLAK (1983) or DIEWERT (2001a) for the details. Section 5.5 below outlines some of these results.

There is another rather severe problem with the model of ALCHIAN and KLEIN. They claimed that asset prices that appear on the right hand side of the consumer's intertemporal budget constraint could be used to approximate futures prices that appear on the left hand side of the consumer's intertemporal budget constraint:

> "If assets are standardized in terms of their present and future service flows, the current vector of asset prices $[P_A(j)]$, can therefore be used as a proxy for current futures prices, $p_A(i,t)$." ARMEN A. ALCHIAN and BENJAMIN KLEIN (1973, p. 177).

This claim is indeed an ambitious one since the assets being held by households will typically be claims to capital stocks being used by firms and *will have nothing to do with the intertemporal consumption prices faced by consumers.*[37]

Summing up the above discussion, it appears that an asset price index is not a useful target for a central bank. However, this negative conclusion does not imply that asset price indexes would not be useful in other contexts. For example, it may well be the case that an asset price index will help us to *forecast* the central bank's target inflation index.[38] Furthermore, asset prices will play a key role as components in indexes that are suitable candidates to be central bank inflation targets. For example, the price of land and the price of owner occupied housing (both asset prices) may be useful in constructing user costs of land and structures for owner occupied housing and these user costs could play a role in an atemporal consumer price index that had a broad domain of definition. Similarly, the prices of land, structures, equipment and inventories (all asset prices) play a role in forming user costs for these inputs and thus would appear in a broadly defined input cost index.

We turn now to the problems involved in picking a specific functional form for the price index.

5. ON CHOOSING THE INDEX NUMBER CONCEPT

Over the years, the theoretical literature on the index number problem has suggested at least 5 different approaches to the problem of choosing a specific index number formula. The five approaches are:

- The fixed basket approach (and averages of fixed baskets);
- The test or axiomatic approach;
- The stochastic or statistical approach;
- The economic approach and
- The approach of DIVISIA.

37. However, owner occupied housing and other long lived consumer durable goods are exceptions to this observation.
38. Papers that follow this line of attack include SHIRATSUKA (1999) and GOODHART and HOFMANN (2000).

The first 4 approaches lead to definite recommendations about the preferred index number formula and we consider each of these approaches below in subsections 5.1 to 5.4. DIVISIA's (1926) approach assumes that prices and quantities are continuous functions of time and hence in order to obtain a practical index number formula, it is necessary to use methods of numerical approximation[39] or make assumptions about the path taken by the price and quantity functions through time and use a line integral approach to the determination of the price index. However, both of these strategies lead to a very large number of possible index number formulae.[40] Hence, the approach of DIVISIA to the problem of choosing an index number formula does not lead to any practical results. Thus we will now move on to other approaches for choosing a functional form for the price index; approaches that lead to some specific functional form recommendations.

5.1. *The Fixed Basket Approach*

We revert to the notation introduced in section 2 above. Historically[41], the simplest way of obtaining a price index that compares the level of prices in period 1 to the level in period 0 is to take the base period basket of commodities that was purchased in period 0, the vector $q^0 \equiv [q_1^0, q_2^0, \ldots, q_N^0]$, price out how much this basket would cost in each of the two periods and take the ratio of these costs. This leads to the *Laspeyres price index*, P_L, defined as follows:

$$P_L(p^0, p^1, q^0) \equiv p^1 q^0 / p^0 q^0 \equiv \sum_{n=1}^{N} p_n^1 q_n^0 \Big/ \sum_{n=1}^{N} p_n^0 q_n^0. \tag{8}$$

However, rather than using the base period basket as the quantity vector which is held constant, the basket that was purchased in period 1, the vector $q^1 \equiv [q_1^1, q_2^1, \ldots, q_N^1]$, is just as valid from an a priori point of view[42] for making comparisons between periods 0 and 1. This leads to the *Paasche price index*, P_P, defined as follows:

39. On this approach, see DIVISIA (1926; p. 40) who showed how the Laspeyres index was a discrete approximation or DIEWERT (2001b; p. 17–18) and the references there. Frisch (1936; p. 8) aptly summed up the difficulties with this discrete approximation approach: "As the elementary formula of the chaining, we may get Laspeyre's or Paasche's or Edgeworth's or nearly any other formula, according as we choose the approximation principle for the steps of the numerical integration."
40. See BALK (2000) for a comprehensive review of the line integral approach.
41. For notes on the early history of index number theory, see DIEWERT (1993a). It should be noted that we focus on the problems involved in making price comparisons between two periods. There are additional problems when we want to make comparisons in a consistent manner between many periods; see HILL (1988).
42. Of course, from the viewpoint of statistical agency practice, the Laspeyres index is to be preferred since it will be difficult to obtain quantity weights for the current period but quite feasible to obtain them for a past base period. However, at this time, we are talking about matters of principle rather than practice.

$$P_P(p^0, p^1, q^1) \equiv p^1 q^1 / p^0 q^1 \equiv \sum_{n=1}^{N} p_n^1 q_n^1 / \sum_{n=1}^{N} p_n^0 q_n^1. \tag{9}$$

If both the Paasche and Laspeyres give much the same answer, then either one could be used as "the" price index. However, if there is significant variation in the relative prices p_n^1/p_n^0, then usually, the Paasche price index will be significantly below the corresponding Laspeyres index. Let me try and explain why this is so.

In order to make this explanation, it is useful to rewrite the Laspeyres and Paasche indexes as functions of the N relative prices, p_n^1/p_n^0 for $n = 1, \ldots, N$, and the N *period t expenditure shares*, s_n^t, defined as follows for $t = 0, 1$:

$$s_n^t \equiv p_n^t / p^t q^t, \quad n = 1, \ldots, N; t = 0, 1. \tag{10}$$

Using definitions (10), we can rewrite the Laspeyres index P_L defined by (8) as follows:

$$\begin{aligned} P_L(p^0, p^1, q^0) &\equiv \sum_{n=1}^{N} p_n^1 q_n^0 / p^0 q^0 \\ &= \sum_{n=1}^{N} (p_n^1/p_n^0) p_n^0 q_n^0 / p^0 q^0 \\ &= \sum_{n=1}^{N} (p_n^1/p_n^0) s_n^0. \end{aligned} \tag{11}$$

Thus the Laspeyres index can be written as a (base period) share weighted *arithmetic average* of the N price relatives. In a similar but slightly more complicated fashion, we can rewrite the Paasche index defined by (9) as follows:

$$\begin{aligned} P_P(p^0, p^1, q^1) &\equiv p^1 \cdot q^1 / \sum_{n=1}^{N} p_n^0 q_n^1 \\ &= \left[\sum_{n=1}^{N} p_n^0 q_n^1 / p^1 \cdot q^1 \right]^{-1} \\ &= \left[\sum_{n=1}^{N} (p_n^0/p_n^1) p_n^1 q_n^1 / p^1 \cdot q^1 \right]^{-1} \\ &= \left[\sum_{n=1}^{N} (p_n^0/p_n^1) s_n^1 \right]^{-1} \quad \text{using definitions (10)for } t = 1 \\ &= \left[\sum_{n=1}^{N} (p_n^1/p_n^0)^{-1} s_n^1 \right]^{-1}. \end{aligned} \tag{12}$$

Thus the Paasche index can be written as a (period 1) share weighted *harmonic average* of the N price relatives.

If the price relatives are not all equal to each other and *if* the period 1 expenditure shares s_n^1 are equal to their period 0 counterparts s_n^0, then it can be shown that a weighted harmonic mean is strictly less than a weighted arithmetic mean (with the same weights in both means)[43] and hence under these conditions, the Paasche index is strictly less than the Laspeyres index; i.e., under these conditions, we have:

$$P_P(p^0, p^1, q^1) < P_L(p^0, p^1, q^0). \tag{13}$$

Of course, it is unlikely that the period 0 and 1 expenditure shares will be *exactly* equal but if they are approximately equal, (which is usually the case empirically), then there is a strong likelihood that the Paasche price index will be numerically smaller than its Laspeyres counterpart.

In any case, if the Paasche and Laspeyres indexes give significantly different numerical answers, then given that we want a *single* number to express the amount of inflation that has taken place going from period 0 to 1, a natural approach is to take an even handed or symmetric average of P_L and P_P as our "final" estimate of inflation. But which type of average should we choose?

The first two types of such *symmetric averages*[44] that might come to mind are the *arithmetic mean* and the *geometric mean* which lead to the DROBISCH (1871, p. 425) SIDGWICK (1883, p. 68) index, P_{DS}, and the FISHER[45] (1922) *ideal index*, P_F defined as follows:

$$P_{DS}(p^0, p^1, q^0, q^1) \equiv (1/2)P_L(p^0, p^1, q^0) + (1/2)P_P(p^0, p^1, q^1); \tag{14}$$

$$P_F(p^0, p^1, q^0, q^1) \equiv \left[P_L(p^0, p^1, q^0)P_P(p^0, p^1, q^1)\right]^{1/2}. \tag{15}$$

The geometric and arithmetic mean are special cases of the *mean of order r*, defined as follows for arbitrary positive numbers a and b:[46]

$$\begin{aligned} m_r(a,b) &\equiv [(1/2)a^r + (1/2)b^r]^{-1} \quad && r \neq 0; \\ &\equiv a^{1/2}b^{1/2} && r = 0. \end{aligned} \tag{16}$$

Thus there are an infinite number of possible averages of P_L and P_P that we could consider. In order to determine which type of symmetric average to choose, we need to consider what *properties* that we would like the resulting index to satisfy.

43. See Hardy, LITTLEWOOD and PÓLYA (1934; p. 26).
44. For a discussion of the properties of symmetric averages, see DIEWERT (1993b). Formally, an average $m(a, b)$ of two numbers a and b is symmetric if $m(a, b) = m(b, a)$.
45. BOWLEY (1899; p. 641) appears to have been the first to suggest the use of this index.
46. For the properties of means of order r, see HARDY, LITTLEWOOD and PÓLYA (1934).

One important property that we would like our chosen index number formula, $P(p^0, p^1, q^0, q^1)$ to satisfy is *the time reversal test*, which is the following property:

$$P(p^0, p^1, q^0, q^1) = 1/P(p^1, p^0, q^1, q^0). \tag{17}$$

Thus if we reverse the roles of periods 0 and 1 in the index number formula, we get the reciprocal of the original index. Obviously, a single price relative, p_1^1/p_1^0, satisfies this property so we would like our index to also satisfy this property. If $P(p^0, p^1, q^0, q^1)$ does not satisfy the time reversal test, then the formula gives essentially different answers depending on which period we choose as the base. Put another way, if the time reversal test is not satisfied, then there exist two price and two quantity vectors, p^0, p^1, q^0, q^1, such that

$$P(p^0, p^1, q^0, q^1)P(p^1, p^0, q^1, q^0) \neq 1. \tag{18}$$

Thus we have a situation where there is a certain amount of price change going from period 0 to 1 and then in period 2, the price and quantity data revert to the data of period 0 but the initial amount of price change is *not reversed* by an index number formula that does not satisfy the time reversal test.[47]

It can be shown that the Fisher ideal index P_F defined by (15) satisfies the time reversal test but the Drobisch Sidgwick index P_{DS} defined by (14) does not. In fact, DIEWERT (1997, p. 138) showed that the geometric mean is the *only* homogeneous mean of the Paasche and Laspeyres indexes that leads to an index number formula that satisfies the time reversal test.[48] Thus this approach of taking a symmetric average of the Paasche and Laspeyres indexes leads to the Fisher ideal formula as being "best" in this class of index number formulae.

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.[49] 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 $n = 1, 2, \ldots, N$.[50] Price statisticians refer to this type of index as a *pure price index*[51] and

47. It is easy to show that the Laspeyres and Paasche price indexes do not satisfy the time reversal test, which is a major problem with the use of these indexes.
48. The mean function $m(a, b)$ need only satisfy two properties to get this result: (i) positivity; $m(a, b) > 0$ if $a > 0$ and $b > 0$ and (ii) (positive) linear homogeneity; $m(\lambda a, \lambda b) = \lambda m(a, b)$ for all $\lambda > 0$, $a > 0$ and $b > 0$.
49. FISHER (1922) considered both averaging strategies in his classic study on index numbers. WALSH (1901) (1921) concentrated on the second averaging strategy,
50. Note that we have chosen the mean function $m(q_n^0, q_n^1)$ to be the same for each commodity n. We assume that $m(a, b)$ has at least the following two properties: $m(a, b)$ is a positive and continuous function, defined for all positive numbers a and b and $m(a, a) = a$ for all $a > 0$.
51. See section 7 in DIEWERT (2001a).

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:

$$P_K(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^{N} p_n^1 m(q_n^0,q_n^1) / \sum_{j=1}^{N} p_j^0 m(q_j^0,q_j^1). \tag{19}$$

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 (19). Again we ask that P_K satisfy the *time reversal test*, (17) above. Under this hypothesis, it is immediately obvious that the mean function m must be a *symmetric mean*[52]; i.e., m must satisfy the following property: $m(a,b) = m(b,a)$ for all $a > 0$ and $b > 0$. This assumption still does not pin down the functional form for the pure price index defined by (19) above. For example, the function $m(a,b)$ could be the *arithmetic mean*, $(1/2)a + (1/2)b$, in which case (19) reduces to the *Marshall* (1887) *Edgeworth* (1925) *price index* P_{ME}, which was the pure price index preferred by KNIBBS (1924, 56):

$$P_{ME}(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^{N} p_n^1 (1/2)(q_n^0 + q_n^1) / \sum_{j=1}^{N} p_j^0 (1/2)(q_j^0 + q_j^1). \tag{20}$$

On the other hand, the function $m(a,b)$ could be the *geometric mean*, $(ab)^{1/2}$, in which case (19) reduces to the *Walsh* (1901, p. 398) (1921, p. 97) *price index*, P_W.[53]:

$$P_W(p^0,p^1,q^0,q^1) \equiv \sum_{n=1}^{N} p_n^1 (q_n^0 q_n^1)^{1/2} / \sum_{j=1}^{N} p_j^0 (q_j^0 q_j^1)^{1/2}. \tag{21}$$

However, there are many other possibilities for the mean function m, including the mean of order r, $[(1/2)a^r + (1/2)b^r]^{1/r}$ for $r \neq 0$. Obviously, in order to completely determine the functional form for the pure price index P_K, we need to impose at least one additional test or axiom on $P_K(p^0,p^1,q^0,q^1)$.

In order to obtain an additional axiom, we note that there is a problem with the use of the Marshall Edgeworth price index (20) in the context of using the formula to make international comparisons of prices. If the price levels of a very large country are compared to the price levels of a small country using formula (20), then the quantity vector of the large country may totally overwhelm the influence of the quantity vector corresponding to the small country.[54] In technical terms, the Marshall Edgeworth formula is

52. For more on symmetric means, see DIEWERT (1993b; p. 361).
53. Walsh endorsed PW 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." WALSH (1921; p. 103). His formula 6 is PW defined by (28) and his 9 is the FISHER ideal defined by (22) above.
54. This is not likely to be a severe problem in the time series context where the change in quantity vectors going from one period to the next is likely to be small.

not homogeneous of degree 0 in the components of both q^0 and q^1. To prevent this problem from occurring in the use of a pure price index $P_K((p^0, p^1, q^0, q^1)$ defined by (19), we ask that P_K satisfy the following *invariance to proportional changes in current quantities test.*[55]

$$P_K(p^0, p^1, q^0, \lambda q^1) = P_K(p^0, p^1, q^0, q^1) \text{ for all } p^0, p^1, q^0, q^1 \text{ and all } \lambda > 0. \tag{22}$$

The two tests, the time reversal test (27) and the invariance test (22), enable us to determine the precise functional form for the pure price index P_K defined by (29) above: the pure price index P_K must be the WALSH index P_W defined by (21).[56]

In order to be of practical use by statistical agencies, an index number formula must be able to be expressed as a function of the base period expenditure shares, s_n^0, the current period expenditure shares, s_n^1, and the N price ratios, p_n^1/p_n^0. The WALSH price index defined by (21) above can be rewritten in this format:

$$\begin{aligned} P_W(p^0, p^1, q^0, q^1) &\equiv \sum_{n=1}^{N} p_n^1 (q_n^0 q_n^1)^{1/2} \Big/ \sum_{j=1}^{N} p_j^0 (q_j^0 q_j^1)^{1/2} \\ &= \sum_{n=1}^{N} \left[p_n^1/(p_n^0 p_n^1)^{1/2} \right] (s_n^0 s_n^1)^{1/2} \Big/ \sum_{j=1}^{N} \left[p_j^0/(p_j^0 p_j^1)^{1/2} \right] (s_j^0 s_j^1)^{1/2} \\ &= \sum_{n=1}^{N} (s_n^0 s_n^1)^{1/2} \left[p_n^1/p_n^0 \right]^{1/2} \Big/ \sum_{j=1}^{N} (s_j^0 s_j^1)^{1/2} \left[p_j^0/p_j^1 \right]^{1/2}. \end{aligned} \tag{23}$$

We sum up the results of this subsection as follows. Our first approach in this subsection was to take an even handed average of the two primary fixed basket indexes: the Laspeyres and Paasche price indices. These two primary indexes are based on pricing out the baskets that pertain to the two periods under consideration. In a sense, they are extreme baskets. Taking an average of them led to the *Fisher ideal price index* P_F defined by (15) above. Our second approach was to average the basket quantity weights and then price out this average basket at the prices pertaining to the two situations under consideration. This approach led to the *Walsh price index* P_W defined by (21) above. Both of these indexes can be written as a function of the base period expenditure shares, s_n^0, the current period expenditure shares, s_n^1, and the N price ratios, p_n^1/p_n^0. Assuming that the statistical agency has information on these three sets of variables, which index should be used? Experience with normal time series data has shown that these two indices will not differ substantially and thus it is a matter of indifference which of these in-

55. This is the terminology used by DIEWERT (1992b; p. 216). VOGT (1980) was the first to propose this test.
56. See section 7 of DIEWERT (2001a).

dexes is used in practice.[57] Both of these indices are examples of *superlative indexes*, which will be defined in subsection 5.4 below. However, note that both of these indexes treat the data pertaining to the two situations in a *symmetric* manner. HILL[58] commented on superlative price indexes and the importance of a symmetric treatment of the data as follows:

> "Thus economic theory suggests that, in general, a symmetric index that assigns equal weight to the two situations being compared is to be preferred to either the Laspeyres or Paasche indices on their own. The precise choice of superlative index – whether FISHER, TÖRNQVIST or other superlative index – may be of only secondary importance as all the symmetric indices are likely to approximate each other, and the underlying theoretic index fairly closely, at least when the index number spread between the Laspeyres and Paasche is not very great." HILL (1993, p. 384).

We turn now to our second general approach to index number theory.

5.2. *The Test or Axiomatic Approach*

The *test* or *axiomatic approach* to index number theory regards the price index, $P(p^0,p^1,q^0,q^1)$, as a function of the price vectors that pertain to the two periods under consideration, p^0 and p^1, and of the quantity vectors that pertain to the two periods under consideration, q^0 and q^1. The basic idea of the axiomatic approach is that the index number formula, $P(p^0,p^1,q^0,q^1)$, is to be regarded as some sort of weighted average of the individual price relatives, $p_1^1/p_1^0,\ldots,p_N^1/p_N^0$, and with this structure in mind, we ask that $P(p^0,p^1,q^0,q^1)$ satisfy a sufficient number of mathematical properties that a weighted average of price relatives would satisfy until the functional form for P is determined. The origins of this approach go back a century or so to WALSH (1901, 1921) and FISHER (1911, 1922)[59] but in more recent years, some key references are EICHHORN and VOELLER (1976), DIEWERT (1992b), Balk (1995) and VON AUER (2001).

We have already listed some tests in the previous subsection: recall the *time reversal test* defined by (24) and the *invariance to proportional changes in current quantities test*, (22). We list a few more tests below.

Our first additional test is the invariance to changes in the units of measurement test or commensurability test:

57. DIEWERT (1978; p. 887–889) showed that these two indices will approximate each other to the second order around an equal price and quantity point. Thus for normal time series data where prices and quantities do not change much going from the base period to the current period, the indices will approximate each other quite closely. Note that HILL (2000) has recently shown that while the commonly used superlative indexes approximate each other closely, this is not the case for the quadratic mean of order r price indexes defined below in section 5.4 for extreme values of r.
58. See also HILL (1988).
59. For references to the early history of the test approach, see DIEWERT (1992b) (1993a) and BALK (1995).

$$\begin{aligned} &P(\alpha_1 p_1^0,\ldots,\alpha_N p_N^0;\alpha_1 p_1^1,\ldots,\alpha_N p_N^1;\alpha_1^{-1}q_1^0,\ldots,\alpha_N^{-1}q_n^0;\alpha_1^{-1}q_1^1,\ldots,\alpha_N^{-1}q_n^1) \\ &= P(p_1^0,\ldots,p_N^0;p_1^1,\ldots,p_N^1;q_1^0,\ldots,q_n^0;q_1^1,\ldots,q_N^1) \text{ for all } \alpha_1>0,\ldots,\alpha_N>0. \end{aligned} \tag{24}$$

This test says that the price index does not change if the units of measurement for each commodity are changed. It is a very important test since the units of measurement for commodities are arbitrary.[60]

Our next two tests are also important ones that restrict the behavior of the index as prices in either period are multiplied by a common scalar factor. Thus consider the *proportionality in current prices test*:

$$P(p^0,\lambda p^1,q^0,q^1) = \lambda P(p^0,p^1,q^0,q^1) \text{ for } \lambda > 0. \tag{25}$$

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. Now consider the *inverse proportionality in base period prices test*:

$$P(\lambda p^0,p^1,q^0,q^1) = \lambda^{-1} P(p^0,p^1,q^0,q^1) \text{ for } \lambda > 0. \tag{26}$$

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.

Our next two 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. Thus we have the *monotonicity in current period prices test*:

$$P(p^0,p^1,q^0,q^1) < P(p^0,p^2,q^0,q^1) \text{ if } p^1 < p^2. \tag{27}$$

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, p. 23) and it is a very reasonable property for a price index to satisfy. Similarly, we have the *monotonicity in base period prices test*:

$$P(p^0,p^1,q^0,q^1) > P(p^2,p^1,q^0,q^1) \text{ if } p^0 < p^2. \tag{28}$$

60. BALK (1995; p. 73) observed that this test had a very important consequence: "Thus P can be written as a function of only $3N$ variables, namely N price ratios p_i^1/p_i^0, N comparison period values $p_i^1 q_i^1$ and N base period values $p_i^0 q_i^0$."

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, p. 23).

There is one more test that we need to list and then we can discuss recent results on which formula is "best" from the axiomatic perspective. The remaining test is the *consistency in aggregation test.* VARTIA (1976) defined an index number formula to be consistent in aggregation if the value of the index calculated in two stages necessarily coincides with the index calculated in a single stage. It turns out to be a bit tricky to provide a rigorous definition of this property; DIEWERT (1978, p. 895; 2001 b, p. 62–64), BALK (1995, p. 85; 1996) and VON AUER (2001, p. 9–10) all provided alternative definitions.

We will explain DIEWERT's definition of consistency in aggregation since it seems to be the most straightforward definition. We suppose that the price and quantity data for period t, p^t and q^t, can be written in terms of M subvectors as follows:

$$p^t = (p^{t1}, p^{t2}, \dots, p^{tM}); \; q^t = (q^{t1}, q^{t2}, \dots, q^{tM}); \; t = 0, 1 \tag{29}$$

where the dimensionality of the subvectors p^{tm} and q^{tm} is N_m for $m = 1, 2, \dots, M$ with the sum of the dimensions N_m equal to N. These subvectors correspond to the price and quantity data for subcomponents of the price index for period t. We construct subindexes for each of these components going from period 0 to 1. For the base period, we set the price for each of these subcomponents, say P_m^0 for $m = 1, 2, \dots, M$, equal to 1 and we set the corresponding base period subcomponent quantities, say Q_m^0 for $m = 1, 2, \dots, M$, equal to the base period value of consumption for that subcomponent for $m = 1, 2, \dots, M$; i.e., we have:

$$P_m^0 \equiv 1; \; Q_m^0 \equiv \sum_{i=1}^{Nm} p_i^{0m} q_i^{0m} \quad \text{for } m = 1, 2, \dots, M. \tag{30}$$

Now we use the chosen index number formula in order to construct a period 1 price for each subcomponent, say P_m^1 for $m = 1, 2, \dots, M$, of the consumer price index. Thus the period 1 subcomponent prices are defined as follows:

$$P_m^1 \equiv P_L^m(p^{0m}, p^{1m}, q^{0m}, q^{1m}) \quad \text{for } m = 1, 2, \dots, M. \tag{31}$$

Once the period 1 prices for the M subindexes have been defined by (31), then corresponding subcomponent period 1 quantities Q_m^1 for $m = 1, 2, \dots, M$ can be defined by deflating the period 1 subcomponent values $\sum_{i=1}^{Nm} p_i^{1m} q_i^{1m}$ by the prices P_m^1 defined by (38); i.e., we have:

$$Q_m^1 \equiv \sum_{i=1}^{Nm} p_i^{1m} q_i^{1m} / P_m^1 \quad \text{for } m = 1, 2, \dots, M. \tag{32}$$

We can now define subcomponent price and quantity vectors for each period $t = 0, 1$ using equations (30) to (32) above. Thus we define the period 0 and 1 subcomponent price vectors P^0 and P^1 as follows:

$$P^0 = (P_1^0, P_2^0, \ldots, P_M^0) \equiv 1_M; \; P^1 = (P_1^1, P_2^1, \ldots, P_M^1) \tag{33}$$

where 1_M denotes a vector of ones of dimension M and the components of P^1 are defined by (31). The period 0 and 1 subcomponent quantity vectors Q^0 and Q^1 are defined as follows:

$$Q^0 = (Q_1^0, Q_2^0, \ldots, Q_M^0); \; Q^1 = (Q_1^1, Q_2^1, \ldots, Q_M^1) \tag{34}$$

where the components of Q^0 are defined in (30) and the components of Q^1 are defined by (32). The price and quantity vectors in (33) and (34) represent the results of the first stage aggregation. We can now use these vectors as inputs into the second stage aggregation problem; i.e., we can now apply the chosen price index formula using the information in (33) and (34) as inputs into the index number formula. Denote this two stage formula as $P^*(P^0, P^1, Q^0, Q^1)$. We ask whether this two stage index equals the corresponding single stage index P; i.e., we ask whether

$$P^*(P^0, P^1, Q^0, Q^1) = P(p^0, p^1, q^0, q^1). \tag{35}$$

If the Laspeyres or Paasche formula is used at each stage of each aggregation, the answer to the above question is *yes*. However, *none* of the other index number formulae defined in this paper satisfy this somewhat stringent definition of consistency in aggregation.[61]

We now turn to the recent papers by DIEWERT, BALK and VON AUER who all had candidates for the "best" index number formula from the viewpoint of the test approach.

DIEWERT (1992b, p. 223) showed that the Fisher price index P_F defined by (15) above satisfied some 21 tests of which 18 were regarded as important and hence concluded that the Fisher index was probably "best" from the viewpoint of the axiomatic approach..[62]

BALK (1995) disputed DIEWERT's conclusion to a certain extent:

> "Finally, the characterizations of the Fisher price index provide no evidence for preferring this index to the other ideal index mentioned, the VARTIA-II price index." Bert M. BALK (1995, p. 87).

The *Vartia-II* (1976) *price index* is defined as follows:

61. However, DIEWERT (1978) showed that many of the indexes defined earlier will be approximately consistent in aggregation. In particular, PDS, PF, PME and PW all have this approximate consistency in aggregation property.
62. The FISHER price index satisfies all of the tests listed above except the consistency in aggregation property.

$$P_V(p^0, p^1, q^0, q^1) \equiv \sum_{n=}^{N} L(s_n^0, s_n^1) \ln(p_n^1/p_n^0) / \sum_{n=1}^{N} L(s_n^0, s_n^1). \tag{36}$$

where the expenditure shares s_n^t are defined by (10) above and the *logarithmic mean function* $L(a, b)$ for a and b positive is defined by

$$\begin{aligned} L(a,b) &\equiv [a-b]/[\ln a - \ln b] \quad \text{for } a \neq b \\ &\equiv a \quad\quad\quad\quad\quad\quad\quad \text{for } a = b. \end{aligned} \tag{37}$$

However, REINSDORF and DORFMAN (1999) and VON AUER (2001, p. 14) showed that the Vartia-II index P_V did not satisfy the important monotonicity tests (27) and (28) above, whereas the Fisher index P_F does satisfy these tests. Hence, in my view, the Fisher index clearly dominates the Vartia-II index from the perspective of the axiomatic approach.

Finally, VON AUER makes a strong case for the Marshall Edgeworth formula P_{ME} defined earlier by (20):

> "In sum, my personal champion is the MARSHALL-EDGEWORTH index P_{ME}, a price index formula which has been thoroughly neglected in past debates on the 'best' price index formula." LUDWIG VON AUER (2001, p. 15).

Indeed, the Marshall Edgeworth index is a worthy competitor to the Fisher index. On the negative side, the Marshall Edgeworth index does not satisfy the invariance to proportional changes in current quantities test[63], (22), whereas the Fisher index does satisfy this test. Both the Fisher and Marshall Edgeworth indexes do not satisfy Diewert's stronger consistency in aggregation property but on the positive side, the Marshall Edgeworth price index satisfies von Auer's weaker consistency in aggregation property whereas the Fisher index does not have this property. However, in this day and age of computer power, I do not think that the consistency in aggregation property is a very important one so I would still prefer the Fisher index over the Marshall Edgeworth.

Summing up the above discussion, our conclusion at this stage is that perhaps the "best" index number formula from the viewpoint of the test approach is the Fisher ideal price index P_F defined by (15) but a case can be made for the Marshall Edgeworth index P_{ME} defined by (20).

63. This means that the quantity index that corresponds using (2) to the Marshall Edgeworth price index, $Q_{ME}(p^0, p^1, q^0, q^1) \equiv p^1 q^1 / [p^0 q^0 P_{ME}(p^0, p^1, q^0, q^1)]$, does not satisfy the important linear homogeneity property for quantity indexes, $Q_{ME}(p^0, p^1, q^0, \lambda q^1) = \lambda Q_{ME}(p^0, p^1, q^0, q^1)$ for all $\lambda > 0$.

5.3 The Stochastic Approach

There are two main branches to the stochastic approach for the determination of the price index: the *weighted* and *unweighted* approaches.

The unweighted approach can be traced back to the work of JEVONS and EDGEWORTH over a hundred years ago[64]. The basic idea behind the unweighted stochastic approach is that each price relative, p_i^1/p_i^0 for $i = 1, 2, \ldots, N$ can be regarded as an estimate of a common inflation rate α between periods 0 and 1[65]; i.e., it is assumed that

$$p_i^1/p_i^0 = \alpha + \varepsilon_i; i = 1, 2, \ldots, N \tag{38}$$

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

$$P_C(p^0, p^1) \equiv \sum_{i=1}^{N} (1/N) p_i^1/p_i^0. \tag{39}$$

A drawback of the Carli price index is that it does not satisfy the time reversal test, i.e., $P_C(p^1, p^0) \neq 1/P_C(p^0, p^1)$[66].

Let us change our stochastic specification and assume that the logarithm of each price relative, $\ln(p_i^1/p_i^0)$, is an unbiased estimate of the logarithm of the inflation rate between periods 0 and 1, β say. The counterpart to (38) is now:

$$\ln(p_i^1/p_i^0) = \beta + \varepsilon_i; i = 1, 2, \ldots, N \tag{40}$$

where $\beta \equiv \ln \alpha$ and the ε_i 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:

$$P_J(p^0, p^1) \equiv \prod_{i=1}^{N} (p_i^1/p_i^0)^{1/N}. \tag{41}$$

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 indices suf-

64. For references to the literature, see DIEWERT (1993a, p. 37–38; 1995a, 1995b).
65. "In drawing our averages the independent fluctuations will more or less destroy each other; the one required variation of gold will remain undiminished." W. STANLEY JEVONS (1884, p. 26).
66. In fact FISHER (1922, p. 66) noted that $P_C(p^0, p^1)P_C(p^1, p^0) \geq 1$ unless the period 1 price vector p^1 is proportional to the period 0 price vector p^0; i.e., FISHER showed that the Carli index has a definite upward bias. He urged statistical agencies not to use this formula.

fer from a fatal flaw: each price relative p_i^1/p_i^0 is regarded as *being equally important* and is given an equal weight in the index number formulae (39) and (41).

WALSH pointed out the problem with the unweighted stochastic approach:

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

However, WALSH did not specify exactly how these economic weights should be determined.

THEIL (1967, p. 136–137) proposed a solution to the lack of weighting in the Jevons index, (41). 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 ith price relative is equal to $s_i^0 \equiv p_i^0 q_i^0 / \sum_{k=1}^N p_k^0 q_k^0$, the period 0 expenditure share for commodity i. Then the overall mean (period 0 weighted) logarithmic price change is $\sum_{i=1}^N s_i^0 \ln(p_i^1/p_i^0)$. 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 $\sum_{i=1}^N s_i^1 \ln(p_i^1/p_i^0)$. 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[67] argued that a nice symmetric index number formula can be obtained if we make the probability of selection for the nth price relative equal to the arithmetic average of the period 0 and 1 expenditure shares for commodity n. Using these probabilities of selection, THEIL's final measure of overall logarithmic price change was

$$\ln P_T(p^0, p^1, q^0, q^1) \equiv \sum_{n=1}^{N} (1/2)(s_n^0 + s_n^1) \ln(p_n^1/p_n^0). \tag{42}$$

The index P_T defined by (42) is equal to an index defined earlier by the Finnish economist LEO TÖRNQVIST (1936).[68]

67. "The price index number defined in (1.8) and (1.9) uses the n individual logarithmic price differences as the basic ingredients. They are combined linearly by means of a two stage random selection procedure: First, we give each region the same chance $\frac{1}{2}$ of being selected, and second, we give each dollar spent in the selected region the same chance ($1/m_a$ or $1/m_b$) of being drawn." HENRI THEIL (1967; p. 138).
68. See also TÖRNQVIST and TÖRNQVIST (1937) where the formula was explicitly defined.

We can give the following statistical interpretation of the right hand side of (42). Define the n-th logarithmic price ratio r_n by:

$$r_n \equiv \ln(p_n^1/p_n^0) \text{ for } n = 1, \ldots, N. \tag{43}$$

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 $n = 1, \ldots, N$. Note that since each set of expenditure shares, s_n^0 and s_n^1, sums to one over n, the probabilities ρ_n will also sum to one. It can be seen that the expected value of the discrete random variable R is

$$E[R] \equiv \sum_{n=1}^{N} \rho_n r_n = \sum_{n=1}^{N} (1/2)(s_n^0 + s_n^1) \ln(p_n^1/p_n^0) = \ln P_T(p^0, p^1, q^0, q^1) \tag{44}$$

using (43) and (42). Thus the logarithm of the index P_T can be interpreted as *the expected value of the distribution of the logarithmic price ratios* in the domain of definition under consideration, where the N discrete price ratios in this domain of definition are weighted according to THEIL's probability weights, $\rho_n \equiv (1/2)[s_n^0 + s_n^1]$ for $n = 1, \ldots, N$.

Taking antilogs of both sides of (44), we obtain the Törnqvist Theil price index, P_T. This index number formula has a number of good properties. In particular, P_T satisfies the proportionality in current prices test (25) and the time reversal test (17) discussed earlier. These two tests can be used to justify THEIL's (arithmetic) method of forming an average of the two sets of expenditure shares in order to obtain his probability weights, $\rho_n \equiv (1/2)[s_n^0 + s_n^1]$ for $n = 1, \ldots, N$. Consider the following *symmetric mean class of logarithmic index number formulae*:

$$\ln P_s(p^0, p^1, q^0, q^1) \equiv \sum_{n=1}^{N} m(s_n^0, s_n^1) \ln(p_n^1/p_n^0) \tag{45}$$

where $m(s_n^0, s_n^1)$ is a positive function of the period 0 and 1 expenditure shares on commodity n, s_n^0 and s_n^1 respectively. In order for P_s to satisfy the time reversal test, it is necessary that the function m be symmetric. Then it can be shown[69] that for P_s to satisfy test (25), m must be the arithmetic mean. This provides a reasonably strong justification for THEIL's choice of the mean function.

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

$$t_n \equiv \ln p_n^0/p_n^1 \text{ for } n = 1, \ldots, N = \ln(p_n^1/p_n^0)^{-1} = -\ln(p_n^1/p_n^0) = -r_n \tag{46}$$

69. See DIEWERT (2000) and BALK and DIEWERT (2001).

where the last equality follows using definitions (43). We can still associate the symmetric probability, $\rho_n \equiv (1/2)[s_n^0 + s_n^1]$, with the n-th reciprocal logarithmic price ratio t_n for $n = 1, \ldots, N$. Now define the discrete random variable, T say, as the random variable which can take on the values t_n with probabilities $\rho_n \equiv (1/2)[s_n^0 + s_n^1]$ for $n = 1, \ldots, N$. It can be seen that the expected value of the discrete random variable T is

$$\begin{aligned} E[T] &\equiv \textstyle\sum_{n=1}^N \rho_n t_n \\ &= -\textstyle\sum_{n=1}^N \rho_n r_n && \text{using (46)} \\ &= -E[R] && \text{using (44)} \\ &= -\ln P_T(p^0, p^1, q^0, q^1). \end{aligned} \tag{47}$$

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

It is possible to consider weighted stochastic approaches to index number theory where we look at the distribution of the price ratios, p_n^1/p_n^0, rather than the distribution of the logarithmic price ratios, $\ln p_n^1/p_n^0$. Thus, again following in the footsteps of THEIL, suppose we draw price relatives at random in such a way that each dollar of expenditure in the *base period* has an equal chance of being selected. Then the probability that we will draw the n-th price relative is equal to s_n^0, the period 0 expenditure share for commodity n. Now the overall mean (period 0 weighted) price change is:

$$P_L(p^0, p^1, q^0, q^1) = \sum_{n=1}^{N} s_n^0 (p_n^1/p_n^0), \tag{48}$$

which turns out to be the Laspeyres price index, P_L (recall (11) above). This stochastic approach is the natural one for studying *sampling problems* associated with implementing a Laspeyres price index.[70]

In the above weighted stochastic approaches to index number theory, the price relatives or their logarithms were regarded as having discrete probability distributions where the probabilities associated with each relative (or logarithmic relative) were functions of the expenditure shares in the two periods under consideration. The index number formula was taken to be the *mean* of the appropriate discrete distribution of these price relatives. However, other measures of central tendency of the distribution could be chosen, such as the weighted median or a trimmed mean. For additional material on

70. We prefer THEIL's stochastic approach to the LASPEYRES approach because the former approach treats the data pertaining to the two periods in a symmetric manner. Moreover, the Laspeyres index does not satisfy the crucial time reversal test.

these alternative stochastic approaches, see DIEWERT (1995b), CECCHETTI (1997) and WYNNE (1997).

In summary, a reasonably strong case can be made for the Törnqvist Theil price index P_T defined by (42) as being the "best" index number formula that falls out of the weighted stochastic approach to index number theory.

We turn now to our last approach to determining the "best" functional form for the price index.

5.4. The Economic Approach

In this subsection, we will outline the theory of the cost of living index for a single consumer (or household) that was first developed by the Russian economist, A. A. KONÜS (1924). This theory relies on the assumption of *optimizing behavior* on the part of economic agents (consumers or producers). Thus given a vector of commodity or input prices p^t that the agent faces in a given time period t, it is assumed that the corresponding observed quantity vector q^t is the solution to a cost minimization problem that involves either the consumer's preference or utility function f or the producer's production function f. Thus in contrast to the axiomatic approach to index number theory, the economic approach does *not* assume that the two quantity vectors q^0 and q^1 are independent of the two price vectors p^0 and p^1. In the economic approach, the period 0 quantity vector q^0 is determined by the consumer's preference function f and the period 0 vector of prices p^0 that the consumer faces and the period 1 quantity vector q^1 is determined by the consumer's preference function f and the period 1 vector of prices p^1.

In the economic approach, it is assumed that the consumer has well defined *preferences* over different combinations of the N consumer commodities or items. Each combination of items can be represented by a positive vector $q \equiv [q_1, \ldots, q_N]$. The consumer's preferences over alternative possible consumption vectors q are assumed to be representable by a continuous, nondecreasing and concave[71] utility function f. Thus if $f(q^1) > f(q^0)$, then the consumer prefers the consumption vector q^1 to q^0. We further assume that the consumer minimizes the cost of achieving the period t utility level $u^t \equiv f(q^t)$ for periods $t = 0, 1$. Thus we assume that the observed period t consumption vector q^t solves the following *period t cost minimization problem*:

$$C(u^t, p^t) \equiv \min_q \left\{ \sum_{n=1}^{N} p_n^t q_n : f(q) = u^t \equiv f(q^t) \right\} = \sum_{n=1}^{N} p_n^t q_n^t;\ t = 0, 1. \tag{49}$$

71. f is concave if and only if $f(\lambda q^1 + (1-\lambda)q^2) \geq \lambda f(q^1) + (1-\lambda) f(q^2)$ for all $0 \leq \lambda \leq 1$ and all $q^1 \gg 0_n$ and $q^2 \gg 0_n$. Note that $q \geq 0_N$ means that each component of the N dimensional vector q is nonnegative, $q \gg 0_N$ means that each component of q is positive and $q > 0_N$ means that $q \geq 0_N$ but $q \neq 0_N$; i.e. q is nonnegative but at least one component is positive.

The period t price vector for the n commodities under consideration that the consumer faces is p^t. Note that the solution to the cost or expenditure minimization problem (49) for a general utility level u and general vector of commodity prices p defines the *consumer's cost function*, $C(u, p)$.

The Konüs (1924) family of *true cost of living indices* pertaining to two periods where the consumer faces the strictly positive price vectors $p^0 \equiv (p_1^0, \ldots, p_N^0)$ and $p^1 \equiv (p_1^1, \ldots, p_N^1)$ in periods 0 and 1 respectively is defined as the ratio of the minimum costs of achieving the same utility level $u \equiv f(q)$ where $q \equiv (q_1, \ldots, q_N)$ is a positive reference quantity vector; i.e., we have

$$P_K(p^0, p^1, q) \equiv C[f(q), p^1]/C[f(q), p^0]. \tag{50}$$

Definition (50) defines a *family* of price indices because there is one such index for each chosen reference quantity vector q.

It is natural to choose two specific reference quantity vectors q in definition (50): the observed base period quantity vector q^0 and the current period quantity vector q^1. The first of these two choices leads to the following *Laspeyres-Konüs true cost of living index*:

$$\begin{aligned} P_K(p^0, p^1, q^0) &\equiv C[f(q^0), p^1]/C[f(q^0), p^0] \\ &= C[f(q^0), p^1]/\sum_{n=1}^{N} p_n^0 q_n^0 \quad \text{using (49) for } t = 0 \qquad (51) \\ &= \min_q \left\{ \sum_{n=1}^{N} p_n^1 q_n : f(q) = f(q^0) \right\} \Big/ \sum_{n=1}^{N} p_n^0 q_n^0 \end{aligned}$$

using the definition of the cost minimization problem that defines $C[f(q^0), p^1]$

$$\leq \sum_{n=1}^{N} p_n^1 q_n^0 \Big/ \sum_{n=1}^{N} p_n^0 q_n^0$$

since $q^0 \equiv (q_1^0, \ldots, q_N^0)$ is feasible for the minimization problem

$$= P_L(p^0, p^1, q^0, q^1)$$

where P_L is the Laspeyres price index defined by (8) above. Thus the (unobservable) Laspeyres-Konüs true cost of living index is bounded from above by the observable Laspeyres price index.[72]

The second of the two natural choices for a reference quantity vector q in definition (50) leads to the *following Paasche-Konüs true cost of living index*:

72. This inequality was first obtained by KONÜS (1924; 1939, p. 17). See also POLLAK (1983).

$$
\begin{aligned}
P_K(p^0, p^1, q^1) &\equiv C[f(q^1), p^1]/C[f(q^1), p^0] \\
&= \sum_{n=1}^{N} p_n^1 q_n^1 / C[f(q^1), p^0] \quad \text{using (49) for } t = 1 \qquad (52) \\
&= \sum_{n=1}^{N} p_n^1 q_n^1 / \min_q \left\{ \sum_{n=1}^{N} p_n^0 q_n : f(q) = f(q^1) \right\}
\end{aligned}
$$

using the definition of the cost minimization problem that defines $C[f(q^1), p^0]$

$$
\geq \sum_{n=1}^{N} p_n^1 q_n^1 / \sum_{n=1}^{N} p_n^0 q_n^1
$$

since $q^1 \equiv (q_1^1, \ldots, q_n^1)$ is feasible for the minimization problem and thus

$$
C[f(q^1), p^0] \leq \sum_{n=1}^{N} p_n^0 q_n^1 \text{ and hence } 1/C[f(q^1), p^0] \geq 1/\sum_{n=1}^{N} p_n^0 q_n^1
$$

$$
= P_P(p^0, p^1, q^0, q^1)
$$

where P_P is the Paasche price index defined by (9) above. *Thus the (unobservable) Paasche-Konüs true cost of living index is bounded from below by the observable Paasche price index.*[73]

The inequality (51) shows that the Laspeyres price index P_L has a nonnegative *substitution bias* relative to the true cost of living index, $P_K(p^0, p^1, q^0)$ while the inequality (52) shows that the Paasche index P_P has a nonpositive *substitution bias* relative to the true cost of living index, $P_K(p^0, p^1, q^1)$. Thus the Laspeyres index will generally have an *upward bias* relative to a cost of living index while the Paasche index will generally have a *downward bias* relative to a cost of living index.

The above inequalities are independent of the functional form for the consumer's utility function $f(q)$ or the corresponding cost function $C(u, p)$. To make further progress, it is necessary to make specific functional form assumptions about f or C.

Suppose that the consumer's *cost function*, $C(u, p)$, has the following *translog functional form*:[74]

$$
\begin{aligned}
\ln C(u, p) \equiv a_0 &+ \sum_{n=1}^{N} a_n \ln p_n + (1/2) \sum_{n=1}^{N} \sum_{k=1}^{N} a_{nk} \ln p_n \ln p_k \\
&+ b_0 \ln u + \sum_{n=1}^{N} b_n \ln p_n \ln u + (1/2) b_{00} [\ln u]^2
\end{aligned} \qquad (53)
$$

73. This inequality is also due to KONÜS (1924) (1939; p. 19). See also POLLAK (1983).
74. CHRISTENSEN, JORGENSON and LAU (1971) introduced this function into the economics literature.

where ln is the natural logarithm function and the parameters a_n, a_{nk}, and b_n satisfy the following restrictions:

$$a_{nk} = a_{kn}; \quad n, k = 1, \ldots, N; \tag{54}$$

$$\sum_{n=1}^{N} a_n = 1; \tag{55}$$

$$\sum_{n=1}^{N} b_n = 0; \tag{56}$$

$$\sum_{k=1}^{N} a_{nk} = 0; \quad n = 1, \ldots, N. \tag{57}$$

The parameter restrictions (54)-(57) ensure that $C(u,p)$ defined by (53) is linearly homogeneous in p, a property that a cost function must have. It can be shown that the translog cost function defined by (53)-(57) can provide a second order Taylor series approximation to an arbitrary cost function[75] and thus it is a *flexible functional form.*

We assume that the consumer has preferences $u = f(q)$ that correspond to the translog cost function and that the consumer engages in cost minimizing behavior during periods 0 and 1 so that (49) holds. Define the geometric average of the period 0 and 1 utility levels as u^*; i.e., define $u^0 \equiv f(q^0), u^1 \equiv f(q^1)$ and

$$u^* \equiv [u^0 u^1]^{1/2}. \tag{58}$$

DIEWERT (1976, p. 122) showed that under the above assumptions, the Törnqvist Theil index number formula P_T defined earlier by (42) is *exactly* equal to the KONÜS true cost of living index defined by (52) where the reference level of utility is u^*, the geometric average of the consumer's period 0 and 1 utility levels; i.e., we have:

$$P_K(u^*, p^0, p^1) \equiv C(u^*, p^1)/C(u^*, p^0) = P_T(p^0, p^1, q^0, q^1). \tag{59}$$

Since the translog cost function defined by (53)-(57) is a flexible functional form, the Törnqvist-Theil price index $\mathrm{P_T}$ is also a *superlative index.*[76]

75. It can also be shown that if all of the $b_i = 0$ and $b_{00} = 0$, then $C(u,p) = uC(1,p) \equiv uc(p)$; i.e., with these additional restrictions on the parameters of the general translog cost function, we have homothetic preferences. Note that we also assume that utility u is scaled so that u is always positive.

76. DIEWERT (1976; p. 117) termed a quantity index $Q(p^0, p^1, q^0, q^1)$ to be superlative if it was exactly equal to $f(q^1)/f(q^0)$ where f can approximate an arbitrary linearly homogeneous utility function to the second order. DIEWERT (1976; p. 134) also termed a price index $P(p^0, p^1, q^0, q^1)$ to be superlative if it was exactly equal to $c(p^1)/c(p^0)$ where c can approximate an arbitrary linearly homogeneous unit cost function to the second order.

For the remainder of this section, we assume that the consumer's utility function f is (positively) *linearly homogeneous.* In the economics literature, this is known as the assumption of *homothetic preferences.* Under this assumption, the consumer's expenditure or cost function, $C(u, p)$ defined by (49) above, decomposes as follows. For positive commodity prices $p \gg 0_n$ and a positive utility level u, we have by the definition of C as the minimum cost of achieving the given utility level u:

$$\begin{aligned} C(u,p) &\equiv \min_q \left\{ \sum_{n=1}^{N} p_n q_n : f(q_1, \ldots, q_N) \geq u \right\} && (60) \\ &= \min_q \left\{ \sum_{n=1}^{N} p_n q_n : (1/u) f(q_1, \ldots, q_N) \geq 1 \right\} \text{ dividing by } u > 0 \\ &= \min_q \left\{ \sum_{n=1}^{N} p_n q_n : f(q_1/u, \ldots, q_n/u) \geq 1 \right\} \text{ using the linear homogeneity of } f \\ &= u \min_q \left\{ \sum_{n=1}^{N} p_n q_n/u : f(q_1/u, \ldots, q_N/u) \geq 1 \right\} \\ &= u \min_z \left\{ \sum_{n=1}^{N} p_n z_n : f(z_1, \ldots, z_N) \geq 1 \right\} \text{ letting } z_n = q_n/u \\ &= uC(1, p) = uc(p) \end{aligned}$$

where $c(p) \equiv C(1, p)$ is the *unit cost function* that is corresponds to f.[77] It can be shown that the unit cost function $c(p)$ satisfies the same regularity conditions that f satisfied; i.e., $c(p)$ is positive, concave and (positively) linearly homogeneous for positive price vectors.[78]

We drop the assumption that the cost function is translog but we continue to assume that the consumer minimizes the cost of achieving the period 0 and 1 utility levels so that equations (49) continue to hold. Substituting (60) into (49) and using $u^t = f(q^t)$ for $t = 0, 1$ leads to the following equations:

$$\sum_{n=1}^{N} p_n^t q_n^t = c(p^t) f(q^t) \qquad \text{for } t = 0, 1. \tag{61}$$

77. Economists will recognize the producer theory counterpart to the result $C(u, p) = uc(p)$: if a producer's production function f is subject to constant returns to scale, then the corresponding total cost function $C(u, p)$ is equal to the product of the output level u times the unit cost $c(p)$.

78. Obviously, the utility function f determines the consumer's cost function $C(u, p)$ as the solution to the cost minimization problem in the first line of (61). Then the unit cost function $c(p)$ is defined as $C(1, p)$. Thus f determines c. But we can also use c to determine f under appropriate regularity conditions. In the economics literature, this is known as duality theory. For additional material on duality theory and the properties of f and c, see SAMUELSON (1953), SHEPHARD (1953) and DIEWERT (1974a; 1993c, p. 107–123).

Thus under the linear homogeneity assumption on the utility function f, observed period t expenditure on the N commodities (the left hand side of (61) above) is equal to the period t unit cost $c(p^t)$ of achieving one unit of utility times the period t utility level, $f(q^t)$, (the right hand side of (61) above). Obviously, we can identify the period t unit cost, $c(p^t)$, as the period t price level P^t and the period t level of utility, $f(q^t)$, as the period t quantity level Q^t.[79]

The linear homogeneity assumption on the consumer's preference function f leads to a simplification for the family of KONÜS true cost of living indices, $P_K(p^0, p^1, q)$, defined by (50) above. Using this definition for an arbitrary reference quantity vector q, we have:

$$\begin{aligned} P_K(p^0, p^1, q) &\equiv C[f(q), p^1]/C[f(q), p^0] \\ &= c(p^1)f(q)/c(p^0)f(q) \quad \text{using (61)} \\ &= c(p^1)/c(p^0). \end{aligned} \tag{62}$$

Thus under the homothetic preferences assumption, the entire family of KONÜS true cost of living indices collapses to a single index, $c(p^1)/c(p^0)$, the ratio of the minimum costs of achieving unit utility level when the consumer faces period 1 and 0 prices respectively. Put another way, under the homothetic preferences assumption, $P_K(p^0, p^1, q)$ is independent of the reference quantity vector q.

If we use the KONÜS true cost of living index defined by the right hand side of (62) as our price index concept, then the corresponding implicit quantity index defined using the product test (2) has the following form:

$$\begin{aligned} Q(p^0, p^1, q^0, q^1) &\equiv \sum_{n=1}^{N} p_n^1 q_n^1 \Big/ \left\{ \sum_{n=1}^{N} p_n^0 q_n^0 P_K(p^0, p^1, q) \right\} \quad \text{using (2)} \\ &= c(p^1)f(q^1)/\{c(p^0)f(q^0)P_K(p^0, p^1, q)\} \quad \text{using (61) twice} \\ &= c(p^1)f(q^1)/\{c(p^0)f(q^0)[c(p^1)/c(p^0)]\} \quad \text{using (62)} \\ &= f(q^1)/f(q^0). \end{aligned} \tag{63}$$

Thus under the homothetic preferences assumption, the *implicit quantity index* that corresponds to the true cost of living price index $c(p^1)/c(p^0)$ is the *utility ratio* $f(q^1)/f(q^0)$. Since the utility function is assumed to be homogeneous of degree one, this is the natural definition for a quantity index.

In addition to the Törnqvist Theil price index P_T, it turns out that there are many other superlative index number formulae; i.e., there exist many quantity indices $Q(p^0, p^1,$

79. There is also a producer theory interpretation of the above theory; i.e., let f be the producer's (constant returns to scale) production function, let p be a vector of input prices that the producer faces, let q be an input vector and let $u = f(q)$ be the maximum output that can be produced using the input vector q. $C(u, p) \equiv \min_q \left\{ \sum_{i=1}^{N} p_n q_n : f(q) \geq u \right\}$ is the producer's cost function in this case and $c(p^t)$ can be identified as the period t input price level while $f(q^t)$ is the period t aggregate input level.

$q^0, q^1)$ that are exactly equal to $f(q^1)/f(q^0)$ and many price indices $P(p^0, p^1, q^0, q^1)$ that are exactly equal to $c(p^1)/c(p^0)$ where the aggregator function f or the unit cost function c is a flexible functional form. We will define two families of superlative indices below.

Suppose the consumer has the following quadratic mean of order r utility function:[80]

$$f^r(q_1, \ldots, q_N) \equiv \left[\sum_{i=1}^{N}\sum_{k=1}^{N} a_{ik} q_i^{r/2} q_k^{r/2}\right]^{1/r} \tag{64}$$

where the parameters a_{ik} satisfy the symmetry conditions $a_{ik} = a_{ki}$ for all i and k and the parameter r satisfies the restriction $r \neq 0$. DIEWERT (1976, p. 130) showed that the utility function f^r defined by (64) is a flexible functional form; i.e., it can approximate an arbitrary twice continuously differentiable linearly homogeneous functional form to the second order.

Define the quadratic mean of order r quantity index Q^r by:

$$Q^r(p^0, p^1, q^0, q^1) \equiv \left\{\sum_{i=1}^{N} s_i^0 (q_i^1/q_i^0)^{r/2}\right\}^{1/r} \left\{\sum_{i=1}^{N} s_i^1 (q_i^1/q_i^0)^{-r/2}\right\}^{-1/r} \tag{65}$$

where $s_i^t \equiv p_i^t q_i^t / \sum_{k=1}^{N} p_k^t q_k^t$ is the period t expenditure share for commodity i as usual. It can be verified that when $r = 2$, Q^r simplifies into Q_F, the FISHER (1922) ideal quantity index.

DIEWERT (1976, p. 132) showed that Q^r is exact for the aggregator function f^r defined by (64); i.e., we have

$$Q^r(p^0, p^1, q^0, q^1) = f^r(q^1)/f^r(q^0). \tag{66}$$

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the N commodities that correspond to the utility function defined by (64), the quadratic mean of order r quantity index Q_F is *exactly* equal to the true quantity index, $f^r(q^1)/f^r(q^0)$. Since Q^r is exact for f^r and f^r is a flexible functional form, we see that the quadratic mean of order r quantity index Q^r is a *superlative index* for each $r \neq 0$. Thus there are an infinite number of superlative quantity indices.

For each quantity index Q^r, we can use the identity (2) in order to define the corresponding *implicit quadratic mean of order r price index Pr*∗:

$$\begin{aligned} P^{r*}(p^0, p^1, q^0, q^1) &\equiv \sum_{i=1}^{N} p_i^1 q_i^1 \Big/ \left\{\sum_{i=1}^{N} p_i^0 q_i^0 Q^r(p^0, p^1, q^0, q^1)\right\} \\ &= c^{r*}(p^1)/c^{r*}(p^0) \end{aligned} \tag{67}$$

80. This terminology is due to DIEWERT (1976; 129).

where c^{r*} is the unit cost function that corresponds to the aggregator function f^r defined by (64) above. For each $r \neq 0$, the implicit quadratic mean of order r price index P^{r*} is also a superlative index.

When $r = 2$, Q^r defined by (65) simplifies to Q_F, the Fisher ideal quantity index and P^{r*} defined by (67) simplifies to P_F, the Fisher ideal price index. When $r = 1$, Q^r defined by (65) simplifies to:

$$
\begin{aligned}
Q^1(p^0,p^1,q^0,q^1) &\equiv \left\{\sum_{i=1}^N s_i^0(q_i^1/q_i^0)^{1/2}\right\}\Big/\left\{\sum_{i=1}^N s_i^1(q_i^1/q_i^0)^{-1/2}\right\} \qquad (68)\\
&= \left[\sum_{i=1}^N p_i^1q_i^1\Big/\sum_{i=1}^N p_i^0q_i^0\right]\left\{\sum_{i=1}^N p_i^0q_i^0(q_i^1/q_i^0)^{1/2}\right\}\Big/\left\{\sum_{i=1}^N p_i^1q_i^1(q_i^1/q_i^0)^{-1/2}\right\}\\
&= \left[\sum_{i=1}^N p_i^1q_i^1\Big/\sum_{i=1}^N p_i^0q_i^0\right]\left\{\sum_{i=1}^N p_i^0(q_i^0q_i^1)^{1/2}\right\}\Big/\left\{\sum_{i=1}^N p_i^1(q_i^0q_i^1)^{1/2}\right\}\\
&= \left[\sum_{i=1}^N p_i^1q_i^1\Big/\sum_{i=1}^N p_i^0q_i^0\right]\Big/\left\{\sum_{i=1}^N p_i^1(q_i^0q_i^1)^{1/2}\Big/\sum_{i=1}^N p_i^0(q_i^0q_i^1)^{1/2}\right\}\\
&= \left[\sum_{i=1}^N p_i^1q_i^1\Big/\sum_{i=1}^N p_i^0q_i^0\right]\Big/P_W(p^0,p^1,q^0,q^1)
\end{aligned}
$$

where P_W is the *Walsh price index* defined previously by (21). Thus P^{1*} is equal to P_W, the *Walsh price index*, and hence it is also a superlative price index.

Suppose the consumer has the following quadratic mean of order r unit cost function:[81]

$$c^r(p_1,\ldots,p_N) \equiv \left[\sum_{i=1}^N\sum_{k=1}^N b_{ik}p_i^{r/2}p_k^{r/2}\right]^{1/r} \qquad (69)$$

where the parameters b_{ik} satisfy the symmetry conditions $b_{ik} = b_{ki}$ for all i and k and the parameter r satisfies the restriction $r \neq 0$. DIEWERT (1976, p. 130) showed that the unit cost function c^r defined by (69) is a flexible functional form; i.e., it can approximate an arbitrary twice continuously differentiable linearly homogeneous functional form to the second order.

Define the quadratic mean of order r price index P^r by:

$$P^r(p^0,p^1,q^0,q^1) \equiv \left\{\sum_{i=1}^N s_i^0(p_i^1/p_i^0)^{r/2}\right\}^{1/r}\left\{\sum_{i=1}^N s_i^1(p_i^1/p_i^0)^{-r/2}\right\}^{-1/r} \qquad (70)$$

where $s_i^t \equiv p_i^tq_i^t/\sum_{k=1}^N p_k^tq_k^t$ is the period t expenditure share for commodity i as usual. It can be verified that when $r = 2$, P^r simplifies into P_F, the Fisher ideal price index defined by (15) above.

81. This terminology is due to DIEWERT (1976; p. 130). This unit cost function was first defined by DENNY (1974).

DIEWERT (1976, p. 134) showed that P^r is exact for the unit cost function c^r defined by (69); i.e., we have

$$P^r(p^0, p^1, q^0, q^1) = c^r(p^1)/c^r(p^0). \tag{71}$$

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the N commodities that correspond to the unit cost function defined by (69), the quadratic mean of order r price index P_F is *exactly* equal to the true price index, $c^r(p^1)/c^r(p^0)$. Since P^r is exact for c^r and c^r is a flexible functional form, we see that the quadratic mean of order r price index P^r is a *superlative index* for each $r \neq 0$. Thus there are an infinite number of superlative price indices.

For each price index P^r, we can use the identity (2) in order to define the corresponding *implicit quadratic mean of order r quantity index* Q^{r*}:

$$\begin{aligned} Q^{r*}(p^0, p^1, q^0, q^1) &\equiv \sum_{i=1}^{N} p_i^1 q_i^1 \Big/ \left\{ \sum_{i=1}^{N} p_i^0 q_i^0 P^r(p^0, p^1, q^0, q^1) \right\} \\ &= f^{r*}(p^1)/f^{r*}(p^0) \end{aligned} \tag{72}$$

where f^{r*} is the aggregator function that corresponds to the unit cost function c^r defined by (69) above. For each $r \neq 0$, the implicit quadratic mean of order r quantity index Q^{r*} is also a superlative index.

When $r = 2$, P^r defined by (70) simplifies to P_F, the Fisher ideal price index and Q^{r*} defined by (72) simplifies to Q_F, the Fisher ideal quantity index.

We may sum up the results of this subsection as follows: the WALSH price index P_W defined by (21), the Fisher ideal index P_F defined by (15) and the Törnqvist Theil index number formula P_T defined by (42) can all be regarded as being equally desirable from the viewpoint of the economic approach to index number theory.

5.5. *Summing up the Results*

Economists and price statisticians have studied the problem of choosing a functional form for the index number formula over the past century from a number of different perspectives. In the previous subsections, we looked at 4 different approaches that led to specific index number formulae as being "best" from each perspective. From the viewpoint of fixed basket approaches, we found that the Fisher and Walsh price indexes, P_F and P_W, defined by (15) and (21) appeared to be "best". From the viewpoint of the test approach, the Fisher index appeared to be "best". From the viewpoint of the stochastic approach to index number theory, the Törnqvist Theil index number formula P_T defined by (42) emerged as being "best". Finally, from the viewpoint of the economic approach to index number theory, the Walsh price index P_W, the Fisher ideal index P_F and the Törnqvist Theil index number formula P_T were all be regarded as being equally desirable (along with two entire families of price indexes). What is amazing is that the same three index number formulae emerge as being "best" from very different perspectives! What is even better is that for normal time series data, the three indexes (Walsh, Fisher, and Törnqvist Theil) all approximate each other very closely and so it will not matter very much which of these alternative indexes is chosen.[82]

This completes the first part of this paper where we have discussed possible transactions domains of definition for a target index of inflation and alternative index number concepts that could be used once the domain of definition has been chosen. We now turn our attention to the properties of the harmonized index of consumer prices (HICP), which each member state of the European Union is required to produce, starting in January, 1997.

6. THE CONCEPTUAL FOUNDATIONS OF THE HARMONIZED INDEX OF CONSUMER PRICES

6.1. The Properties of the HICP

The papers by ASTIN (1999) and BERGLUND (1999) lay out the main features of the Harmonized Index of Consumer Prices (HICP) published by Eurostat. In their view, the HICP should have the following properties:

(a) It should encompass only market transactions[83]; i.e., *imputations* such as user costs or imputed rental prices for owner occupied housing would not be included.[84]
(b) It should not include interest rates[85] or interest costs since "such costs are neither a good or a service but the instrument for balancing the supply and demand of money" (BERGLUND, 1999, p. 69).
(c) The index should treat owner occupied housing in one of two ways: either exclude owner occupied housing from the index or to include new purchases of dwelling units[86], essentially treating purchases of new dwelling units like any other purchase of a consumer durable.

82. Theorem 5 in DIEWERT (1978; p. 888) shows that P_F, P_T and P_W will approximate each other to the second order around an equal price and quantity point; see DIEWERT (1978; p. 894) and HILL (2000) for some empirical results.
83. See ASTIN (1999; p. 3).
84. "Firstly, the harmonized indices would be concerned only with actual monetary transactions. So, for example, in the area of housing costs, we would not use the imputed rents method to measure the 'price' of owner-occupied housing (such a method is motivated in the measurement of the volume of consumption of housing services, but is irrelevant in the context of measurement of price change)." BERGLUND (1999; p. 69).
85. See also ASTIN (1999; p. 3).
86. See ASTIN (1999; p. 5). At present, the HICP excludes both the services of owner occupied dwellings and purchases of new dwellings. "However, consideration is at present being given to a future possible inclusion of the net acquisition prices of new dwellings." BERGLUND (1999; p. 71).

(d) The harmonized index should use the Laspeyres formula but the basket must be updated between one and seven years with a preference for more frequent reweighting.[87]

(e) "Expenditure incurred for business purposes should be excluded." (BERGLUND, 1999, p. 72).

(f) The harmonized CPI for a country should include the consumption expenditures made by foreign visitors and exclude the expenditure by residents while visiting in a foreign country.[88]

(g) The prices, which should be used in the HICP, are consumer prices (or final demand prices) rather than producer prices.[89] Thus harmonized prices should include commodity and value added taxes in principle.

(h) The prices of highly subsidized consumer goods and services should be the prices faced by the consumer; i.e., prices after the subsidies.[90]

An implication of (a) is that the HICP should not include *new commodities* in the domain of definition of the price index (until there is a sample rotation); i.e., if a commodity is present in one of the two periods being compared but not the other, then that commodity should be excluded from the price index. This point follows from (a) because a new commodity cannot be matched precisely to a corresponding commodity in the previous period and hence an *imputation* must be made to the price of the new commodity in order to adjust for quality change. However, in his remarks on the Panel Discussion on Quality Adjustment at this Conference, JOHN ASTIN made clear that the HICP has undertaken quality adjustment of prices in a systematic way.

We turn now to a discussion of properties (a) to (h) above for a harmonized price index.[91] These 7 properties for the HICP enable us to distinguish it from a cost of living (COL) index or from a producer price index (PPI) based on producer theory.[92] The HICP shares properties (e), (g) and (h) with a COL and the HICP also shares properties (a), (b) and (c) (if new houses are treated like other consumer durables) with a PPI.[93] However properties (a)-(c) are not consistent with a COL index, which should use either a rental equivalence approach to the consumption of housing services or a user cost ap-

87. See ASTIN (1999; p. 3) and BERGLUND (1999; p. 70). BERGLUND (1999; p. 70) notes that the Paasche formula is equally valid from a theoretical perspective but its use "is ruled out on practical grounds".
88. See ASTIN (1999; p. 6) and BERGLUND (1999; p. 72).
89. According to BERGLUND (1999; p. 70), the HICP "shall be based on the price of goods and services available for purchase in the economic territory of the Member State for the purposes of directly satisfying consumer needs."
90. ASTIN (1999; p. 4).
91. None of the discussion that follows should detract from the fact that national and Eurostat statisticians got the HICP up and running in a very short time to fulfill a critical need. This is a tremendous accomplishment.
92. For the economic approach to output price indexes, see FISHER and SHELL (1972), SAMUELSON and SWAMY (1974), ARCHIBALD (1977), DIEWERT (1980, 1983b, 2001a), CAVES, CHRISTENSEN and DIEWERT (1982), BALK (1998), and ALTERMAN, DIEWERT and FEENSTRA (1999).
93. Sales of new dwelling units would be included in a PPI.

proach to the consumption of owner-occupied housing services. In the following section, we will look at alternative treatments of housing in more detail.

Thus from the viewpoint of economic theory[94], a HICP is not clearly based on consumer theory (which leads to a COL) or on producer theory (which leads to a PPI): *it is a mixture of a consumer and producer price index for consumption expenditures.*

However, it should be noted that it is not easy to define and implement either a cost of living index or a producer price index that covers consumer expenditures in a consistent fashion. The problem with implementing a COL is that households engage in household production such as work at home, home renovations, etc. In theory, all of the services and materials that households buy as inputs into their household production functions should be segregated from their consumption purchases and be placed in a set of household production accounts. In practice, it is very difficult to determine whether a specific household purchase is an intermediate input or a final demand purchase. Similar measurement problems apply to the construction of a Producer Price Index that covers business sales of consumer goods and services to households for final demand. The problem is that some fraction of business sales of "consumer commodities" will be to other business units who use them as intermediate inputs but it is very difficult to determine these fractions commodity by commodity.

As noted in point (d) above, the HICP is based on the use of a Laspeyres formula; i.e., the base period quantity basket is repriced over time. It is argued that this must be done on practical grounds. However, this argument is not completely convincing since SHAPIRO and WILCOX (1997) have shown that the LLOYD (1975) MOULTON (1996) formula can be used to form a close approximation to a superlative index like the Fisher (1922) ideal or the Törnqvist price indexes defined earlier in this paper. The Lloyd-Moulton formula makes use of the same information set as the usual Laspeyres index except that an estimate of the elasticity of substitution between the various commodities must be provided to the statistical agency.[95] Finally, even if superlative indexes cannot be produced in as timely a fashion as the Laspeyres index, they can be produced with a lag as new household expenditure data or new national accounts data become available. Producing a superlative index on a delayed basis would surely give some useful information to the central bank as to the probable extent of upper level substitution bias.

We will conclude this section by discussing points (b), (f) and (h) in a bit more detail and then in the following section, we will look at some of the problems involved in providing a firmer methodological base for the HICP.

94. Of course, it must be kept in mind that the European Union did not introduce its harmonized index of consumer prices in order to be consistent with economic theory. Initially, HICPs were introduced because the convergence criteria in the Maastricht Treaty required that price inflation be measured comparably across EU countries. Subsequently, the HICP was given a prominent role in the ECB'S quantitative definition of price stability for the simple reason that is was (and still remains) the best available measure.

95. More recently, some recent results due to SCHULTZ (1999) and OKAMOTO (2001) show how various *midyear price indexes* can approximate superlative indexes fairly closely under certain conditions. These midyear indexes also do not require information on current quantities or expenditures.

6.2. *Imputations and the Treatment of Interest*

"In practice, 'inflation' is what happens to be the index used to measure it! We decided at an early stage that inflation is essentially a monetary phenomenon. It concerns the changing power of money to produce goods and services. This led us down two important paths. Firstly, the HICPs would be concerned only with actual monetary transactions. So, for example, in the field of housing, we would not use the imputed rents method to measure the price of owner-occupied housing. (This is a valuable concept in the context of the measurement of the volume of consumption of housing services, but it is irrelevant in the context of the measurement of price change). Secondly, we would not include the cost of borrowed money, which is neither a good nor a service. So interest payments were to be excluded. This immediately set the HICP apart from some national CPIs which include interest payments on the grounds that they form part of the regular outgoings of households: a perfectly reasonable argument in the context of a compensation index, but less so for an inflation index." JOHN ASTIN (1999, p. 2–3).

Thus a harmonized index can only have actual transactions that took place in the two periods being compared in its domain of definition and there are to be no imputed prices in the index. We have already stressed that the domain of definition problem needs to be very carefully specified. However, the above description of the HICP does not explain why monetary transactions in certain classes of consumption goods were *excluded* from the domain of definition of the HICP; i.e., why were *actual transactions* in second hand houses excluded? On the other hand, in a Cost of Living (COL) approach to the Consumer Price Index (CPI), the consumption of owner-occupied housing would be valued according to a *rental equivalence approach* or a *user cost approach*. In the rental equivalence approach, the services of an owner-occupied home would be valued at a comparable market rental price. It is true that this price would be an *imputed* or *estimated* one but is this a very different procedure from say *estimating* the aggregate price of television sets in a country from 30 representative price quotes? It is true that homes are a more complex product but it seems to me that the two estimation or imputation situations are not all that different. On the other hand, in the user cost approach to the purchase of a consumer durable, it is explicitly recognized that *not all of the good is consumed in the period of purchase*. Thus the purchase price should be decomposed into two parts: the first part which is the cost to the consumer of using the *services of the commodity* during the period of purchase, and a second part, which is a form of *investment* that will yield either a return or services to the consumer in future periods. Moreover, the user cost approach provides us with a way of valuing the services of the older vintages of household consumer durable goods and thus allows us to build up a more comprehensive picture of actual household consumption as opposed to the money purchases approach advocated for the HICP, which includes only *new* purchases of consumer durables. In order to estimate these user costs, it is necessary to have information on the prices of used consumer durables at the beginning and end of each period. Thus one could argue that the user cost approach uses *more* information on actual asset transactions than the money purchases or acquisitions approach to the treatment of durables. We will return to a more technical discussion of these alternative approaches in the following section.

This is perhaps not the appropriate place to get into an extensive discussion of the role of interest in economics but many economists would be somewhat puzzled at the meaning of the statement that interest is the cost of borrowed money and hence is not a good or a service. Most economists would regard interest as the payment for the use of financial capital for a specified period of time and hence regard it as a service. Hence interest is a price just like any other price: it is the price a borrower must pay to a lender for the use of financial capital for a specified time period.[96] However, KEITH WOOLFORD (1999) has suggested an interesting reason for the possible exclusion of interest from a price index. Namely, interest is not a *contemporaneous price*; i.e., an interest rate necessarily refers to *two* points in time; a beginning point when the capital is loaned and an ending point when the capital loaned must be repaid. Thus if for some reason, one wanted to restrict attention to a domain of definition that consisted of only contemporaneous prices, interest rates would be excluded. However, interest rates *are* prices (even though they are more complex than contemporaneous prices).

6.3. *The Treatment of Nonmarket or Highly Subsidized Services*

"In most cases goods and services on the market are sold at a price determined by normal market processes. But in several important sectors, especially healthcare and education, it is common to have partial or total subsidies provided by the state. This raises difficult problems in CPI construction, regarding both concept and measurement.

Some experts argued that the full, unsubsidised, price of such products should be included...

Others argued that the HICP does not aim to measure total inflation, but just that part impacting on the private household sector...

The solution finally adopted owes much to the work of PETER HILL. He showed that within the ESA [European System of Accounts] structure it was possible to define an element of expenditure, which he named HFMCE, which related solely to that part of the expenditure actually paid by private households. So that, for example, if 80% of a

96. One of the first economists to realize that interest was an intertemporal price and analogous to an exchange rate that compares the price of a currency in one location with another currency in a different location was the Italian monsignore and civil servant FERDINANDO GALIANI (1751; p. 303): "Hence arose exchange and interest, which are brothers. One is the equalizing of present money and money distant in space, made by an apparent premium, which is sometimes added to the present money, and sometimes to the distant money, to make the intrinsic value of both equal, diminished by the less convenience or the greater risk. Interest is the same thing done between present money and money that is distant in time, time having the same effect as space; and the basis of the one contract, as of the other, is the equality of the true intrinsic value."

chemist's prescription charge is reimbursed by the government, only the remaining 20 % would be included in the HICP. A change in the subsidy would have a similar effect on the 'market' price to a change in VAT [Value Added Tax], which, of course, is also included in all CPIs." JOHN ASTIN (1999, p. 4).

The treatment of subsidized goods chosen by the HICP is exactly the right one if our domain of definition is the transactions of *households*, (which is a consumer theory perspective). However, if our domain of definition is the consumer goods and services produced by *firms*, then the treatment is not correct. From this perspective (a producer theory perspective), the "correct" price is the full, unsubsidized price.

6.4. The Geographic Domain of Definition of the Index

Recall point (f) above; i.e., that the harmonized CPI for a country should include the consumption expenditures made by foreign visitors and exclude the expenditure by residents while visiting in a foreign (non EU) country. ASTIN describes the motivation for this treatment as follows:

> "A quite different aspect of HCIPs is the question of geographic coverage. This is a matter of special interest in the EU, given the fact that the Monetary Union (MU) is only a subset of the EU, and is likely to be a subset for some time, as the memberships of both the MU and the EU are likely to increase – at different rates – over the coming years.
>
> At the heart of this question are two concepts well known to national accountants: the domestic concept and the national concept...
>
> In principle, a price statistician has two choices. First, he can choose to measure the changes in prices faced by consumers normally resident in the country – in which case the prices paid by these consumers when they are outside the country also have to be included in the index. This is known as the 'national' concept of measurement.
>
> Alternatively, he can choose to measure the changes in prices faced by all consumers in the country itself – in which case one must measure only domestic prices, but the weights applied must relate to the total consumption within the country, whether by the resident population or by foreign visitors. This is known as the 'domestic' concept of measurement.
>
> There are both theoretical and practical aspects to this question. On a practical level, it would obviously be difficult, if not impossible, for a national price statistician to measure price changes in other countries where consumption is made by residents of his own country. In practice, he would have to use the CPIs of a range of foreign countries – many of which, of course, would not be in the EU.
>
> But theoretically (fortunately) this approach is not called for. National inflation should surely measure national price changes, even if some of them are faced by foreign visitors." JOHN ASTIN (1999, p. 6–7).

The transactions domain of definition that is suggested by the above quotation is: all consumer expenditures in the EU, including those of tourists from outside the EU. This domain of definition does not fit neatly into the usual categories. If our domain of definition was EU household consumer expenditures, then tourist expenditures by EU residents made in non EU countries should be *included* and the tourist expenditures of non EU residents in the EU should be *excluded*. If we took a producer theory perspective to the domain of definition, then all sales of consumer commodities made by EU suppliers should be included and this domain of definition would almost coincide with the HICP domain.[97]

How could the above problems be resolved in a way that would make the HICP domain of definition fit into the usual national accounts categories? A straightforward way of proceeding would be to make the domain of definition household expenditures of EU residents on consumer goods and services but excluding tourist expenditures in non EU countries.[98] A national CPI would add tourist expenditures in non EU countries back into its domain of definition.[99] Tourist expenditures made by non EU residents in the national country would not appear in either the national CPI or the revised HICP.[100]

6.5. *Conclusions*

We summarize the above discussion as follows. The "theory" of the Harmonized Index of Consumer Prices seems to lack an underlying firm theoretical basis. Evidently, its primary purpose is as a measure of inflation that is based on actual transactions that use money. However, as we argued in section 2 above, a measure of inflation based on "monetary" transactions is too broad to be useful. Thus when the inflation measurement goal of the harmonized index is narrowed down to focus on purchases of consumer goods and services in the economic territory of the Member State, the "general theory" of the HICP does not constrain the index as much as an explicit producer or consumer theory approach would. As a result, the HICP does not fit into either the consumer or producer domains of definition. *Thus the HICP introduces a third class of index numbers, which is a mixture of consumer and producer price indexes.* This class of index numbers can be contrasted with the two classes of consumption price indexes that emerge from the national accounts framework (and economic theory). One member of this family would look at the *consumption transactions of households* (a consumer theory approach)

97. Unfortunately, EU suppliers of consumer goods and services also sell these commodities to other EU producers as intermediate inputs so the HICP domain of definition is not quite a producer domain either.
98. Thus for HICP purposes, a national consumer expenditures survey would collect information on tourist expenditures of national residents in the usual way. These tourist expenditures would be further classified according to whether they were made in a non EU country or an EU country. The HICP would use only the information on the latter class of expenditures whereas a national CPI would use the information on both classes of tourist expenditures in forming base period expenditure shares.
99. Of course, there are practical difficulties in collecting (foreign) prices for these tourist expenditures by the national statistical agency. However, there are difficulties of a similar nature in the present HICP: how can EU national price statisticians obtain information on the expenditures of non EU tourists in their country?
100. However, these expenditures would appear in the country's PPI and in the country's export price index.

and another branch of the family would look at *the domestic production by firms of consumer goods and services* (a producer theory approach). This suggested dual approach to index number theory would help fill out the boxes in the *System of National Accounts: 1993*, where there are *basic prices* (which correspond roughly to producer prices) and *final demand prices* (which correspond to consumer prices in the case of the household components of final demand).[101]

The above remarks on the lack of complete consistency of the harmonized price indexes are not meant to denigrate the accomplishments of the price statisticians who got the HICP up and running. After all, they faced many time and political constraints and did the best job that they could in a very short time. Moreover, it is difficult to construct a completely consistent index of consumer prices no matter what methodology one uses as the starting point.

In the following section, we will offer some suggestions on how the HICP might be put on a more consistent theoretical foundation.

7. DISCUSSION OF THE PROBLEM AREAS IN CONSTRUCTING A CONSUMER PRICE INDEX

7.1. The Treatment of Quality Change and New Commodities

As we have seen in the previous section, the HICP is basically a fixed base Laspeyres type price index where the base must be changed at least every 10 years but with a preference for more frequent rebasing. The advantage of this methodological approach is its simplicity and ease of explanation. In the base period, expenditure shares for 100 or so basic commodity classes are estimated and a sample of representative items is chosen for each of these basic commodity classes. These items are priced every month, long term price relatives (relating the current month price to the corresponding base period price) are calculated and then averaged for each commodity class (we will discuss exactly how these item price relatives are to be averaged in section 7.4 below) and then these long term "average" price relatives are inserted into the Laspeyres formula, (11) above. Everything seems quite straightforward.

However, as we saw in the paper by Heravi and Silver (2001d), even for a relatively simple commodity like a washing machine, by the end of the year 1998, about 50 % of

101. There are some problems with the *System*'s methodology on the producer side; e.g., there is no user cost methodology for capital input, the role of interest is not completely recognized, the role of land, natural resources and inventories as inputs is not recognized and so on. On the consumer side, the user cost or rental equivalence approach to consumer durables is ruled out except for housing services. There is also a reluctance to make any imputations associated with the introduction of new commodities. However, the next revision of the Accounts will surely deal with these problems.

the washing machines sold in the UK were not available at the beginning of the year! This is not atypical of the type of sample degradation that occurs in modern economies: the U.S. Bureau of Labor Statistics (1984, p. 13) has estimated that approximately 3 % of the price quotes it collected in the previous month are no longer available in the following month. This rapid rate of disappearance of old goods and the rapid introduction of new goods and services creates tremendous methodological problems for the fixed base Laspeyres price index: *a substantial fraction of items simply cannot be matched exactly for more than a few months at a time.*

How can we deal with this lack of item matching on a conceptual level? The obvious answer is to *use hedonic regression methods*[102] to *quality adjust* every item in a basic commodity classification into units of a "standard" item in the classification. Then as an item disappeared, the missing price could immediately be replaced by another (quality adjusted) item price. Of course there is a very big cost in doing these hedonic regressions and there will be some loss of reproducibility and objectivity because the various "operators" of the hedonic regressions will not always end up with exactly the same quality adjustments. However, conceptually, I do not see any other way of overcoming the problem of rapid sample degradation.

Note that once the quality adjustment is done, there is no conceptual problem with applying the fixed base Laspeyres methodology using the universe of quality adjusted prices in each basic commodity classification.

7.2. Substitution Bias or Representativity Bias

Substitution bias is the difference between a cost of living index and the corresponding Laspeyres or Paasche price indexes.[103] These latter two fixed basket indexes essentially assume that the consumer does not substitute away from commodities that have become more expensive going from one period to another.

Now it might be thought that substitution bias is not a relevant consideration if we adopt a fixed basket approach to price measurement; i.e., substitution bias arises *only* in the context of the economic approach to index number theory and we are not obliged to adopt the economic approach. However, in section 5.1 above, we argued that initially, there were two "natural" fixed basket indexes to choose from in making price comparisons between two periods: the Laspeyres and Paasche indexes. We also argued that if

102. See Silver (1999), Silver and Heravi (2001a, 2001b, 2001c, 2001d), Diewert (2001c) and Triplett (1987, 1990, 2002) for recent material on hedonic regression methods. For nonhedonic methods for dealing with new commodities based on the idea of reservation prices, see Hicks (1940, p. 114), Diewert (1980, p. 498–501; 1987, p. 779) and Hausman (1997, 1999).

103. Substitution bias for the Laspeyres price index is the difference between the right and left hand sides of (51) in section 5.4 above while substitution bias for the Paasche price index is the difference between the right and left hand sides of (52).

these two indexes gave different answers, then in order to obtain a single representative estimate of price change between the two periods, it would be necessary to either take an average of the Paasche and Laspeyres estimates or take an average of the two "natural" baskets as a more representative basket. These two approaches led to the Fisher ideal price index P_F defined by (15) and the Walsh price index P_W defined by (21). Put another way, the quantity weights that are used in the Paasche and Laspeyres indexes are representative of only *one* of the two periods under consideration and hence are in general, *not representative of both periods*. Hence, when we adopt the symmetric fixed basket approach to index number theory, we can speak of the Paasche and Laspeyres indexes suffering from *representativity bias* [104] as opposed to *substitution bias*, which is relevant when we adopt the economic approach to index theory.[105]

It will be useful to obtain a rough estimate of the numerical size of the representativity bias of the Laspeyres and Paasche price indexes under some simplifying assumptions. Our first simplifying assumption is to assume that expenditure shares do not change between the two periods. Thus we assume:

$$s_n^0 = s_n^1 \equiv s_n, \qquad n = 1, \ldots, N \tag{73}$$

where the period t expenditure shares s_n^t were defined by (10) above. Assumption (73) will not be satisfied exactly in real life but it will usually be satisfied empirically to a reasonable degree of approximation, at least over short time periods.[106]

Define the *inflation rate for the nth commodity* going from period 0 to 1, i_n, as follows:

$$1 + i_n \equiv p_n^1/p_n^0, \qquad n = 1, \ldots, N. \tag{74}$$

Recall the share formula for the Laspeyres price index, (11) above. Using (73) and (74), we can rewrite this formula as a function of the commodity specific inflation rates as follows:

104. In recent times, the idea that the Paasche and Laspeyres baskets are not representative of *both* periods being compared can be traced to PETER HILL (1998; p. 46): "When inflation has to be measured over a specified sequence of years, such as a decade, a pragmatic solution to the problems raised above would be to take the middle year as the base year. This can be justified on the grounds that the basket of goods and services purchased in the middle year is likely to be much more *representative* of the pattern of consumption over the decade as a whole than baskets purchased in either the first or the last years. Moreover, choosing a more representative basket will also tend to reduce, or even eliminate, any bias in the rate of inflation over the decade as a whole as compared with the increase in the CoL index." Thus in addition to introducing the concept of representativity bias, Hill also introduced the idea of *midyear indexes*, which has also been pursued by SCHULTZ (1999) and OKAMOTO (2001).

105. Thus the representativity bias of the Laspeyres price index is either $P_L - P_F$ or $P_L - P_W$, depending on which symmetric fixed basket approach is preferred. Since typically P_F will be very close to P_W, the difference between $P_L - P_F$ and $P_L - P_W$ will not be material. If the true cost of living can be approximated by either P_F or P_W, then representativity bias is equal to substitution bias.

106. Note that assumption (73) does not mean that the quantity vectors, q^0 and q^1, remain unchanged as prices change. If we take the economic approach to index number theory, assumption (73) means that the consumer has Cobb Douglas preferences; e.g., see DIEWERT (1995a; p. 18).

$$\begin{aligned} P_L(i_1, \ldots, I_N) &\equiv \sum_{n=1}^{N} s_n^0(1 + i_n) \\ &\approx \sum_{n=1}^{N} s_n(1 + i_n) \quad \text{using (73)} \\ &= 1 + \sum_{n=1}^{N} s_n i_n \equiv 1 + i^* \end{aligned} \tag{75}$$

where

$$i^* \equiv \sum_{n=1}^{N} s_n i_n \tag{76}$$

is the weighted sample mean of the individual commodity inflation rates.

Recall the share formula for the Paasche price index, (12) above. Using (73) and (74), we can rewrite this formula as a function of the commodity specific inflation rates as follows:

$$\begin{aligned} P_P(i_1, \ldots, i_N) &\equiv \left[\sum_{n=1}^{N} s_n^1(1 + i_n)^{-1}\right]^{-1} \\ &\approx \left[\sum_{n=1}^{N} s_n(1 + i_n)^{-1}\right]^{-1} \quad \text{using (73)} \\ &\approx 1 + \sum_{n=1}^{N} s_n i_n + \left[\sum_{n=1}^{N} s_n i_n\right]^2 - \sum_{n=1}^{N} s_n [i_n]^2 \end{aligned} \tag{77}$$

where we have approximated the line above by a second order Taylor series approximation around

$$\begin{aligned} &i_n = 0 \text{ for } n = 1, \ldots, N \\ &= P_L + i^{*2} - \sum_{n=1}^{N} s_n [i_n]^2 \quad \text{using (75) and (76)} \\ &= P_L - \sum_{n=1}^{N} s_n [i_n - i^*]^2. \end{aligned}$$

Thus the Paasche price index P_P, under the simplifying assumptions (73), is approximately equal to the Laspeyres price index P_L, minus the variance of the sample distribution of the individual commodity inflation rates i_n.[107]

107. Let X be a discrete distribution that takes on the values i_n with probability s_n for $n = 1, \ldots, N$. Then $i^* \equiv \sum_{n=1}^{N} s_n i_n$ is the mean of this random variable and $\sum_{n=1}^{N} s_n [i_n - i^*]^2$ is its variance. The approximation result (77) is a generalization of a result in Diewert (1998; p. 56–57).

We can obtain a similar second order Taylor series approximation for the Fisher price index P_F:

$$\begin{aligned} P_F(i_1,\ldots,i_N) &\equiv \left[\sum_{n=1}^{N} s_n^0(1+i_n)\right]^{1/2}\left[\sum_{n=1}^{N} s_n^1(1+i_n)^{-1}\right]^{-1/2} \\ &\approx \left[\sum_{n=1}^{N} s_n(1+i_n)\right]^{1/2}\left[\sum_{n=1}^{N} s_n(1+i_n)^{-1}\right]^{-1/2} \quad \text{using (73)} \\ &\approx P_L - (1/2)\sum_{n=1}^{N} s_n[i_n - i^*]^2 \end{aligned} \tag{78}$$

where we have approximated the line above the last line by a second order Taylor series approximation around $i_n = 0$ for $n = 1,\ldots,N$.

Finally, we can obtain a similar second order Taylor series approximation for the Walsh price index P_W defined by (23) above:

$$\begin{aligned} P_W(i_1,\ldots,i_N) &\equiv \left[\sum_{n=1}^{N}(s_n^0 s_n^1)^{1/2}(1+i_n)^{1/2}\right]\Big/\left[\sum_{n=1}^{N}(s_n^0 s_n^1)^{1/2}(1+i_n)^{-1/2}\right] \\ &\approx \left[\sum_{n=1}^{N} s_n(1+i_n)^{1/2}\right]\Big/\left[\sum_{n=1}^{N} s_n(1+i_n)^{-1/2}\right] \quad \text{using (73)} \\ &\approx P_L - (1/2)\sum_{n=1}^{N} s_n[i_n - i^*]^2 \end{aligned} \tag{79}$$

where we have approximated the line above the last line by a second order Taylor series approximation around $i_n = 0$ for $n = 1,\ldots,N$.

Note that the Fisher and Walsh price indexes have the same second order Taylor series approximations.

Now we can subtract (78) or (79) from (75) and we obtain the following expression for the *approximate representativity bias for the Laspeyres formula*:

$$B_L(i_1,\ldots,i_N) \equiv (1/2)\sum_{n=1}^{N} s_n[i_n - i^*]^2. \tag{80}$$

Thus the approximate bias for the Laspeyres price index is equal to one half of the variance of the commodity specific inflation rates between the two periods under consideration. This approximate representativity bias is always nonnegative; i.e., the Laspeyres price index will generally give an answer that is too high compared to an index that uses more representative quantity weights.

In a similar fashion, we can subtract (78) or (79) from (77) and we obtain the following expression for the *approximate representativity bias for the Paasche formula*:

$$B_P(i_1, \ldots, i_N) \equiv -(1/2) \sum_{n=1}^{N} s_n [i_n - i^*]^2. \tag{80}$$

Thus the approximate bias for the Paasche price index is equal to minus one half of the variance of the commodity specific inflation rates between the two periods under consideration. This approximate representativity bias is always nonpositive; i.e., the Paasche price index will generally give an answer that is too low compared to an index that uses more representative quantity weights.

Our conclusion is that the HICP Laspeyres type index suffers from representativity bias and hence it will generally show higher rates of inflation than a pure price index that uses more representative quantity weights. Formula (80) above gives a useful approximation to this representativity bias.

7.3. Fixed Base versus Chain Indexes

As we saw in section 6 above, Eurostat gives member EU countries a considerable amount of leeway in deciding how often they should change their base year: member countries are allowed to keep their base year fixed for up to 7 years![108]

Unfortunately, this lack of harmonization on how often to rebase will lead to a lack of comparability between the member country HICP's under certain conditions. We will now proceed to give such a set of conditions.

Let us make assumption (73) again; i.e., that expenditure shares remain constant from period to period. We now define the *period t inflation rate for commodity n relative to the base period 0*, i_n^t, as follows:

$$1 + i_n^t \equiv p_n^t / p_n^0, \quad n = 1, \ldots, N; \quad t = 1, \ldots, T. \tag{81}$$

If we look at the Laspeyres formula going from period 0 to t, as in the previous section, we can derive the following expression for the *approximate representativity bias for the period t Laspeyres fixed base formula*:

108. However, Article 3 of Commission Regulation (EC) 2454/97, which deals with HICP weightings, requires Member States to use weightings which reflect consumers' expenditure patterns in a weighting reference period ending no more than seven years before the preceding December. In practice this means a base year revision once every five years since budget survey results usually become available with a considerable time lag. In addition, each year, Eurostat price statisticians have to critically examine the expenditure weightings and where reliable evidence shows that there have been important changes since the weighting reference period that would affect the change in the HICP by more than 0.1 percentage point, the weightings of the HICP must be adjusted appropriately.

$$B_L^{0,t}(i_1^t,\ldots,i_n^t) \equiv (1/2)\sum_{n=1}^{N} s_n[i_n^t - i^{t*}]^2 \tag{82}$$

where

$$i^{t*} \equiv \sum_{n=1}^{N} s_n i_n^t. \tag{83}$$

If the long term price relatives $p_n^t/p_n^0 \equiv 1 + i_n^t$ trend *linearly* with time t, then from (82), it can be seen that the approximate representativity bias for the period t Laspeyres fixed base formula will grow *quadratically* with time. Thus under the assumption of linear trends in prices over time, the fixed base Paasche and Laspeyres price indexes will diverge at a rate that is quadratic in time whereas under the same assumptions, the chained Paasche and Laspeyres price indexes will diverge at a rate that is only linear in time. Hence under these conditions, the fixed base Laspeyres price index will grow much more quickly than its chained counterpart.

However, if the long term price relatives p_n^t/p_n^0 do not grow linearly with time but simply fluctuate randomly around a constant, then the conclusion in the previous paragraph will not hold and both the fixed base and chained Laspeyres price indexes will exhibit much the same behavior.

The implications for Eurostat of all this seem fairly clear: each country should compute the variance of their aggregate long term price relatives and determine whether these variances are growing at faster than linear rates in recent years. If this is the case, then the frequency of rebasing will make a difference to the aggregate country index. Under these conditions, in order to make the country inflation rates comparable, Eurostat should eliminate any choice in the frequency of rebasing.[109]

7.4. The Choice of Formula at the Elementary Level

In section 5.1 above, we gave a brief overview of how the basic fixed base Laspeyres price index is constructed. In particular, we noted that in the base period, expenditure shares for 100 or so basic commodity classes are estimated and a sample of representative items is chosen for each of these basic commodity classes. These items are priced every month, long term price relatives are calculated and then averaged for each commodity and then these long term "average" price relatives are inserted into the Laspeyres formula, (11) above. The question that we want to address in this section is: exactly how should the sampled long term price relatives be averaged?

HICP regulations allow the use of two types of averaging: either the DUTOT (1738) formula can be used or the geometric mean of the sample of price relatives can be used;

109. Of course, annual rebasing will typically lead to the smallest representativity bias and so I would favor this alternative.

i.e., the formula of Jevons can be used (see formula (41) above). If there are K prices in the sample of prices for the commodity class, the *Dutot formula* P_D is defined as a ratio of average prices as follows:[110]

$$P_D(p^0, p^1) \equiv (1/K)\sum_{k=1}^{K} p_k^1/(1/K)\sum_{k=1}^{K} p_k^0 = \sum_{k=1}^{K} p_k^1/\sum_{k=1}^{K} p_k^0. \tag{84}$$

The Carli formula, (39) above, was explicitly banned as an aggregation formula at the first stage of aggregation due to its systematic failure of the time reversal test.[111]

However, given that the underlying index concept for the HICP is a fixed base Laspeyres price index, it is necessary to ask whether the use of the JEVONS or Dutot formulae at the first stage of aggregation is consistent with the overall Laspeyres index methodology? In the case of the Jevons formula, our tentative answer to this question is *no* as we shall explain below.

Before we can address the above question, it is necessary to discuss another problem. In section 2 above where we first introduced the value aggregates V^0 and V^1 and the price and quantity indexes, $P(p^0, p^1, q^0, q^1)$ and $Q(p^0, p^1, q^0, q^1)$ that decomposed the value ratio V^1/V^0 into the price change part $P(p^0, p^1, q^0, q^1)$ and the quantity change part $Q(p^0, p^1, q^0, q^1)$, we took it for granted that the period t price and quantity for commodity n, p_n^t and q_n^t respectively, were well defined. However, is this definition really straightforward? Again, the answer to this question is *no* since individual consumers may purchase the same item during period t *at different prices.* Similarly, if we look at the sales of a particular shop or outlet that sells to consumers, *the same item may sell at very different prices during the course of the period.* Hence before we can apply a traditional bilateral price index of the form $P(p^0, p^1, q^0, q^1)$ considered in previous sections of this paper, there is a non trivial *first stage aggregation problem* in order to obtain the basic prices p_n^t and q_n^t that are the components of the price vectors $p^0$and p^1 and the quantity vectors q^0 and q^1.

DIEWERT (1995a, p. 20–21), following WALSH[112] and DAVIES (1924, 1932), suggested

110. We have abused our notation in letting p^0 and p^1 now denote K dimensional vectors of sampled item prices in a particular expenditure category. Later in this section, we also let p^0 and p^1 have their original meaning as N dimensional vectors and finally, we also let p^0 and p^1 denote M dimensional vectors that represent the universe of item prices in a particular expenditure category. However, the meaning of p^0 and p^1 will be clear from the context.

111. Recall FISHER's (1922; p. 66) observation that $P_C(p^0, p^1)P_C(p^1, p^0) > 1$ unless p^1 is proportional to p^0.

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

that the appropriate quantity at this first stage of aggregation is the total quantity purchased of the narrowly defined item and the corresponding price is the value of purchases of this item divided by the total amount purchased (which is a unit value). We will adopt this suggestion as our concept for the price and quantity at the first stage of aggregation.

Having decided on an appropriate definition of price and quantity for an item at the very lowest level of aggregation, we now consider how to aggregate these elementary prices and quantities into the 100 or so higher level aggregates. Let us choose one of these 100 categories and suppose that there are M lowest level items or specific commodities in this category. If we take the Laspeyres perspective to index number theory, then we can use the Laspeyres formula at this elementary level of aggregation:[113]

$$P_L(p^0, p^1, q^0) \equiv \sum_{m=1}^{M} p_m^1 q_m^0 \Big/ \sum_{m=1}^{M} p_m^0 q_m^0 = \sum_{m=1}^{M} \rho_m^0 p_m^1 \Big/ \sum_{m=1}^{M} \rho_m^0 p_m^0 \tag{85}$$

where the *base period item probabilities* ρ_m^0 are defined as follows:

$$\rho_m^0 \equiv q_m^0 \Big/ \sum_{i=1}^{m} q_i^0; \quad m = 1, \ldots, M. \tag{86}$$

Thus the base period probability for item m, ρ_m^0, is equal to the purchases of item m in the base period relative to total purchases of all items in the commodity class in the base period. We note that these definitions require that all items in the commodity class have the same units (or can be quality adjusted into "standard" units).

Now it is easy to see how formula (85) could be turned into a rigorous sampling framework for sampling prices in the particular commodity class under consideration. If item prices in the commodity class were sampled proportionally to their base period probabilities ρ_m^0, then the Laspeyres index (85) could be estimated by the Dutot index defined by (84). In general, with an appropriate sampling scheme, the use of the Dutot formula at the elementary level of aggregation can be perfectly consistent with a Laspeyres index concept.

The Dutot formula can also be consistent with a Paasche index concept. If we use the Paasche formula at the elementary level of aggregation, we obtain the following formula:

$$P_P(p^0, p^1, q^1) \equiv \sum_{m=1}^{M} p_m^1 q_m^1 \Big/ \sum_{m=1}^{M} p_m^0 q_m^1 = \sum_{m=1}^{M} \rho_m^1 p_m^1 \Big/ \sum_{m=1}^{M} \rho_m^1 p_m^0 \tag{87}$$

where the *period one item probabilities* ρ_m^1 are defined as follows:

113. Recall that the Laspeyres formula is consistent in aggregation so that first constructing Laspeyres indexes for each of the 100 commodity classes and then doing a second stage Laspeyres index will be equivalent to doing a single stage Laspeyres index. BALK (1994) considers in some detail the problems involved in setting up a sampling framework for the Laspeyres index.

$$\rho_m^1 \equiv q_m^1 / \sum_{i=1}^{M} q_i^1; \quad m = 1, \ldots, M. \tag{88}$$

Thus the period one probability for item m, ρ_m^1, is equal to the quantity purchased of item m in period one relative to total purchases of all items in the commodity class in that period.

Again, it is easy to see how formula (87) could be turned into a rigorous sampling framework for sampling prices in the particular commodity class under consideration. If item prices in the commodity class were sampled proportionally to their period one probabilities ρ_m^1, then the Paasche index (87) could be estimated by the Dutot index defined by (84). In general, with an appropriate sampling scheme, the use of the Dutot formula at the elementary level of aggregation can be perfectly consistent with a Paasche index concept.[114]

Rather than use the fixed basket representations for the Laspeyres and Paasche indexes, formulae (8) and (9) above, and use the quantity shares ρ_m^0 or ρ_m^1 as probability weights for prices, we could use the expenditure share representations for the Laspeyres and Paasche indexes, formulae (11) and (12) above, and use the expenditure shares s_m^0 or s_m^1 as probability weights for price relatives. Thus if the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their base period expenditure shares in the commodity class, then the following *Carli index*

$$P_C(p^0, p^1) \equiv \sum_{k=1}^{K} (1/K) p_k^1 / p_k^0 \tag{89}$$

can be consistent with the estimation of a Laspeyres price index for that commodity class.[115] On the other hand, if the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their period one expenditure shares in the commodity class, then the following *harmonic index*

$$P_H(p^0, p^1) \equiv \left\{ \sum_{k=1}^{K} (1/K) [P_k^1 / p_k^0]^{-1} \right\}^{-1} \tag{90}$$

can be consistent with the estimation of a Paasche price index for that commodity class.[116]

114. Of course, the Dutot index based on the Paasche sampling framework will usually be greater than the Dutot index based on the Laspeyres sampling framework due to representativity or substitution bias.

115. We have abused our notation a bit here in not inventing a new notation for the sampled prices. We have also let the price vectors p^0 and p^1 denote both vectors of sampled prices as well as the complete list of item prices in the commodity class for periods 0 and 1.

116. If the same items are chosen in (89) and (90), then since a harmonic mean is equal or less than the corresponding arithmetic mean, we have $P_H(p^0, p^1) \leq P_C(p^0, p^1)$; see HARDY, LITTLEWOOD and PÓLYA (1934, p. 26). Equation (77) can be used to estimate the difference between P_H and P_C if we let $N = K$ and let the shares $s_n = 1/K$; i.e., (77) implies $P_H \approx P_C - \sum_{k=1}^{K} (1/K)[i_k - i^*]^2$, where $i_k \equiv p_k^1 / p_k^0$.

The above results show that the Dutot and Carli elementary indexes, P_D and P_C defined by (84) and (89), can be justified as approximations to an underlying Laspeyres price index for the commodity class under appropriate price sampling schemes. They also show that the Dutot and harmonic elementary indexes, P_D and P_H defined by (84) and (90), can be justified as approximations to an underlying Paasche price index for the commodity class under appropriate price sampling schemes. However, we have *not* been able to justify the use of the Jevons elementary index P_J defined by (41) as an approximation to an underlying Laspeyres index. Hence, there appears to be an inconsistency in the HICP, which adopts a fixed base Laspeyres methodology but yet allows the use of the geometric mean of a sample of price relatives as an admissible elementary index.

The above inconsistency could be avoided if the HICP were to adopt a more symmetric approach to index number theory as we shall now explain.

Recall THEIL's (1968) stochastic approach to index number theory explained in section 5.3 above. Let us apply this framework to a particular commodity classification. The logarithm of our index number target is:

$$\ln P_T(p^0, p^1, q^0, q^1) = \sum_{m=1}^{M}(1/2)[s_m^0 + s_m^1] \ln p_m^1/p_m^0 = \sum_{m=1}^{M} \rho_m \ln p_m^1/p_m^0 \tag{91}$$

where the probability of selecting item m in this category is ρ_m defined as the average of the category expenditure shares, s_m^0 and s_m^1, pertaining to the two periods under consideration:

$$\rho_m \equiv (1/2)[s_m^0 + s_m^1]; \quad m = 1, \ldots, M. \tag{92}$$

If items are selected proportionally to the above probabilities, then $\sum_{k=1}^{K}(1/K)\ln p_k^1/p_k^0$ could be an estimator of the category logarithmic price change and hence the Jevons category price index P_J defined as

$$\ln P_J(p^0, p^1) \equiv \sum_{k=1}^{K}(1/K) \ln p_k^1/p_k^0 \tag{93}$$

can be justified as a symmetric measure of price change for the category under consideration. Thus under an item sampling scheme where the probability of selection is proportional to the probabilities defined by (92)[117], we can justify the Jevons elementary index as an approximation to the Törnqvist Theil price index for the category, P_T.

We now consider another sampling framework. Suppose that item expenditure shares during the two periods under consideration are equal so that

117. This item sampling framework is not particularly "practical" but our goal here is to see under what conditions the Jevons elementary index emerges as an appropriate one for the underlying index number concept.

$$s_m^0 = s_m^1; \quad m = 1, \ldots, M. \tag{94}$$

Now if the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their base period expenditure shares in the commodity class, then the Carli index P_C defined by (89) can be justified as an approximation to the Laspeyres price index. However, under assumption (94), the same sampling frame can be used in order to justify the harmonic index P_H defined by (90) as an approximation to the Paasche price index. Hence, taking the geometric mean of P_C and P_H gives us an elementary index that will be a good approximation to the Fisher ideal price index P_F for the category under consideration. Thus define the elementary index P_{CSWD} advocated by CARRUTHERS, SELLWOOD and WARD (1980) and DALÉN (1992) as follows:

$$P_{CSWD}(p^0, p^1) \equiv \left[P_C(p^0, p^1) P_H(p^0, p^1)\right]^{1/2} \tag{95}$$

where P_C and P_H are defined by (89) and (90) respectively.[118]

A major advantage of P_{CSWD} over the Carli and Harmonic elementary indexes is that P_{CSWD} satisfies the time reversal test whereas the other two indexes do not.[119] A second major advantage of P_{CSWD} is that this index under some conditions is consistent[120] with the use of a Fisher index; i.e., *it is consistent with taking a symmetric average of the Paasche and Laspeyres indexes.*

A natural question to ask at this stage is: how different will the Carruthers, Sellwood, Ward and Dalén index P_{CSWD} be from the Jevons elementary index P_J? As in section 7.2 above, define the individual item inflation rates i_k for the prices in the sample as follows:

$$1 + i_k \equiv p_k^1 / p_k^0, \quad k = 1, \ldots, K. \tag{96}$$

It is easy to show that both $P_J(p^0, p^1)$ and $P_{CSWD}(p^0, p^1)$ can be written as functions of the K individual item inflation rates, $i_1, \ldots, i_K$, so that we can write P_J and P_{CSWD} as P_J

118. P_{CSWD} was first suggested by FISHER (1922; p. 472) as his formula number 101. FISHER (1922; p. 211) observed that PCSWD was numerically very close to the unweighted geometric mean index P_J defined by (93) for his data set. FISHER (1922; p. 245) regarded P_J and P_{CSWD} as being the best unweighted index number formula although he regarded both formulae as being "poor". FISHER (1922; p. 244–245) also classified P_H as the worst "poor" and P_C as the second best "worthless" index number formula. In more recent times, CARRUTHERS, SELLWOOD and WARD (1980; p. 25) and DALÉN (1992; p. 140) have advocated the use of P_{CSWD} as an elementary index.

119. Recall that FISHER (1922; p. 66) showed that $P_C(p^0, p^1) P_C(p^1, p^0) > 1$ unless p^1 is proportional to p^0. In a similar fashion, it can be shown that $P_H(p^0, p^1) P_H(p^1, p^0) < 1$ unless p_1 is proportional to p^0. Thus P_C will generally have an upward bias while P_H will generally have a downward bias. DALÉN (1994; p. 150–151) gives some nice explanations for the upward bias of the Carli index.

120. We require assumptions (94) plus an appropriate sampling framework. DALÉN (1994; p. 151) regarded the CARRUTHERS, SELLWOOD, WARD and DALÉN formula as an approximation to the Fisher ideal price index.

$(i_1, \ldots, i_K)$ and $P_{CSWD}(i_1, \ldots, i_K)$. DALÉN (1992, p. 143) showed that the second order Taylor series approximations to $P_J(i_1, \ldots, i_K)$ and $P_{CSWD}(i_1, \ldots, i_K)$ around the point $(i_1, \ldots, i_K) = (0, \ldots, 0)$ are:

$$P_J(i_1, \ldots, i_K) \approx 1 + i^* - (1/2) \sum_{k=1}^{K} (1/K)[i_k - i^*]^2; \tag{97}$$

$$P_{CSWD}(i_1, \ldots, i_K) \approx 1 + i^* - (1/2) \sum_{k=1}^{K} (1/K)[i_k - i^*]^2 \tag{98}$$

where $i^* \equiv \sum_{k=1}^{K}(1/K)i_k$ is the sample mean of the individual item inflation rates. The approximations on the right hand side of (97) and (98) are identical and hence show that *P_J and P_{CSWD} approximate each other to the second order* around the point $(i_1, \ldots, i_K) = (0, \ldots, 0)$.[121] Thus for normal time series data, these two elementary indexes will usually be very close.[122]

The approximation results in the above paragraph indicate that the use of the Jevons elementary index is *not consistent* with the fixed base Laspeyres methodology that is at the heart of the HICP. However, the Jevons index *is consistent* with the stochastic approach of THEIL and is *approximately consistent* with a *symmetric* fixed basket approach to index number theory. Our discussion above indicates that a great deal of care needs to be taken in order to work out a sampling framework that is consistent with the overall approach to index number theory that is chosen by the statistical agency.

We conclude this section with some quotations which summarize the results of recent index number studies that make use of *scanner data*; i.e., of detailed data on the prices and quantities of individual items that are sold in retail outlets.

"A second major recent development is the willingness of statistical agencies to experiment with scanner data, which are the electronic data generated at the point of sale by the retail outlet and generally include transactions prices, quantities, location, date and time of purchase and the product described by brand, make or model. Such detailed data may prove especially useful for constructing better indexes at the elementary level. Recent studies that use scanner data in this way include SILVER (1995), REINSDORF (1996), BRADLEY, COOK, LEAVER and MOULTON (1997), DALÉN (1997), DE HAAN and OPPERDOES (1997) and HAWKES (1997). Some estimates of elementary index bias (on an annual basis) that emerged from these studies were: 1.1 percentage points for television sets in the United Kingdom; 4.5 percentage points for coffee in the United States; 1.5 percentage points for ketchup, toilet tissue, milk and tuna in the United States; 1 percentage point for fats, detergents, breakfast cereals and frozen fish in Sweden; 1 percen-

121. DIEWERT (1995a; p. 29) generalized Dalén's result, allowing the point of approximation to be an arbitrary vector of constants instead of the vector of zeros.

122. The corresponding second order approximation to P_C is $1 + i^*$ and to P_H is $1 + i^* - \sum_{k=1}^{K} (1/K) [i_k - i^*]^2$, which is 1 plus the sample mean minus the sample variance of the item inflation rates.

tage point for coffee in the Netherlands and 3 percentage points for coffee in the United States respectively. These bias estimates incorporate both elementary and outlet substitution biases and are significantly higher than our earlier Ballpark estimates of.255 and.41 percentage points. On the other hand, it is unclear to what extent these large bias estimates can be generalized to other commodities." W. ERWIN DIEWERT (1998, p. 54–55).

"Before considering the results it is worth commenting on some general findings from scanner data. It is stressed that the results here are for an experiment in which the same data were used to compare different methods. The results for the U.K. Retail Prices Index can not be fairly compared since they are based on quite different practices and data, their data being collected by price collectors and having strengths as well as weaknesses (FENWICK, BALL, SILVER and MORGAN, 2002). Yet it is worth following up on DIEWERT's (2001c) comment on the U.K. Retail Prices Index electrical appliances section, which includes a wide variety of appliances, such as irons, toasters, refrigerators, etc. which went from 98.6 to 98.0, a drop of 0.6 percentage points from January 1998 to December 1998. He compares these results with those for washing machines and notes that '..it may be that the non washing machine components of the electrical appliances index increased in price enough over this period to cancel out the large apparent drop in the price of washing machines but I think that this is somewhat unlikely.' A number of studies on similar such products have been conducted using scanner data for this period. Chained Fishers indices have been calculated from the scanner data, (the RPI (within year) indices are fixed base Laspeyres ones), and have been found to fall by about 12 % for televisions (SILVER and HERAVI, 2001b), 10 % for washing machines (Table 7 below), 7.5 % for dishwashers, 15 % for cameras and 5 % for vacuum cleaners (SILVER and HERAVI, 2001c). These results are quite different from those for the RPI section and suggest that the washing machine disparity, as DIEWERT notes, may not be an anomaly. Traditional methods and data sources seem to be giving much higher rates for the CPI than those from scanner data, though the reasons for these discrepancies were not the subject of this study." MICK SILVER and SAEED HERAVI (2001d, p. 25).

The above quotations summarize the results of many category index number studies that are based on the use of scanner data. These studies indicate that when detailed price and quantity data are used in order to compute superlative indexes or hedonic indexes for an expenditure category, the resulting measures of price change are generally below the corresponding official statistical agency estimates of price change for that category. Often the measures of price change based on the use of scanner data are *considerably below* the corresponding official measures.[123] These results are very troubling. They

123. However, scanner data studies do not always show large potential biases in official CPIs. Masato Okamoto has informed me that a large scale comparative study in Japan is underway. Using scanner data for about 250 categories of processed food and daily necessities collected over the period 1997 to 2000, it was found that the indexes based on scanner data averaged only about 0.2 percentage points below the corresponding official indexes per year. Japan uses the Dutot formula at the elementary level in its official CPI.

seem to indicate that the sampling procedures and index number formulae used by statistical agencies to calculate measures of price change at the lowest levels of aggregation are leading to estimates of price change that are considerably higher (in many cases) than corresponding estimates of price change that are based on the use of superlative indexes or hedonic regression methods.

We turn now to one of the most difficult problems associated with the HICP and that is the treatment of housing.

7.5. *The Treatment of Housing*

"We have noticed also that though the benefits which a man derives from living in his own house are commonly reckoned as part of his real income, and estimated at the net rental value of his house; the same plan is not followed with regard to the benefits which he derives from the use of his furniture and clothes. It is best here to follow the common practice, and not count as part of the national income or dividend anything that is not commonly counted as part of the income of the individual." ALFRED MARSHALL (1898, p. 594–595).

When a durable good (other than housing) is purchased by a consumer, national Consumer Price Indexes (and the HICP) attribute all of that expenditure to the period of purchase even though the use of the good extends beyond the period of purchase. However, the treatment of owner occupied housing in national CPI's is more diverse. The Bureau of Labor Statistics in the U.S. follows the treatment suggested by MARSHALL above and estimates a price for the use of an owner occupied dwelling that is equal to the rental of an equivalent dwelling. This is the *rental equivalence approach* to the treatment of owner occupied housing. Statistics Iceland estimates a user cost for the dwelling; we will discuss this *user cost approach* in more detail below.

There are two additional approaches to the treatment of owner occupied housing in a CPI. The first of these two approaches is the *net acquisitions approach*, which is nicely described by GOODHART as follows:

> "The first is the net acquisition approach, which is the change in the price of newly purchased owner occupied dwellings, weighted by the net purchases of the reference population. This is an asset based measure, and therefore comes close to my preferred measure of inflation as a change in the value of money, though the change in the price of the stock of existing houses rather than just of net purchases would in some respects be even better. It is, moreover, consistent with the treatment of other durables. A few countries, e.g., Australia and New Zealand, have used it, and it is, I understand, the main contender for use in the Euro-area Harmonized Index of Consumer Prices (HICP), which currently excludes any measure of the purchase price of (new) housing, though it does include minor repairs and maintenance by home owners, as well as all expenditures by tenants." CHARLES GOODHART (2001, F350).

Thus the weights for the net acquisitions approach are the net purchases of the household sector of new houses in the base period and the long term price relative for this category is the price of new houses (quality adjusted) in the current period relative to the price of new houses in the base period. Note that this price does not include the land that the dwelling sits on. Finally, note that this treatment of housing is identical to the treatment of purchases of other consumer durables.

Our fourth approach to the treatment of owner occupied housing, the *payments approach*, is described by GOODHART as follows:

> "The second main approach is the payments approach, measuring actual cash outflows, on down payments, mortgage repayments and mortgage interest, or some subset of the above. This approach always, however, includes mortgage interest payments. This, though common, is analytically unsound. First, the procedure is not carried out consistently across purchases. Other goods bought on the basis of credit, e.g., credit card credit, are usually not treated as more expensive on that account (though they have been in New Zealand). Second, the treatment of interest flows is not consistent across persons. If a borrower is worse off in some sense when interest rates rise, then equivalently a lender owning an interest bearing asset is better off; why measure one and not the other? If I sell an interest earning asset, say a money market mutual fund holding, to buy a house, why am I treated differently to someone who borrows on a (variable rate) mortgage? Third, should not the question of the price of any purchase be assessed separately from the issue of how that might be financed? Imports, inventories and all business purchases tend to be purchased in part on credit. Should we regard imports as more expensive, when the cost of trade credit rises? Money, moreover, is fungible. As we know from calculations of mortgage equity withdrawal, the loan may be secured on the house but used to pay for furniture. When interest rates rise, is the furniture thereby more expensive? Moreover, the actual cash out-payments totally ignore changes in the on going value of the house whether by depreciation, or capital loss/gain, which will often dwarf the cash flow. Despite its problems, such a cash payment approach was used in the United Kingdom until 1994 and still is in Ireland." CHARLES GOODHART (2001, F350-F351).

Thus the payments approach to owner occupied housing is a kind of a cash flow approach to the costs of operating an owner occupied dwelling. I agree with GOODHART in being critical of this approach.[124] My main objection to the approach is that it ignores the opportunity costs of holding the equity in the owner occupied dwelling and it ignores depreciation. However, once adjustments are made for these imputed costs, we have drifted into a rather complicated user cost approach to the treatment of housing. In general, this approach will tend to lead to much smaller monthly expenditures on owner occupied housing than the other 3 main approaches.

With the above four approaches to the treatment of owner occupied housing in mind, we turn now to its treatment in the HICP:

> "A special coverage problem concerns owner-occupied housing. This has always been one of the most difficult sectors to deal with in CPIs.
>
> Strictly, the price of housing should not be included in a CPI because it is classified as capital. On the other hand, the national accounts classifies imputed rents of owner-occupiers as part of consumers' expenditure. This is a reasonable thing to do if the aim is to measure the volume of consumption of the capital resource of housing. But that is not what a CPI is measuring.
>
> Some countries, following the compensation index concept, would prefer to have mortgage interest included in the HICP. This approach could indeed be defended for a compensation index,

124. I agree with most of GOODHART's criticisms of the payments approach except that when (real) interest rates rise on a sustained basis, I would argue that furniture is thereby made more expensive from at least two perspectives. To cover these increased real interest rate costs, rental prices of furniture should indeed rise (the rental equivalence perspective) and the furniture owner's opportunity cost of using the furniture should also rise (the user cost perspective).

because there is not doubt that the monthly mortgage payment is an important element in the budget of many households: a rise in the interest rate acts in exactly the same way as a price increase from the point of view of the individual household. But this is not acceptable for a wider inflation index.

So, after many hours of debate, the Working Party came to the conclusion that there were just two options. The first was to simply exclude owner-occupied housing from the HICP. One could at least argue that this was a form of harmonization, although it is worrying that there are such large differences between Member States in the percentages of the population which own or rent their dwellings. Exclusion also falls in line with the international guideline issued 10 years ago by the ILO. Furthermore, it would be possible to supplement the HICP with a separate house price index, which could be used by analysts as part of a battery of inflation indicators.

The second option was to include owner-occupied housing on the basis of acquisition costs, essentially treating them like any other durable. Most secondhand housing would be excluded: in practice the index would include new houses plus a small volume of housing new to the household sector (sales from the company or government sectors to the household sector).

The main problem here is practical: several countries do not have new house price indices and their construction could be difficult and costly. A Task Force is at present examining these matters. Final recommendations are due at the end of 1999." JOHN ASTIN (1999, p. 5).

Thus the HICP seems to be leaning towards a fifth approach to the treatment of owner occupied housing; i.e., to just omit it entirely from the index (which is the current treatment)! The problem with this solution is the fact that the proportion of owner occupied dwellings differs dramatically across EU countries: for example, as noted earlier, only 40 % of Germans live in owner occupied dwellings while about 85 % of Spaniards live in owner occupied dwellings. Thus omitting owner occupied housing from the HICP will tend to make the indexes *incomparable* across EU countries.

The next most preferred approach to the treatment of owner occupied housing mentioned by ASTIN is the acquisitions approach. The problem with this approach is that *purchases* of new houses simply do not *reflect the actual consumption of housing services* for the population of owner-occupiers. Thus if our purpose is to measure the real consumption of the population during a period and a price index is required to deflate nominal consumption expenditures into real consumption, then the acquisitions approach to the treatment of owner occupied housing will not be satisfactory.

We now consider the remaining three approaches to the treatment of owner occupied housing. We agree with GOODHART and ASTIN that the payments approach is not very suitable as a measure of general housing inflation. That leaves the rental equivalence and user cost approaches. From the viewpoint of HICP methodology which tries to avoid imputations, both of these approaches are not suitable, since they involve imputations. However, we have already noted that the HICP endorses quality change adjustments and of course, these are imputations. Moreover, the HICP endorses sampling of prices at the elementary level of aggregation and the resulting sample average measures of price change are also imputations.[125] Thus I do not think that the "no imputations" rule in the HICP should be taken too seriously.

125. Thus the rental equivalence approach to owner occupied housing essentially collects a sample of rents for various dwellings and then uses this sample of rents to impute rents to owner occupied dwellings with similar characteristics.

We shall conclude this section by trying to make a case for the use of either the rental equivalence or user cost approaches for the treatment of owner occupied housing.

The *rental equivalence approach* simply values the services yielded by the use of a consumer durable good for a period by the corresponding market rental value for the same durable for the same period of time (if such a rental value exists). This is the approach taken by the Bureau of Labor Statistics in the U.S. and in the *System of National Accounts: 1993* for owner occupied housing:

> "As well-organized markets for rented housing exist in most countries, the output of own-account housing services can be valued using the prices of the same kinds of services sold on the market with the general valuation rules adopted for goods and services produced on own account. In other words, the output of housing services produced by owner-occupiers is valued at the estimated rental that a tenant would pay for the same accommodation, taking into account factors such as location, neighbourhood amenities, etc. as well as the size and quality of the dwelling itself." Eurostat and others (1993, p. 134).

However, the *System of National Accounts: 1993* follows MARSHALL (1898, p. 595) and does *not* extend the rental equivalence approach to consumer durables other than housing. This seemingly inconsistent treatment of durables is explained in the *SNA 1993* as follows:

> "The production of housing services for their own final consumption by owner-occupiers has always been included within the production boundary in national accounts, although it constitutes an exception to the general exclusion of own-account service production. The ratio of owner-occupied to rented dwellings can vary significantly between countries and even over short periods of time within a single country, so that both international and intertemporal comparisons of the production and consumption of housing services could be distorted if no imputation were made for the value of own-account services." EUROSTAT and others (1993, p. 126).

As mentioned earlier, the BLS uses the rental equivalence approach to price the use of owner occupied housing. This is an opportunity cost approach: the owner values the services yielded by his or her dwelling by the amount of rental income it could generate during each period. This seems to me to be a very reasonable approach but it could fail under two conditions:

(i) Rental markets for some classes of owner occupied housing could be nonexistent or very thin or
(ii) Rental markets for some classes of owner occupied housing could be unrepresentative of arms length transactions; e.g., expensive houses could be rented to "friends" at reduced rates in exchange for house sitting services.

If either of the two conditions listed above are relevant to the country's housing markets, then the rental equivalence approach to the treatment of owner occupied housing will fail and in order to price the services yielded by owner occupied housing, it will be necessary to use an alternative opportunity cost approach: the user cost approach. We will now consider the user cost approach in more detail and contrast it to the acquisitions approach.

The *acquisitions approach* to the treatment of a consumer durable like housing is very simple: if one unit of the good costs P^0 dollars and the reference group of households purchases q^0 units of it in period 0, then the observed total purchase cost P^0q^0 is attributed to period 0.

The problem with this approach is that the *services* of the purchased goods are not confined to period 0. By the definition of a durable good (it lasts longer than one period), the purchase will yield a flow of services to the consumer for periods that follow period 0. Thus it does not seem appropriate to charge the *entire* purchase price P^0 to the initial period of purchase. But how should the purchase price be distributed or allocated across periods? This is a *fundamental problem of accounting*, where a similar cost allocation problem occurs when a firm purchases a durable input.

One solution to this cost allocation problem is the *historical cost accounting* solution, which works as follows. If the durable good lasts $T+1$ periods, then the cost accountant somehow obtains a set of $T+1$ depreciation rates, $d_0, d_1, \ldots, d_T$, such that $d_0 + d_1 + \ldots + d_T = 1$. Then d_tP^0 is allocated to period t for $t = 0, 1, 2, \ldots, T$.

Economists have tended to take a different approach to the cost allocation problem – an approach based on opportunity costs. Thus to determine the net cost of using the durable good during period 0, we assume that one unit of the durable good is purchased at the beginning of period 0 at the price P^0. The "used" or "second-hand" durable good can be sold at the end of period 0 at the price P_s^1. It might seem that a reasonable net cost for the use of one unit of the consumer durable during period 0 is its initial purchase price P^0 less its end of period 0 "scrap value" P_s^1. However, *money received at the end of the period is not as valuable as money that is received at the beginning of the period*. Thus in order to convert the end of period value into its beginning of the period equivalent value, it is necessary to *discount* the term P_s^1 by the term $1 + r^0$ where r^0 is the beginning of period 0 nominal interest rate that the consumer faces. Hence we define the *period 0 user cost* u^0 for the consumer durable[126] as

$$u^0 \equiv P^0 - P_s^1/(1 + r^0). \tag{99}$$

There is another way to view the user cost formula (99): the consumer purchases the durable at the beginning of period 0 at the price P^0 and charges himself or herself the rental price u^0. The remainder of the purchase price, I^0, defined as

$$I^0 \equiv P^0 - u^0 \tag{100}$$

is regarded as an *investment*, which is to yield the appropriate opportunity cost of capital r^0 that the consumer faces. At the end of period 0, this rate of return could be realized provided that I^0, r^0 and the selling price of the durable at the end of the period P_s^1 satisfy the following equation:

126. This approach to the derivation of a user cost formula was used by DIEWERT (1974b) who in turn based it on an approach due to HICKS (1946; p. 326).

$$I^0(1 + r^0) = P_s^1. \tag{101}$$

Given P_s^1 and r^0, (101) determines I^0, which in turn, given P^0, determines the user cost u^0 via (100).[127]

The user cost formula (99) can be put into more familiar form if we first define the period 0 *economic depreciation rate* δ and the period 0 *ex post asset inflation rate* i^0. Define δ by:

$$(1 - \delta) \equiv P_s^1/P^1 \tag{102}$$

where P_s^1 is the price of a used asset at the end of period 0 and P^1 is the price of a new asset at the end of period 0. The period 0 inflation rate for the new asset i^0 is defined by:

$$1 + i^0 \equiv P^1/P^0. \tag{103}$$

Substituting (103) into (102) gives us the following formula for the end of period 0 used asset price:

$$P_s^1 = (1 - \delta)(1 + i^0)P^0. \tag{104}$$

Substitution of (104) into (99) yields the following expression for the *period 0 user cost* u^0:

$$u^0 = \left[(1 + r^0) - (1 - \delta)(1 + i^0)\right]P^0/(1 + r^0) = \left[r^0 - i^0 + \delta(1 + i^0)\right]P^0/(1 + r^0). \tag{105}$$

Note that $r^0 - i^0$ can be interpreted as a period 0 *real interest rate* and $\delta(1 + i^0)$ can be interpreted as an *inflation adjusted depreciation rate.*

The user cost u^0 is expressed in terms of prices that are discounted to the *beginning* of period 0. However, it is also possible to express the user cost in terms of prices that are "discounted" to the *end* of period 0. Thus define the *end of period 0 user cost* p^0 as:[128]

127. This derivation for the user cost of a consumer durable was also made by DIEWERT (1974b; p. 504).

128. CHRISTENSEN and JORGENSON (1969) derived a user cost formula similar to (106) in a different way. If the inflation rate i equals 0, then the user cost formula (106) reduces to that derived by WALRAS (1954; p. 269) (first edition 1874). This zero inflation rate user cost formula was also derived by the industrial engineer A. HAMILTON CHURCH (1901; p. 907–908), who perhaps drew on the work of Matheson: "In the case of a factory where the occupancy is assured for a term of years, and the rent is a first charge on profits, the rate of interest, to be an appropriate rate, should, so far as it applies to the buildings, be equal (including the depreciation rate) to the rental which a landlord who owned but did not occupy a factory would let it for." EWING MATHESON (1910; p. 169), first published in 1884.

$$p^0 \equiv (1+r^0)u^0 = \left[r^0 - i^0 + \delta(1+i^0)\right]P^0 \tag{106}$$

where the last equation follows using (105). In the case where the asset inflation rate i^0 is zero, the end of the period user cost defined by (106) reduces to:

$$p^0 = (r^0 + \delta)P^0. \tag{107}$$

If the historical cost depreciation rate d_0 is equal to the economic depreciation rate δ, it can be seen that the no inflation user cost p^0 is *greater* than the corresponding historical cost period 0 cost allocation, δP^0, by the amount of the interest rate term, $r^0 P^0$. It is this difference that explains why the user cost (or rental equivalence) approach to the consumption of consumer durables will tend to give a *larger* value for consumption than the acquisitions approach, as we shall see shortly.

Abstracting from transactions costs and inflation, it can be seen that the end of the period user cost defined by (107) is an *approximate rental cost*; i.e., the rental cost for the use of a consumer (or producer) durable good should equal the opportunity cost of the capital tied up, $r^0 P^0$, plus the decline in value of the asset over the period, δP^0. When asset inflation is brought into the picture, the situation is more complicated. As it stands, the end of the period user cost formula (106) is an *ex post* (or after the fact) *user cost*: we cannot calculate the asset inflation rate i^0 until we have reached the end of period 0. Formula (106) can be converted into an *ex ante* (or before the fact) *user cost* formula if we interpret i^0 as an *anticipated asset inflation rate*. The resulting formula should approximate a market rental rate for the asset under inflationary conditions.

Note that in the user cost approach to the treatment of consumer durables, the *entire* user cost formula (106) is the period 0 price. Thus in the time series context, it is *not* necessary to deflate each component of the formula *separately*; the period 0 price $p^0 \equiv [r^0 - i^0 + \delta(1+i^0)]P^0$ is compared to the corresponding period 1 price, $p^1 \equiv [r^1 - i^1 + \delta(1+i^1)]P^1$ and so on.

We now want to compare the user cost approach to the treatment of consumer durables to the acquisitions approach. Obviously, in the short run the value flows associated with each approach could be very different. For example, if real interest rates, $r^0 - i^0$, are very high and the economy is in a severe recession or depression, then purchases of new consumer durables, Q^0 say, could be very low and even approach 0 for very long lived assets, like houses. On the other hand, using the user cost approach, existing stocks of consumer durables would be carried over from previous periods and priced out at the appropriate user costs and the resulting consumption value flow could be quite large. Thus in the short run, the monetary values of consumption under the two approaches could be vastly different. Hence, we will restrict ourselves in what follows to a (hypothetical) longer run comparison.[129]

129. The following material is taken from DIEWERT (2001a).

Suppose that in period 0, the reference population of households purchased q^0 units of a consumer durable at the purchase price P^0. Then *the period 0 value of consumption from the viewpoint of the acquisitions approach* is:

$$V_A^0 \equiv P^0 q^0. \tag{108}$$

Recall that the end of period user cost for one new unit of the asset purchased at the beginning of period 0 was p^0 defined by (106) above. In order to simplify our analysis, we assume *declining balance depreciation*; i.e., at the beginning of period 0, a one period old asset is worth $(1-\delta)P^0$; a two period old asset is worth $(1-\delta)^2 P^0$; . . .; a t period old asset is worth $(1-\delta)^t P^0$; etc. Under these hypotheses, the corresponding end of period 0 user cost for a new asset purchased at the beginning of period 0 is p^0; the end of period 0 user cost for a one period old asset at the beginning of period 0 is $(1-\delta)p^0$; the corresponding user cost for a two period old asset at the beginning of period 0 is $(1-\delta)^2 p^0$; . . .; the corresponding user cost for a t period old asset at the beginning of period 0 is $(1-\delta)^t p^0$; etc.[130] Our final simplifying assumption is that household purchases of the consumer durable have been growing at the geometric rate g into the indefinite past. This means that if household purchases of the durable were q^0 in period 0, then in the previous period they purchased $q^0/(1+g)$ new units; two periods ago, they purchased $q^0/(1+g)^2$ new units; . . .; t periods ago, they purchased $q^0/(1+g)^t$ new units; etc. Putting all of these assumptions together, it can be seen that *the period 0 value of consumption from the viewpoint of the user cost approach* is:

$$\begin{aligned} V_U^0 &\equiv p^0 q^0 + \left[(1-\delta)p^0 q^0/(1+g)\right] + \left[(1-\delta)^2 p^0 q^0/(1+g)^2\right] + \ldots \\ &= (1+g)(g+\delta)^{-1} p^0 q^0 \end{aligned} \tag{109}$$

summing the infinite series

$$= (1+g)(g+\delta)^{-1}\left[r^0 - i^0 + \delta(1+i^0)\right] P^0 q^0 \tag{110}$$

using (10).

We simplify (110) by letting the asset inflation rate i^0 be 0 (so that r^0 can be interpreted as a real interest rate) and we take the ratio of the user cost flow of consumption (110) to the acquisitions measure of consumption in period 0, (108):

$$V_U^0/V_A^0 = (1+g)(r^0+\delta)/(g+\delta). \tag{111}$$

130. For most consumer durables, the one hoss shay assumption for depreciation is more realistic than the declining balance model. To see the sequence of one hoss shay user costs, see HULTEN (1990) and DIEWERT and LAWRENCE (2000).

Using formula (111), it can be seen that if $1 + g > 0$ and $\delta + g > 0$, then V_U^0/V_A^0 will be greater than unity if

$$r^0 > g(1 - \delta)/(1 + g), \tag{112}$$

a condition that will usually be satisfied.[131] Thus under normal conditions and over a longer time horizon, household expenditures on consumer durables using the user cost approach will tend to exceed the corresponding money outlays on new purchases of the consumer durable. The difference between the two approaches will tend to grow as the life of the asset increases (i.e., as the depreciation rate δ decreases).

To get a rough idea of the possible magnitude of the value ratio for the two approaches, V_U^0/V_A^0, we evaluate (111) for a "housing" example where the depreciation rate is 2 % (i.e., $\delta = .02$), the real interest rate is 4 % (i.e., $r^0 = .04$) and the growth rate for the production of new houses is 1 % (i.e., $g = .01$). In this base case, the ratio of user cost expenditures on housing to the purchases of new housing in the same period, V_U^0/V_A^0, is 2.02. If we *increase* the depreciation rate to 3 %, then V_U^0/V_A^0 *decreases* to 1.77; if we *decrease* the depreciation rate to 1 %, then V_U^0/V_A^0 *increases* to 2.53. Again looking at the base case, if we *increase* the real interest rate to 5 %, then V_U^0/V_A^0 *increases* to 2.36 while if we *decrease* the real interest rate to 3 %, then V_U^0/V_A^0 *decreases* to 1.68. Finally, if we *increase* the growth rate for new houses to 2 %, then V_U^0/V_A^0 *decreases* to 1.53 while if we *decrease* the growth rate to 0, then V_U^0/V_A^0 *increases* to 3.00. *Thus an acquisitions approach to housing in the CPI is likely to give about one half the expenditure weight that a user cost approach would give.*

For shorter lived assets, the difference between the acquisitions approach and the user cost approach will not be so large and hence justifies the acquisitions approach as being approximately "correct" as a measure of consumption services.[132]

We conclude this section by listing some of the problems and difficulties that might arise in implementing a user cost approach to purchases of owner occupied housing.

131. Note that if the real interest rate r_0 equals g, the real rate of growth in housing investment, then from (111), $V_U^0/V_A^0 = (1 + g)$ and the acquisitions approach will be more or less equivalent to the user cost approach over the long run.

132. The simplified Icelandic user cost approach should be considered for other consumer durables as well. In formula (111), let $r^0 = .04, g = .01$ and $\delta = .15$ and under these conditions, $V_U^0/V_A^0 = 1.20$; i.e., for a declining balance depreciation rate of 15 %, the user cost approach will give us an estimated value of consumption that is 20 % higher than the acquisitions approach under the conditions specified. Thus for consumer durable depreciation rates that are lower than 15 %, it would be useful for the statistical agency to produce Icelandic user costs for these goods and for the national accounts division to produce the corresponding consumption flows as "analytic series". It should be noted that this extends the present national accounts treatment of housing to other long lived consumer durables. Note also that this revised treatment of consumption in the national accounts would tend to make rich countries richer, since poorer countries hold fewer long lived consumer durables on a per capita basis.

- It is difficult to determine what the relevant nominal interest rate r^0 is for each household. It may be necessary to simply use a benchmark interest rate that would be determined by either the government, a national statistical agency or an accounting standards board.
- It is difficult to determine what the relevant depreciation rate is for housing.[133] For example, the Bureau of Economic Analysis in the U.S. assumes that the declining balance depreciation rate for housing in the U.S. is 1.2% per year. Using Statistics Canada data on investment in residential construction in Canada for the years 1926 to 1999 as well as data from the Statistics Canada National Balance Sheet Accounts on the value of residential structures, we estimate that the comparable declining balance depreciation rate for residential structures in Canada is 2.9% per year. This is a huge difference in depreciation rates and is probably not justified.
- The user cost of housing is made up of two main parts: the user cost of structures and the user cost of the land that the dwelling sits on. Constructing the user cost of land is simpler than constructing the user cost of structures since we can assume that the land depreciation rate is zero. However, we still have to worry about the treatment of capital gains on land.
- *Ex post user costs* will be too volatile to be acceptable to users[134] and hence an *ex ante user cost* concept will have to be used. This creates difficulties in that different national statistical agencies will generally make different assumptions and use different methods in order to construct forecasted structures and land inflation rates and hence the resulting ex ante user costs of housing may not be comparable across countries.[135]
- The user cost formula (106) must be generalized to accommodate various taxes that may be associated with the purchase of a durable or with the continuing use of the durable.[136]
- A final problem with the user cost approach to valuing the services of owner occupied housing concerns the treatment of renovation expenditures. In most cases, renovation expenditures will yield a benefit to the home owner for a period longer than a

133. As mentioned earlier, it is not necessary to assume declining balance depreciation in the user cost approach: any pattern of depreciation can be accommodated, including one hoss shay depreciation, where the durable yields a constant stream of services over time until it is scrapped. See DIEWERT and LAWRENCE (2000) for some empirical examples for Canada using different assumptions about the form of depreciation.
134. GOODHART (2001; F351) comments on the practical difficulties of using ex post user costs for housing as follows: "An even more theoretical user cost approach is to measure the cost foregone by living in an owner occupied property as compared with selling it at the beginning of the period and repurchasing it at the end ... But this gives the absurd result that as house prices rise, so the opportunity cost falls; indeed the more virulent the inflation of housing asset prices, the more negative would this measure become. Although it has some academic aficionados, this flies in the face of common sense; I am glad to say that no country has adopted this method."
135. For additional material on the difficulties involved in constructing ex ante user costs, see DIEWERT (1980; p. 475–486). For empirical comparisons of different user cost formulae, see HARPER, BERNDT and WOOD (1989) and DIEWERT and LAWRENCE (2000).
136. For example, property taxes are associated with the use of housing services.

year and hence, in principle, all such expenditures should be capitalized and depreciated over time.[137]

In view of the difficulties involved in obtaining comparable user costs across EU countries, it may be useful to implement the very simple version of the theory that is used by Iceland in its CPI. In this implementation of user cost theory, the following simplified user cost formula is used:

$$p_S^t = (r^* + \delta_S + \tau_S^t)P_S^t \tag{113}$$

where p_S^t is the period t user cost for housing structures, P_S^t represents a period t price index for new structures, r^* is an assumed real interest rate, δ_S is an assumed declining balance depreciation rate for structures and τ_S^t is the period t property tax rate on structures.[138] In addition to the above structures user cost, there is also a user cost for the land that the structures sit on:

$$p_L^t = (r^* + \tau_L^t)P_L^t \tag{114}$$

where p_L^t is the period t user cost for housing land, P_L^t represents a period t price index for housing land, r^* is the assumed real interest rate and τ_L^t is the period t property tax rate on housing land.[139]

While the above Icelandic user cost of housing approach to the treatment of owner occupied housing is not conceptually perfect, it does give a reasonable approximation to an ex ante user cost approach. Moreover, this approach would be comparable across countries (if the same real interest rate r^* were chosen and if the structures depreciation rates δ_S were not too different) and it would probably be acceptable to users, since users in Iceland have not complained.[140]

137. "Normal" maintenance expenditures can be immediately expensed and hence should appear as a separate CPI category (that can be associated with housing expenditures). However, it may be difficult to distinguish between renovation expenditures (that can be capitalized and then depreciated over time) and maintenance expenditures. Also, if there are changes over time in the intensity of maintenance expenditures or renovation expenditures, this will affect the depreciation rate for housing structures.
138. This term could also include insurance premiums for the structure.
139. The Icelandic owner occupied housing user cost model in actual fact has only one user cost that covers both the structure and the land that the structure sits on. The real interest rate that is used is approximately 4 % per year and the combined depreciation rate for land and structures is assumed to equal 1.25 % per year. The depreciation rate for structures alone is estimated to be 1.5 % per year. Property taxes are accounted for separately in the Icelandic CPI. Housing price information is provided by the State Evaluation Board based on property sales data of both new and old housing. The SEB also estimates the value of the housing stock and land in Iceland, using a hedonic regression model based on property sales data. The value of each household's dwelling is collected in the Household Budget Survey.
140. Personal communication with Rósmundur Gudnason from Statistics Iceland. Gudnason also notes that the housing rentals part of the Icelandic CPI has closely tracked the user cost estimates for owner occupied housing in recent years. Gudnason also states that "the Icelandic Central Bank considers the inclusion of housing in this way into the CPI as one of the most important parts for their monetary targeting."

Given the importance of owner occupied housing in many EU countries, the recommendations that seem to follow from the discussion above are:

- The HICP should attempt to implement the acquisitions approach, the rental equivalence approach and the simplified Icelandic user cost approach to the treatment of owner occupied housing. Users can then decide which approach best suits their purpose.
- Any one of the three approaches could be chosen as the approach that would be used in the "headline" HICP. The other two approaches would be made available to users as "analytic tables".
- It may be that the rental equivalence approach fails due to thinness (or nonrepresentativeness) of rental markets for some types of owner occupied housing. However, I would still recommend that the HICP construct rental equivalence estimates for the value of owner occupied housing services since these estimates will be needed for national accounts purposes in any case. Users could be alerted to the weakness of these estimates.

We turn now to our final difficult measurement issue in the HICP (or in any national CPI for that matter): namely, the treatment of seasonal commodities in the index.

7.6. The Treatment of Seasonal Commodities

Seasonal commodities are commodities whose consumption varies substantially and systematically as the month of the year changes. A *strongly seasonal commodity* is one where the commodity is simply not available at certain seasons of the year. Seasonal commodities usually comprise 20 to 30 per cent of the commodities in a typical CPI.

Obviously, strongly seasonal commodities cause difficulties for the HICP or any monthly CPI: how can we compare the price of a commodity in a month when it is available to its (nonexistent) price in a month when it is not available in the marketplace? Even if a seasonal commodity is available for all months in a year, the fact that monthly quantities vary substantially creates difficulties for a typical CPI since the base period expenditure weights are usually annual average expenditures for that base year. Hence for at least some months of the year, these annual average weights will not reflect actual base period expenditures for that month for the seasonal commodity under consideration.

What are possible solutions to the problem of seasonal commodities in the CPI? [141]

141. We will not cover imputation techniques as a possible solution to the seasonal unavailability of prices. For recent surveys of imputation techniques, see ARMKNECHT and MAITLAND-SMITH (1999) and FEENSTRA and DIEWERT (2000).

If our goal is to construct an annual index, then there is a satisfactory theoretical solution that will enable us to deal with seasonal commodities. This solution is due to MUDGETT (1955) and STONE (1956): simply regard each commodity in each month as a *separate* commodity and then use normal index number methodology to compare the twelve months of price data in the current year with the corresponding twelve months of data in the base year. Thus if there are N monthly commodities in the domain of definition of the index, the index number formula will compare the $12N$ prices in the current year with the corresponding $12N$ prices in the base year.

DIEWERT (1983c) took the MUDGETT STONE approach one step further: he argued that the price data pertaining to the last 12 months could be compared with the corresponding monthly data in the base year so that each month, the statistical agency could produce such a *moving year price index*. The resulting monthly series is a nice seasonally adjusted series but it is not subject to the arbitrariness that plagues existing seasonal adjustment procedures.[142] The main disadvantage of this method is that it requires *monthly population expenditure information* for the base year if the usual Laspeyres methodology is used and if a superlative index is calculated, then monthly expenditure weight information is required on an ongoing basis. Unfortunately, the typical consumer expenditure survey that collects population expenditure weights is usually quite expensive and not particularly accurate. Thus in order to calculate moving year superlative indexes, it will be necessary for the HICP to invest in a continuing consumer expenditure survey or to make use of national accounts information and produce these moving year superlative indexes with a lag.

There is another major problem with the moving year price index concept: namely, it will not tell us very much about short term month to month movements in prices. *Hence in addition to constructing moving year indexes, it will be necessary to also construct a separate month to month index but omitting strongly seasonal commodities.* There are a number of ways in which this month to month index could be implemented. Conceptually, the "best" way of proceeding would be to construct month to month chained superlative indexes using current price and quantity information on the set of commodities that are available in the two consecutive months. However, given the difficulties in obtaining current month quantity weights on a timely basis, it may be necessary to use lagged monthly quantity weights or lagged monthly expenditure shares as proxies for current period information.[143] If lagged expenditure data are not available, then

142. In order to obtain a centered seasonally adjusted series, it will be necessary to wait 7 months. However, traditional statistical methods of seasonal adjustment usually require 18 months of additional data before seasonally adjusted estimates are finalized. For an empirical example of the moving year method and its theoretical consistency with economic theory, see DIEWERT (1996) (1999). For additional examples of the method and a discussion of the advantages and disadvantages of this index number method of seasonal adjustment compared to traditional econometric methods of seasonal adjustment, see ALTERMAN, DIEWERT and FEENSTRA (1999).

143. Alternatively, various forecasting methods could be used in order to predict current period expenditure shares.

monthly expenditure data for a base year can be used to construct *Laspeyres like or Paasche like* indexes for consecutive months in the current year. Thus let $s_n^{0,m}$ be the expenditure share of commodity n in month m of the base year 0 (so that $\sum_{n=1}^{N} s_n^{0,m} = 1$ for $m = 1, \ldots, 12$). Using the base year expenditure share for commodity n in month m, $s_n^{0,m}$, as an approximation to the corresponding year t expenditure share for commodity n in month m, $s_n^{t,m}$, leads to the following *approximate chain link Laspeyres index* comparing prices in month m to month $m + 1$ in year t:

$$P_{AL}^{t,m,m+1} \equiv \sum_{n=1}^{N} s_n^{0,m} p_n^{t,m+1} / p_n^{t,m}; \quad m = 1, 2, \ldots, 11 \tag{115}$$

where $p_n^{t,m}$ is the price of commodity n in month m of year t.[144] In a similar manner, using the base year expenditure share for commodity n in month $m + 1$, $s_n^{0,m+1}$, as an approximation to the corresponding year t expenditure share for commodity n in month $m + 1$, $s_n^{t,m+1}$, leads to the following *approximate chain link Paasche index* comparing prices in month m to month $m + 1$ in year t:

$$P_{AP}^{t,m,m+1} \equiv \left\{ \sum_{n=1}^{N} s_n^{0,m+1} \left[p_n^{t,m+1} / p_n^{t,m} \right]^{-1} \right\}^{-1}; \quad m = 1, 2, \ldots, 11. \tag{116}$$

Once approximate Laspeyres and Paasche links have been constructed, we can define the *approximate chain link Fisher index* comparing prices in month m to month $m + 1$ in year t as the geometric mean of the two indexes defined by (115) and (116) above:

$$\begin{aligned} &P_{AF}(p^{t,m}, p^{t,m+1}, s^{0,m}, s^{0,m+1}) \\ &\equiv \left\{ \sum_{n=1}^{N} s_n^{0,m} \left[p_n^{t,m+1} / p_n^{t,m} \right] \right\}^{1/2} \left\{ \sum_{n=1}^{N} s_n^{0,m+1} \left[p_n^{t,m+1} / p_n^{t,m} \right]^{-1} \right\}^{-1/2} \end{aligned} \tag{117}$$

where $p^{t,m}$ and $p^{t,m+1}$ are vectors of the year t prices in months m and $m + 1$ respectively and $s^{0,m}$ and $s^{0,m+1}$ are vectors of the year 0 expenditure shares in months m and $m + 1$ respectively.

The advantage of the approximate Fisher index defined by (117) over the approximate Laspeyres and Paasche indexes defined by (115) and (116) is that the former index satisfies the following important *time reversal test* whereas the latter indexes do not:

$$P_{AF}(p^{t,m}, p^{t,m+1}, s^{0,m}, s^{0,m+1}) P_{AF}(p^{t,m+1}, p^{t,m}, s^{0,m+1}, s^{0,m}) = 1. \tag{118}$$

Thus using the approximate Fisher index, it does not matter whether we use month m or month $m + 1$ as the base month: we get essentially the same answer either way.

144. If $m = 12$, then the price ratios in (115) become $p_n^{t+1,1} / p_n^{t,12}$.

The same approximation of current month expenditure shares by base month expenditure shares could be used in order to define an *approximate Törnqvist Theil chain link index*, P_{AT}, that compares the prices in month m to month $m+1$ of year t:[145]

$$\ln P_{At}(p^{t,m}, p^{t,m+1}, s^{0,m}, s^{0,m+1}) \equiv \sum_{n=1}^{N} (1/2)(s_n^{0,m} + s_n^{0,m+1}) \ln[p_n^{t,m+1}/p_n^{t,m}]; \quad m = 1, 2, \ldots, 11. \tag{119}$$

Finally, recalling formula (23) for the Walsh price index, we can define an *approximate Walsh chain link index*, P_{AW}, that compares the prices in month m to month $m+1$ of year t as follows:

$$P_{AW}(p^{t,m}, p^{t,m+1}, s^{0,m}, s^{0,m+1}) \equiv \sum_{n=1}^{N} (s_n^{0,m} s_n^{0,m+1})^{1/2} [p_n^{t,m+1}/p_n^{t,m}]^{1/2} / \sum_{j=1}^{N} (s_j^{0,m} s_j^{0,m+1})^{1/2} [p_j^{t,m}/p_j^{t,m+1}]^{1/2}. \tag{120}$$

It is straightforward to show that the approximate Walsh and approximate Törnqvist Theil indexes defined by (119) and (120) both satisfy the time reversal test (118).

If current monthly expenditure share information is not available, then in order to deal adequately with seasonal commodities in the context of producing a short term month to month consumer price index, I would recommend the use of monthly chaining, using one of the three approximate formulae P_{AF}, P_{AT} or P_{AW} defined by (118)–(120) above.[146] If monthly expenditure shares do not change much going from the base year 0 to the current year t, then these approximate indexes will approximate their Fisher, Törnqvist Theil and Walsh counterparts fairly closely and the latter indexes were the three indexes that emerged as being "best" from four different approaches to index number theory. This is the first major recommendation that emerges from the analysis that was presented in this section.

A short term month to month CPI cannot deal adequately with strongly seasonal commodities; i.e., commodities which are not available in all months of the year. In order to deal with this problem, our second recommendation is that a moving year index be produced, where the prices in the past 12 months would be compared with their counterpart seasonal prices in a base year. Ideally, one of our three "best' index number formulae would be used in order to construct this moving year index but since information

145. As usual, if $m = 12$, then the price ratios in (119) become $p_n^{t+1,1}/p_n^{t,12}$.
146. A word of caution is in order here. These month to month chained indexes should be cumulated over 12 months and compared to their year over year counterparts, which will normally be much more accurate. If the cumulated indexes differ considerably from their year over year counterparts, then at least some of the seasonal commodities should be dropped from the domain of definition of the month to month index until there is a reasonable correspondence between the cumulated month to month indexes and their year over year counterparts.

on current year expenditures is not likely to be available on a timely basis, it may be useful to construct a preliminary version of this moving year index. Such a preliminary version could be constructed adapting the techniques we used in order to construct the three approximate formulae, (118)-(120). In other words, the prices pertaining to the last 12 months would be compared to their base year counterparts but the monthly expenditure shares that are used in the various formulae would be replaced by the monthly expenditure shares for the most recent year available.

Our final recommendation emerges from the prior two recommendations. Both the short term month to month index and the moving year index will require monthly information on consumer expenditures, ideally on an ongoing basis. Hence our final recommendation is that an ongoing consumer expenditure survey be funded that is large enough so that monthly expenditure shares can be estimated with some degree of accuracy.

8. CONCLUSION

In this paper, we have discussed various domains of definition that might be used as the set of transactions for a harmonized inflation index. Our preference is to use the broad domains that are suggested by the system of national accounts (rather than inventing a new index that does not fit into the national accounting framework), noting that imports should be treated as a primary input rather than as a negative export. In practice, the set of consumer final demand expenditures C (or a somewhat larger aggregate of final demand expenditures up to $C + G + I + X$) is a suitable domain of definition for a harmonized inflation index.

Since central bankers and monetary economists are not usually specialists in index number theory, we reviewed four main approaches to index number theory in some detail in section 5 above. These four approaches led to three index number formulae as being "best" and fortunately, these three formulae will approximate each other quite closely using normal time series data. Hence, it is not necessary to make a definite choice between the four alternative approaches.

In section 6, we reviewed the methodology used by the HICP and pointed out a few problems that should perhaps be addressed.

In section 7, we looked at the main problems that make the construction of a Harmonized Index of Consumer Prices difficult. These problem areas are:

- The treatment of quality change;
- Variations in the frequency of rebasing across countries;
- The use of the Laspeyres formula at higher levels of aggregation which is subject to substitution or representativity bias;
- The lack of quantity or expenditure weights at lower levels of aggregation that seems to lead to a substantial overestimate of inflation (in many cases) at the elementary level of aggregation;

- The treatment of owner occupied housing; and
- The treatment of seasonal commodities.

We have attempted to address many of the above problems in the text above. Even if our suggested solutions turn out to be off the mark, there is a need for the HICP to provide users with a systematic overview of its methodology.[147]

Finally, we noted that in many cases, there is a need for users to have access to more than one index; e.g., recall our discussions of owner occupied housing and the treatment of seasonal commodities. I have no problem with the HICP declaring that one particular index should be its "official" index, but I would hope that alternative indexes could be provided as analytical tables for interested users.

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.

ARMKNECHT, P.A. and F. MAITLAND-SMITH (1999), "Price Imputation and Other Techniques for Dealing with Missing Observations, Seasonality and Quality Changes in Price Indices", IMF Working Paper, 99/78, Statistics Department, International Monetary Fund, Washington D.C.

ALCHIAN, A.A. and B. KLEIN (1973), "On a Correct Measure of Inflation", *Journal of Money, Credit and Banking*, 5 (1), p. 173–191.

ARROW, K.J. (1953), "Le role des valeurs boursières pour le répartition la meilleure des risques", in: *Econométrie*, Paris, p. 41–48; reprinted in English as: "The Role of Securities in the Optimal Allocation of Risk Bearing", *The Review of Economic Studies*, 31 (1964), p. 91–96.

ASTIN, J. (1999), "The European Union Harmonized Indices of Consumer Prices (HICP)", in: R. Gudnason and Thora Gylfadottir (eds.), *Proceedings of the Ottawa Group Fifth Meeting*, text available at: http://www.statcan.ca/secure/english/ottawa group, also published in: *Statistical Journal of the United Nations ECE*, 16 (1999), p. 123–135.

ARCHIBALD, R.B. (1977), "On the Theory of Industrial Price Measurement: Output Price Indexes", *Annals of Economic and Social Measurement*, 6, p. 57–72.

BALK, B. (1994), "On the First Step in the Calculation of a Consumer Price Index", in: *International Conference on Price Indices: Papers and Final Report*, First Meeting of the International Working Group on Price Indices, November, Ottawa, p. 129–145.

BALK, B.M. (1995), "Axiomatic Price Index Theory: A Survey", *International Statistical Review*, 63 (1), p. 69–93.

147. WYNNE (2001) also notes some of the areas where it would be useful for the HICP to provide more documentation on the methods it uses.

BALK, B.M. (1996), "Consistency in Aggregation and Stuvel Indices", *The Review of Income and Wealth*, 42, p. 353–363.

BALK, B.M. (1998), *Industrial Price, Quantity and Productivity Indices*, Boston.

BALK, B.M. (2000), "DIVISIA Price and Quantity Indexes 75 Years After", Department of Statistical Methods, Statistics Netherlands, Voorburg NL.

BALK, B.M. (2001), "The Residual: On Monitoring and Benchmarking Firms, Industries and Economies with respect to Productivity", Inaugural Addresses Research in Management Series, Erasmus University Rotterdam, Rotterdam.

BALK, B.M. and W.E. DIEWERT (2001), "A Characterization of the Törnqvist Price Index", *Economic Letters*, 72, p. 279–281.

BERGLUND, A. (1999), "New Inflation Measure Used as Main Indicator in the ECB/ESCB Monetary Policy for the Euro-Zone", in: M. Silver and D. Fenwick (eds.), *Proceedings of the Measurement of Inflation Conference*, Cardiff UK, p. 67–79.

BERNDT, E.R. (1991), *The Practice of Econometrics: Classic and Contemporary*, Reading.

BERNDT, E.R. and N.J. RAPPAPORT (2001), "Price and Quality of Desktop and Mobile Personal Computers: A Quarter Century Historical Overview", *The American Economic Review*, forthcoming.

BOVER, O. and P. VELLILA (2001), "Hedonic House Prices without Characteristics: The Case of New Multi-unit Housing", Paper presented at the Joint Meeting of the Centre for Economic Policy Research and the European Central Bank: Issues in the Measurement of Price Indices, November 16–17, Frankfurt.

BOWLEY, A.L. (1899), "Wages, Nominal and Real", in: R.H.I. Palgrave (ed.), *Dictionary of Political Economy*, 3, London, p. 640–651.

BRADLEY, R., B. COOK, S.E. LEAVER and B.R. MOULTON (1997), "An Overview of Research on Potential Uses of Scanner Data in the U. S. CPI", Third Meeting of the International Working Group on Price Indices, Statistics Netherlands, April 16–18, Voorburg.

BUREAU OF ECONOMIC ANALYSIS (2001), "BEA News Release: Gross domestic Product, October 31, 2001", BEA, U.S. Department of Commerce, http://www.bea.doc.-gov/newsrel/gdp301 a.htm.

BUREAU OF LABOR STATISTICS (1984), *BLS Handbook of Methods, Volume 2: The Consumer Price Index*, Bulletin 2134–2, Washington D.C.

CARLI, GIAN-RINALDO (1804), "Del valore e della proporzione de' metalli monetati", in: *Scrittori classici italiani di economia politica*, 13, Milano (originally published in 1764), p. 297–366.

CARRUTHERS, A.G., D.J. SELLWOOD and P.W. WARD (1980), "Recent Developments in the Retail Prices Index", *The Statistician*, 29, p. 1–32.

CAVES, D.W., L.R. CHRISTENSEN and W.E. DIEWERT (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica*, 50, p. 1393–1414.

CECCHETTI, S.G. (1997), "Measuring Short-Run Inflation of Central Bankers", *Federal Reserve Bank of St. Louis Review*, 79 (3), p. 143–155.

CHRISTENSEN, L.R. and D.W. JORGENSON (1969), "The Measurement of U.S. Real Capital Input, 1929–1967," *Review of Income and Wealth*, 15 (4), p. 293–320.

CHRISTENSEN, L.R., D.W. JORGENSON and L. J. LAU (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica*, 39, p. 255–256.

CHURCH, A.H. (1901), "The Proper Distribution of Establishment Charges, Part III", *The Engineering Magazine*, 21, p. 904–912.

DALÉN, J. (1992), "Computing Elementary Aggregates in the Swedish Consumer Price Index", *Journal of Official Statistics*, 8, p. 129–147.

DALÉN, J. (1994), "Sensitivity Analyses for Harmonizing European Consumer Price Indices", in: *International Conference on Price Indices: Papers and Final Report*, First Meeting of the International Working Group on Price Indices, November, Ottawa, p. 147–171.

DALÉN, J. (1997), "Experiments with Swedish Scanner Data", Third Meeting of the International Working Group on Price Indices, April 16–18, Voorburg.

DAVIES, G.R. (1924), "The Problem of a Standard Index Number Formula", *Journal of the American Statistical Association*, 19, p. 180–188.

DAVIES, G.R. (1932), "Index Numbers in Mathematical Economics", *Journal of the American Statistical Association*, 27, p. 58–64.

DEBREU, G. (1959), *Theory of Value*, New York.

DE HAAN, J. and E. OPPERDOES (1997), "Estimation of the Coffee Price Index Using Scanner Data: Simulation of Official Practices", Third Meeting of the International Working Group on Price Indices, April 16–18, Voorburg.

DENNY, M. (1974), "The Relationship Between Functional Forms for the Production System", *Canadian Journal of Economics*, 7, p. 21–31.

DIEWERT, W.E. (1974a), "Applications of Duality Theory," in: M. D. Intriligator and D.A. Kendrick (eds.), *Frontiers of Quantitative Economics*, II, Amsterdam, p. 106–171.

DIEWERT, W.E. (1974b), "Intertemporal Consumer Theory and the Demand for Durables", *Econometrica*, 42, p. 497–516.

DIEWERT, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics*, 4, p. 114–145.

DIEWERT, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica*, 46, p. 883–900.

DIEWERT, W.E. (1980), "Aggregation Problems in the Measurement of Capital", in: Dan Usher (ed.), *The Measurement of Capital*, Chicago, p. 433–528.

DIEWERT, W.E. (1983a), "The Theory of the Cost of Living Index and the Measurement of Welfare Change", in: W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa, p. 163–233; reprinted in: W. E. Diewert (ed.), *Price Level Measurement*, Amsterdam, 1990, p. 79–147.

DIEWERT, W.E. (1983b), "The Theory of the Output Price Index and the Measurement of Real Output Change", in: W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa, p. 1049–1113.

DIEWERT, W.E. (1983c), "The Treatment of Seasonality in a Cost of Living Index", in: W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa, p. 1019–1045.

DIEWERT, W.E. (1987), "Index Numbers", in: J. Eatwell, M. Milgate and P. Newman (eds.), *The New Palgrave: A Dictionary of Economics*, 2, London, p. 767–780; reprinted in: W.E. Diewert and A.O. Nakamura (eds.), *Essays in Index Number Theory*, 1, Amsterdam, 1993, p. 71–104.

DIEWERT, W.E. (1992a), "The Measurement of Productivity", *Bulletin of Economic Research*, 44 (3), p. 163–198.

DIEWERT, W.E. (1992b), "Fisher Ideal Output, Input and Productivity Indexes Revisited", *Journal of Productivity Analysis*, 3, p. 211–248; reprinted in: W.E. Diewert and A.O. Nakamura (eds.), *Essays in Index Number Theory*, 1, Amsterdam, 1993, p. 317–353.

DIEWERT, W.E. (1993a), "The Early History of Price Index Research", in: W.E. Diewert and A.O. Nakamura (eds.), *Essays in Index Number Theory*, 1, Amsterdam, p. 33–65.

DIEWERT, W.E. (1993b), "Symmetric Means and Choice under Uncertainty", in: W.E. Diewert and A.O. Nakamura (eds.), *Essays in Index Number Theory*, 1, Amsterdam, p. 355–433.

DIEWERT, W.E. (1993c), "Duality Approaches To Microeconomic Theory", in: W.E. Diewert and A.O. Nakamura (eds.), *Essays in Index Number Theory*, 1, Amsterdam, p. 105–175.

DIEWERT, W.E. (1995a), "Axiomatic and Economic Approaches to Elementary Price Indexes", Discussion Paper, 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. (1996a), "Seasonal Commodities, High Inflation and Index Number Theory", Discussion Paper, 96–06, Department of Economics, University of British Columbia, Vancouver, Canada.

DIEWERT, W.E. (1996b), "Comment on CPI Biases", *Business Economics*,31 (2), p. 30–35.

DIEWERT, W.E. (1997), "Commentary", *Federal Reserve Bank of St. Louis Review*, 79 (3), p. 127–137.

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

DIEWERT, W.E. (1999), "Index Number Approaches to Seasonal Adjustment", *Macroeconomic Dynamics*, 3, p. 1–21.

DIEWERT, W.E. (2000), "Notes on Producing an Annual Superlative Index Using Monthly Price Data", Discussion Paper, 00–08, Department of Economics, University of British Columbia, Vancouver, Canada, p. 30ff.

DIEWERT, W.E. (2001a), "The Consumer Price Index and Index Number Purpose", *Journal of Economic and Social Measurement*, forthcoming; also available on the web at: http://www.econ.ubc.ca/diewert/hmpgdie.htm.

DIEWERT, W.E. (2001b), "The Consumer Price Index and Index Number Theory: A Survey", Discussion Paper, 01–02, Department of Economics, University of British Columbia, Vancouver, Canada.

DIEWERT, W.E. (2001c), "Hedonic Regressions: A Consumer Theory Approach", Discussion Paper, 01–12, Department of Economics, University of British Columbia, Vancouver, Canada; forthcoming in: R. Feenstra and M. Shapiro (eds.), *Scanner Data and Price Indexes*, Studies in Income and Wealth, 61, Chicago.

DIEWERT, W.E. and K.J. Fox (1998), "The Measurement of Inflation after Tax Reform", *Economics Letters*, 61, p. 279–284.

DIEWERT, W.E. and K.J. Fox (1999), "Can Measurement Error Explain the Productivity Paradox?", *Canadian Journal of Economics*, 33 (2), p. 251–280.

DIEWERT, W.E. and K.J. Fox (2001), "The Productivity Paradox and Mismeasurement of Economic Activity", in: *Monetary Policy in a World of Knowledge Based Growth, Quality Change and Uncertain Measurement*, London, p. x-x.

DIEWERT, W.E. and D.A. LAWRENCE (2000), "Progress in Measuring the Price and Quantity of Capital", in: L.J. Lau (ed.), *Econometrics Volume 2: Econometrics and the cost of Capital: Essays in Honor of Dale W. Jorgenson*, Cambridge, MA, p. 273–326.

DIVISIA, F. (1926), *L'indice monetaire et la theorie de la monnaie*, Paris.

DROBISCH, M.W. (1871), "Ueber einige Einwürfe gegen die in diesen Jahrbüchern veröffentlichte neue Methode, die Veränderungen der Waarenpreise und des Geldwerths zu berechten", *Jahrbücher für Nationalökonomie und Statistik*, 16, p. 416–427.

DUTOT, C. (1738), *Réflections politiques sur les finances et le commerce*, 1, La Haye.

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

EDGEWORTH, F.Y. (1925), *Papers Relating to Political Economy*, 1, New York.

EUROSTAT (European Statistical Agency), World Bank, International Monetary Fund, Organization for Economic Cooperation and Development, and United Nations (1993), *System of National Accounts 1993*, United Nations, New York.

FEENSTRA, R.C. and W.E. DIEWERT (2000), "Imputation and Price Indexes: Theory and Evidence from the International Price Program", BLS Working Paper, 335, Office and Prices and Living Conditions, Bureau of Labor Statistics, Washington D.C.

FENWICK, D., A. BALL, M. SILVER and P.H. MORGAN (2002), "Price Collection and Quality Assurance of Item Sampling in the Retail Price Index: How Can Scanner Data Help?", forthcoming in: Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, 61, Chicago.

FIELDING, D. and P. MIZEN (2001), "The Relationship between Price Dispersion and Inflation: a Reassessment", European University Institute Working Paper, 2001/10, Badia Fiesolana, San Domenico, Italy.

F.M. and K. SHELL (1972), "The Pure Theory of the National Output Deflator", in: *The Economic Theory of Price Indexes*, New York, p. 49–113.

FISHER, I. (1911), *The Purchasing Power of Money*, London.

FISHER, I. (1922), *The Making of Index Numbers*, Boston.

FISHER, I. (1930), *The Theory of Interest*, New York.

FRISCH, R. (1930), "Necessary and Sufficient Conditions Regarding the form of an Index Number which shall Meet Certain of Fisher's Tests', *Journal of the American Statistical Association*, 25, p. 397–406.

FRISCH, R. (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica*, 4, p. 1–38.

GALIANI, F. (1751), *Della Moneta*, excerpts translated and published as: *Money*, in: A.E. Monroe (ed.), *Early Economic Thought*, Cambridge, 1930, p. 281–307.

GOODHART, C.A.E. (1995), "Price Stability and Financial Fragility", in: K. Sawamoto, Z. Nakajima and H. Taguchi (eds.), *Financial Stability in a Changing Environment*, London, p. 439–497.

GOODHART, C. (2001), "What Weights should be Given to Asset Prices in the Measurement of Inflation?", *The Economic Journal*, 111 (June), p. F335-F356.

GOODHART, C. and B. HOFMANN (2000), "Financial Variables and the Conduct of Monetary Policy", Sveriges Riksbank Working Paper, 112, Information Riksbank, Stockholm.

HARPER, M.J., E.R. BERNDT and D.O. WOOD (1989), "Rates of Return and Capital Aggregation Using Alternative Rental Prices", in: D.W. Jorgenson and R. Landau (eds.), *Technology and Capital Formation*, Cambridge, MA, p. 331–372.

HAUSMAN, J.A. (1997), "Valuation of New Goods under Perfect and Imperfect Competition", in: T.F. Bresnahan and R.J. Gordon (eds.), *The Economics of New Goods*, Chicago, p. 209–237.

HAUSMAN, J.A. (1999), "Cellular Telephone, New Products and the CPI", *Journal of Business and Economic Statistics*, 17, p. 1–7.

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

HICKS, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, p. 105–140.

HICKS, J.R. (1946), *Value and Capital*, Second Edition, Oxford.

HILL, R.J. (2000), "Superlative Index Numbers: Not All of them Are Super", School of Economics, University of New South Wales, Sydney.

HILL, T.P. (1971), *The Measurement of Real Product*, Paris.

HILL, T.P. (1988), "Recent Developments in Index Number Theory and Practice", *OECD Economic Studies*, 10, p. 123–148.

HILL, T.P. (1993), "Price and Volume Measures", in: Eurostat, IMF, OECD, UN and World Bank (eds.), *System of National Accounts 1993*, Luxembourg, Washington, D.C., Paris, New York, p. 379–406.

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

HILL, P. (1998), "The Measurement of Inflation and Changes in the Cost of Living", *Statistical Journal of the United Nations*, ECE 15, p. 37–51.

HOFFMAN, J. and C. KURZ (2001), "Rent Indices for Housing in West Germany 1985–1998", Paper presented at the Joint Meeting of the Centre for Economic Policy Research and the European Central Bank: Issues in the Measurement of Price Indices, November 16–17, Frankfurt.

HULTEN, C.R. (1990), "The Measurement of Capital", in: E. R. Berndt and J. E. Triplett (eds.), *Fifty Years of Economic Measurement*, Chicago, p. 119–58.

ISSING, O. (2001), "Why Price Stability?", in: A.C. Herrero, V. Gaspar, L. Hoogduin, J. Morgan and B. Winkler (eds.), *Why Price Stability*, Proceedings of the First ECB Central Banking Conference, Frankfurt, p. 179–202.

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

JORGENSON, D.W. and Z. GRILICHES (1967), "The Explanation of Productivity Change", *The Review of Economic Studies*, 34, p. 249–283.

KOHLI, U. (1978), "A Gross National Product Function and the Derived Demand for Imports and Supply of Exports", *Canadian Journal of Economics*, 11, p. 167–182.

KOHLI, U.R. (1982), "Relative Price Effects and the Demand for Imports", *The Canadian Journal of Economics*, 15, p. 203–219.

KOHLI, U. (1983), "Technology and the Demand for Imports", *Southern Economic Journal*, 50, p. 137–150.

KOHLI, U. (1991), *Technology, Duality and Foreign Trade*, Ann Arbor.

KONÜS, A.A. (1924), "The Problem of the True Index of the Cost of Living", translated in: *Econometrica*, 7 (1939), p. 10–29.

KNIBBS, SIR G.H. (1924), "The Nature of an Unequivocal Price Index and Quantity Index", *Journal of the American Statistical Association*, 19, p. 42–60 and p. 196–205.

LEONTIEF, W. (1936), "Composite Commodities and the Problem of Index Numbers", *Econometrica*, 4, p. 39–59.

LLOYD, P.J. (1975), "Substitution Effects and Biases in Nontrue Price Indices", *American Economic Review* 65, p. 301–313.

MARSHALL, A. (1887), "Remedies for Fluctuations of General Prices', *Contemporary Review*, 51, p. 355–375.

MARSHALL, A. (1898), *Principles of Economics*, Fourth Edition, London.

MATHESON, E. (1910), *The Depreciation of Factories and their Valuation*, Fourth Edition, London.

MEYER, L.H. (2001), "Inflation Targets and Inflation Targeting", Remarks by Governor Laurence H. MEYER at the University of California at San Diego Economics Roundtable, July 17; available at: www.federalreserve.gov

MOULTON, B.R. (1996), "Constant Elasticity Cost-of-Living Index in Share Relative Form", Bureau of Labor Statistics, December, Washington D.C.

MUDGETT, B.D. (1955), "The Measurement of Seasonal Movements in Price and Quantity Indexes", *Journal of the American Statistical Association*, 50, p. 93–98.

OKAMOTO, M. (2001), "Mid-Year Basket Index as a Practical Approximation to a Superlative Index", Paper presented at the Sixth Meeting of the International Working Group on Price Indices, April 2–6, Canberra.

POLLAK, R.A. (1975), "The Intertemporal Cost of Living Index", *Annals of Economic and Social Measurement*, 4 (1), p. 179–195; reprinted in: R.A. Pollak, *The Theory of the Cost-of-Living Index*, Oxford, 1989, p. 70–89.

POLLAK, R.A. (1980), "Group Cost-of-Living Indexes", *American Economic Review*, 70, p. 273–278.

POLLAK, R.A. (1981), "The Social Cost-of-Living Index", *Journal of Public Economics*, 15, p. 311–336.

POLLAK, R.A. (1983), "The Theory of the Cost-of-Living Index", in: W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa, p. 87–161; reprinted in: R.A. Pollak, *The Theory of the Cost-of-Living Index*, Oxford, 1989, p. 3–52; also reprinted in: W.E. Diewert (ed.), *Price Level Measurement*, Amsterdam, 1990, p. 5–77.

PRAIS, S.J. (1959), "Whose Cost of Living?", *The Review of Economic Studies*, 26, p. 126–134.

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

REINSDORF, M. and R. DORFMAN (1999), "The Sato-Vartia Index and the Monotonicity Axiom", *Journal of Econometrics*, 90, p. 45–61.

ROMER, P. (1994), "New Goods, Old Theory and the Welfare Costs of Trade Restrictions", *Journal of Development Economics*, 43, p. 5–38.

SAMUELSON, P.A. (1953), "Prices of Factors and Goods in General Equilibrium", *Review of Economic Studies*, 21, p. 1–20.

SAMUELSON, P.A. and S. SWAMY (1974), "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review*, 64, p. 566–593.

SCHULTZ, B. (1999), "Effects of Using Various Macro-Index Formulae in Longitudinal Price and Comparisons: Empirical Studies", in: W. Lane (ed.), *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, Bureau of Labor Statistics, U.S. Department of Labor, Washington D.C., p. 236–249.

SHAPIRO, M.D. and D.W. WILCOX (1997), "Alternative Strategies for Aggregating Prices in the CPI", *Federal Reserve Bank of St. Louis Review*, 79 (3), p. 113–125.

SHEPHARD, R.W. (1953), *Cost and Production Functions*, Princeton.

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

SILVER, M. (1999), "An Evaluation of the Use of Hedonic Regressions for Basic Components of Consumer Price Indices", *Review of Income and Wealth*, 45 (1), p. 41–56.

SILVER, M.S. and S. HERAVI (2001a), "The Measurement of Quality Adjusted Price Changes", forthcoming in: R. Feenstra and M. Shapiro (eds.), *Scanner Data and Price Indexes*, Studies in Income and Wealth, 61, Chicago.

SILVER, M.S. and S. HERAVI (2001b), "Scanner Data and the Measurement of Inflation", *The Economic Journal*, 111 (June), p. F384-F405.

SILVER, M.S. and S. HERAVI (2001c), "Hedonic Price Indices and the Matched Models Approach", Mimeo, Cardiff Business School, Cardiff, U.K.

SILVER, M.S. and S. HERAVI (2001d), "Why the CPI Matched Models Method May Fail Us: Results from a Hedonic and Matched Experiment Using Scanner Data", Paper presented at the Joint Meeting of the Centre for Economic Policy Research and the European Central Bank: Issues in the Measurement of Price Indices, November 16–17, Frankfurt.

SHIRATSUKA, S. (1999), "Asset Price Fluctuation and Price Indices", *Monetary and Economic Studies*, 17 (3), p. 103–128.

SIDGWICK, H. (1883), *The Principles of Political Economy*, London.

STONE, R. (1956), *Quantity and Price Indexes in National Accounts*, Paris.

TAYLOR, J.B. (1993), "Discretion versus Policy Rules in Practice", *Carnegie-Rochester Conferences Series on Public Policy*, 39 (December), p. 195–214.

TAYLOR, J.B., (ed.) (1999), *Monetary Policy Rules*, Chicago.

THEIL, H. (1967), *Economics and Information Theory*, Amsterdam.

TIERNEY, C. (1963), "The Index Number Problem", in: *Reporting the Financial Effects of Price-Level Changes*, Accounting Research Study, 6, New York, p. 61–117.

TÖRNQVIST, LEO (1936), "The Bank of Finland's Consumption Price Index," *Bank of Finland Monthly Bulletin*, 10, p. 1–8.

TÖRNQVIST, L. and E. TÖRNQVIST (1937), "Vilket är förhållandet mellan finska markens och svenska kronans köpkraft?", *Ekonomiska Samfundets Tidskrift*, 39, p. 1–39; reprinted in: *Collected Scientific Papers of Leo Törnqvist*, Helsinki, 1981, p. 121–160.

TRIPLETT, J.E. (1983), "Escalation Measures: What is the Answer? What is the Question?", in: W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa, p. 457–482.

TRIPLETT, J.E. (1987), "Hedonic Functions and Hedonic Indexes", in: *The New Palgrave: A Dictionary of Economics*, 2, London, p. 630–634.

TRIPLETT, J.E. (1990), "Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy", in: E.R. Berndt and J.E. Triplett (eds.), *Fifty Years of Economic Measurement*, Studies in Income and Wealth, 56, Chicago, p. 207–238.

TRIPLETT, J.E. (2001), "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?", *The Economic Journal*, 111 (June), p. F311-F334.

TRIPLETT, J.E. (2002), *Handbook on Quality Adjustment of Price Indexes for Information and Technology Products*, Paris.

VARTIA, Y.O. (1976), "Ideal Log Change Index Numbers", *Scandinavian Journal of Statistics*, 3, p. 121–126.

VOGT, A. (1980), "Der Zeit und der Faktorumkehrtest als 'Finders of Tests'", *Statistische Hefte*, 21, p. 66–71.

VON AUER, L. (2001), "An Axiomatic Checkup for Price Indices", Working Paper, 1/2001, Faculty of Economics and Management, Otto von Guericke University Magdeburg, Magdeburg, Germany.

WALRAS, L. (1954), *Elements of Pure Economics*, translated by W. Jaffe (first published in 1874), London.

WALSH, C.M. (1901), *The Measurement of General Exchange Value*, New York.

WALSH, C.M. (1921), *The Problem of Estimation*, London.

WOOLFORD, K. (1999), "Measuring Inflation: A Framework Based on Domestic Final Purchases", in: M. Silver and D. Fenwick, *Proceedings of the Measurement of Inflation Conference*, Cardiff, UK, p. 534–543.

WYNNE, M.A. (1997), "Commentary", *Federal Reserve Bank of St. Louis Review*, 79 (3), p. 161–167.

WYNNE, M.A. (2001), "Measurement Bias in the HICP: What do we Know and What do we need to Know?", Paper presented at the Joint Meeting of the Centre for Economic Policy Research and the European Central Bank: Issues in the Measurement of Price Indices, November 16–17, Frankfurt.

ZIESCHANG, K.D., A.M. BLOEM and P.A. ARMKNECHT (2001), "Price Indices for Inflation Targeting", paper presented at the Seminar on the Statistical Implications of Inflation Targeting at the International Monetary Fund, September 28, Washington D.C.

SUMMARY

The Harmonized Index of Consumer Prices (HICP) is the single most important indicator of inflation used by the European Central Bank. Sections 2 to 4 of the paper look at the theory of inflation indexes that could be used as target indexes of inflation. A Consumer Price Index (CPI) emerges as perhaps the most useful target index. Four different approaches to index number theory are reviewed and the "best" index number formula from each perspective is determined. Section 6 looks at the methodology of the HICP in the light of the previous sections. Section 7 looks at some of the difficult measurement problems that must be addressed in a CPI or an HICP. These problems include the treatment of quality change, substitution or representativity bias, chained versus fixed base indexes, the choice of formula at the lowest level of aggregation and the treatment of owner occupied housing and seasonal commodities.

ZUSAMMENFASSUNG

Der harmonisierte Konsumentenpreisindex (HICP) ist der wichtigste einzelne Indikator für die Inflation, der von der Europäischen Zentralbank verwendet wird. In den Anschnitten 2 bis 4 der Arbeit werden die aus theoretischer Sicht für ein Inflationsziel möglichen Indizes dargestellt. Dabei erweist sich der Konsumentenpreisindex (CPI) als vielleicht der für ein Inflationsziel als der geeignetste Index. Nach einer Übersicht über die vier verschiedenen Ansätze der Theorie der Indexzahlen wird in diesem Rahmen jeweils die "beste" Indexformel bestimmt. In Abschnitt 6 wird der HIPC vor diesem Hintergrund analysiert. Die schwierigen Messprobleme, die sich bei der Berechnung eines CIP oder HICP ergeben, sind Gegenstand von Abschnitt 7. Dabei handelt es sich um die Behandlung von Qualitätsveränderungen, die Substitutionsverzerrung, Wahl eines Kettenindex oder eines Index mit fester Basis, die Wahl der Formel auf dem tiefsten Aggregationsniveau, die Behandlung von Eigenmieten und saisonalen Gütern.

RÉSUMÉ

L'indice harmonisé des prix à la consommation (HICP) est le seul indice d'inflation utilisé par la banque centrale européenne. Les sections 2 et 4 de cet article présentent les indices d'inflation qui pourraient en théorie être utilisés comme des indices cibles d'inflation. L'indice des prix à la consommation (CPI) ressort comme l'indice cible probablement le plus adéquat. Après un aperçu de quatre différentes approches de la théorie des nombres d'indices, l'on déterminera la "meilleure" formule de chacune des perspectives. La section 6 analyse la méthodologie de l'HICP à la lumière des sections précédentes. Quelques problèmes de mesure difficile devant être pris en considération dans un CPI ou un HICP sont traités dans la section 7. Ces problèmes concernent la gestion du changement de qualité, de la déformation de substitution ou de représentativité, des indices en chaîne contre les indices à base fixe, le choix de formule au niveau le plus bas d'agrégation, la gestion de logements occupés par le propriétaire et les biens saisonniers.