INDEX NUMBER THEORY AND MEASUREMENT ECONOMICS

W. Erwin Diewert,[1] April, 2009.

CHAPTER 16: Index Number Theory: Past Progress and Problems for the Future

1. Introduction

The past quarter century has seen a remarkable amount of progress in both the theory and practice of index number theory and the closely related problems associated with the measurement of output, input and productivity. In section 2 below, we will review some of the significant developments in these areas over the past 30 years or so.

In section 3, we will take a look at some of the significant challenges that still face us in the *price measurement area* while in section 4, we will discuss some of the challenges that face measurement economists and price statisticians in measuring the *productivity performance* of establishments, firms, industries and economies.

2. Past Progress in the Measurement of Price and Quantity Change

In this section, we will discuss ten areas where progress in measuring price change has been made over the past 30 years.

2.1 Alternative Approaches to Index Number Theory Have Converged Substantially

Index number theory gives statistical agencies some guidance on what is the "right" theoretical *target index*.[2] The problem historically has been that there have been *many* alternative index number theories and so statistical agencies have been unable to agree on a single target index to guide them in the preparation of their consumer price indexes or their indexes of real output. Most of the *theoretical literature* on index numbers centers on the case where complete price and quantity information is available for *two* periods where it is desired to compare say the level of prices in one of the periods with those of the other period. This is called *bilateral* index number theory as opposed to *multilateral* index number theory, which deals with many periods instead of just two. However, multilateral approaches can readily be built up using bilateral index number theory. There are five main approaches to bilateral index number theory:

1. Fixed basket approaches and symmetric averages of fixed baskets;
2. The stochastic approach to index number theory;

---

[1] This chapter was presented at the ESRI "Conference on the Next Steps for the Japanese System of National Accounts: Towards More Accurate Measurement and More Comprehensive Accounts", held in Tokyo, March 24-25, 2005. The present paper draws freely on some recent papers on index number and economic measurement problems by the author; see Diewert (2001a) (2001b) (2003).

[2] Of course, the target index may not be achievable by a statistical agency but it is necessary to have some sort of theoretical target so that procedures can be adjusted so as to come closer to the target concept. Having a target concept is also necessary so that the index that is actually produced by a statistical agency can be evaluated from the perspective of how close the actual index comes to the theoretical ideal.

3. Test approaches;
4. The economic approach and
5. The approach of Divisia (1926).

Approaches 3 and 4 will be familiar enough to many price statisticians and expert users of the CPI but perhaps a few words about the other approaches are in order.

The Laspeyres index is an example of a fixed basket index. The problem from a theoretical point of view is that it has an equally valid "twin" between the same two periods under consideration, the Paasche index. If we have two equally valid estimators for the same concept, then statistical theory tells us to take the average of the two estimators in order to obtain a more accurate estimator. However, there is more than one way of taking an average so the question of the "best" average to take of the Paasche and Laspeyres indexes is not trivial. The new ILO (2004) CPI Manual suggests that the two "best" averages that emerge are the Fisher (1922) ideal and the Walsh (1901) (1921) price indexes.[3]

The unweighted stochastic approach to index number theory is also an easy one for price statisticians to follow: if we have lots of independent item price relatives between two periods, then some sort of average of them ought to be a pretty good estimator for the average amount of price change between the two periods. Moreover, this approach has the advantage of giving us a standard error for the estimated aggregate price change. Unfortunately, this straightforward stochastic approach neglects one key variable: namely, the *economic importance* of each price relative. Thus to get a more accurate stochastic approach to index number theory, it is necessary to bring into the picture *expenditure weights* for each item. When this is done, the Törnqvist (1936) Theil (1967) formula emerges as being perhaps "best" from the viewpoint of *weighted* stochastic approaches to index number theory.[4]

It turns out that the test and economic approaches to bilateral index number theory also end up endorsing the Fisher, Walsh and Törnqvist Theil price indexes as being "best" from their perspectives as well.[5]

The fifth approach to index number theory, the continuous time Divisia approach, does not lead to a single discrete time bilateral index number formula that is most consistent with this approach[6] so it provides little practical advice for statistical agencies, although it can be conceptually useful at times.[7]

---

[3] See Chapter 15 in the ILO (2004) CPI Manual. These recommendations are based on the research of Diewert (1997; 138) (2001c; section 7). The IMF (2004) PPI Manual contains the same basic information on index number theory.
[4] See Chapter 16 in the ILO (2004) CPI Manual.
[5] See Chapter 16 in the ILO (2004) CPI Manual. These three index number formulae are all examples of superlative index number formulae; see Diewert (1976) for an explanation of this concept and some examples of superlative formulae.
[6] See Frisch (1936; 7-9) and Chapter 15 of the ILO (2004) CPI Manual, where it is shown that both the Paasche and Laspeyres indexes can be regarded as discrete time approximations to the Divisia index.
[7] See the excellent survey paper on Divisia indexes by Balk (2000).

Thus four of the five major approaches to bilateral index number theory lead to the same three formulae as being best. Which formula should then be used by a statistical agency as their target index? It turns out that for "typical" time series data, it will not matter much, since the three indexes approximate each other very closely.[8]

The fact that four rather different approaches to index number theory lead to the same small number of index number formulae as being "best" and the fact that these formulae closely approximate each other for annual time series data has been a positive development. Fifteen years ago, measurement economists and price statisticians from North America tended to favor the economic approach to index number theory whereas their counterparts in Europe tended to favor the test[9] or stochastic approaches. This difference in views led to a great deal of counterproductive discussion on the relative merits of the various approaches to index number theory at international meetings on price measurement. Since for all practical purposes, the various approaches lead to the same small number of index number formulae as being "best", recent international meetings have been far more productive, with everyone focused on how to improve price measurement rather than fighting methodological wars.

2.2 New Insights into Fixed Base versus Chained Indexes

The *chain system*[10] measures the change in prices going from one period to a subsequent period using a bilateral index number formula involving the prices and quantities pertaining to the two adjacent periods. These one period rates of change (the links in the chain) are then cumulated to yield the relative levels of prices over the entire period under consideration. On the other hand, the *fixed base system* of price levels using the same bilateral index number formula P simply computes the level of prices in period t relative to the base period 0 in one step using the long term price relatives between the two periods.

For at least 70 years, economists and statisticians have been arguing about the relative merits of fixed base versus chained index numbers.[11] Thanks to the contributions of Szulc (1983), T.P. Hill (1988) (1993) and R.J. Hill (1995) (1999a) (1999b) (2001) (2004), I think that we have come to a much better understanding of the conditions when it will be useful to chain or not.

The main advantage of the chain system is that under normal conditions, chaining will reduce the spread between the Paasche and Laspeyres indexes.[12] These two indexes each

---

[8] See Diewert (1978; 888) on this point but see also Hill (2004) for some limitations of Diewert's results.

[9] The work of the Europeans Eichhorn and Voeller (1976), Eichhorn (1978) and Balk (1995) have been very influential in popularizing the test approach to index number theory.

[10] The chain principle was introduced independently into the economics literature by Lehr (1885; 45-46) and Marshall (1887; 373). Both authors observed that the chain system would mitigate the difficulties due to the introduction of new commodities into the economy, a point also mentioned by Hill (1993; 388). Fisher (1911; 203) introduced the term "chain system".

[11] See the discussion in Frisch (1936).

[12] See Diewert (1978; 895) and Hill (1988) (1993; 387-388).

provide an asymmetric perspective on the amount of price change that has occurred between the two periods under consideration and it could be expected that a single point estimate of the aggregate price change should lie between these two estimates. Thus under these as yet to be specified normal conditions, the use of either a chained Paasche or Laspeyres index will usually lead to a smaller difference between the two and hence to estimates that are closer to the "truth".

Hill (1993; 388), drawing on the earlier research of Szulc (1983) and Hill (1988; 136-137), noted that it is not appropriate to use the chain system when prices oscillate (or "bounce" to use Szulc's (1983; 548) term). This phenomenon can occur in the context of regular seasonal fluctuations or in the context of price wars. However, in the context of roughly monotonically changing prices and quantities, Hill (1993; 389) recommended the use of chained symmetrically weighted indexes.[13] The Fisher, Törnqvist and Walsh indexes are examples of symmetrically weighted indexes.

Under what conditions one should chain or not chain? Basically, one should chain if the prices and quantities pertaining to adjacent periods are *more similar* than the prices and quantities of more distant periods, since this strategy will lead to a narrowing of the spread between the Paasche and Laspeyres indexes at each link.[14] Of course, one needs a measure of how similar are the prices and quantities pertaining to two periods. The similarity measures could be *relative* ones or *absolute* ones. In the case of absolute comparisons, two vectors of the same dimension are similar if they are identical and dissimilar otherwise. In the case of relative comparisons, two vectors are similar if they are proportional and dissimilar if they are nonproportional.[15] Once a similarity measure has been defined, the prices and quantities of each period can be compared to each other using this measure and a "tree" or path that links all of the observations can be constructed where the most similar observations are compared with each other using a

---

[13] Note that all known superlative indexes are symmetrically weighted.

[14] Walsh in discussing whether fixed base or chained index numbers should be constructed, took for granted that the precision of all reasonable bilateral index number formulae would improve, provided that the two periods or situations being compared were more similar and hence, for this reason, favored the use of chained indexes: "The question is really, in which of the two courses [fixed base or chained index numbers] are we likely to gain greater exactness in the comparisons actually made? Here the probability seems to incline in favor of the second course; for the conditions are likely to be less diverse between two contiguous periods than between two periods say fifty years apart." Correa Moylan Walsh (1901; 206). Walsh (1921; 84-85) later reiterated his preference for chained index numbers. Fisher also made use of the idea that the chain system would usually make bilateral comparisons between price and quantity data that was more similar and hence the resulting comparisons would be more accurate: "The index numbers for 1909 and 1910 (each calculated in terms of 1867-1877) are compared with each other. But direct comparison between 1909 and 1910 would give a different and more valuable result. To use a common base is like comparing the relative heights of two men by measuring the height of each above the floor, instead of putting them back to back and directly measuring the difference of level between the tops of their heads." Irving Fisher (1911; 204). "It seems, therefore, advisable to compare each year with the next, or, in other words, to make each year the base year for the next. Such a procedure has been recommended by Marshall, Edgeworth and Flux. It largely meets the difficulty of non-uniform changes in the Q's, for any inequalities for successive years are relatively small." Irving Fisher (1911; 423-424).

[15] Diewert (2002b) takes an axiomatic approach to defining various indexes of absolute and relative dissimilarity.

bilateral index number formula.[16] Hill (1995) defined the price structures between the two countries to be more dissimilar the bigger is the spread between P_L and P_P; i.e., the bigger is max $\{P_L/P_P, P_P/P_L\}$. The problem with this measure of dissimilarity in the price structures of the two countries is that it could be the case that $P_L = P_P$ (so that the Hill measure would register a maximal degree of similarity) but the base period prices could be very different than the current period prices. Thus there is a need for a more systematic study of similarity (or dissimilarity) measures in order to pick the "best" one that could be used as an input into Hill's (1999a) (1999b) (2001) spanning tree algorithm for linking observations.[17] However, there is no doubt that the recent research by the Hills has put the question of whether to chain or not on a much more scientific basis. This is a very useful recent advance.

2.3 The Importance of Quality Change

Another element of progress in index number theory is the widespread recognition of the importance of adjusting prices for quality change. Thus there are a substantial number of papers that are now being devoted to this extremely important but conceptually difficult topic in recent years. I view this as a very positive development. It might be argued that this is not really a new development, since many of the early index number theorists were very concerned about the problem of introducing new goods into their preferred indexes.[18] Index number practitioners have also been interested in the problems of quality adjustment for a long time as well.[19] However, interest in this topic is now at unprecedented levels, perhaps due to the fact that about 2 per cent of the price quotes collected by a typical statistical agency in one month are no longer available in the following month. Some of these disappearing price quotes can be traced to seasonal and other factors but a substantial amount of the problem of disappearing quotes can be traced to new products replacing old products.

The main practical method that can be used to deal with quality change is the hedonic regression methodology, due initially to Court (1939) and popularized by Griliches (1971a) (1971b). A recent up to date survey of hedonic regression techniques is Triplett (2004).

In the Japanese context, the main contributor to the quality adjustment debate has been Shiratsuka while in the UK context, the main contributor has been Silver; see Shiratsuka

[16] Fisher (1922; 271-276) hinted at the possibility of using spatial linking; i.e., of linking countries that are similar in structure. However, the modern literature has grown due to the pioneering efforts of Robert Hill (1995) (1999a) (1999b) (2001). Hill (1995) used the spread between the Paasche and Laspeyres price indexes as an indicator of similarity and showed that this criterion gives the same results as a criterion that looks at the spread between the Paasche and Laspeyres quantity indexes.

[17] For a more systematic discussion of the properties of various measures of dissimilarity, see Diewert (2002).

[18] Walsh (1901; 207) argued that Lehr (1885; 45-46) was motivated to introduce the chain system in order to facilitate the introduction of new goods into price indexes. Marshall (1887; 373), Fisher (1911; 204) and Divisia (1926; 44-47) all made similar arguments for the use of the chain principle as an aid to introducing new products into price indexes.

[19] The Stigler (1961) report stressed the importance of quality change as did the more recent Boskin (1996) and Schultze and Mackie (2002) reports on the Consumer Price Index.

(1995a) (1995b) (1999a) (1999b) and Silver (1995) and Silver and Heravi (2001a) (2001b) (2002a) (2002b) (2003). We will discuss the problem of quality change in more detail in section 3.2 below.

2.4 The Usefulness of Multiple Consumer Price Indexes to Suit Different Purposes

Many years ago, Jack Triplett (1983) pointed out that more than one CPI may be required to meet the needs of different users.[20] For example, some users may require information on the month to month movement of prices in a timely fashion. This requirement leads to a Laspeyres type CPI along the lines of existing CPI's, where current information on weights is not necessarily available. However, other users may be more interested in a more accurate or representative measure of price change and may be willing to sacrifice timeliness for increased accuracy. Thus the Bureau of Labor Statistics in the U.S. is providing, on a delayed basis, a superlative index that uses current period weight information as well as base period weight information.[21] This is an entirely reasonable development, recognizing that different users have different needs. A second example where multiple indexes would be useful occurs in the context of the treatment of owner occupied housing. Researchers have made solid cases for at least three different treatments of owner occupied housing: the acquisitions approach (just price out purchases of new dwelling units), the rental equivalence approach (impute a rent for the dwelling based on market rents for comparable housing units) and the user cost approach (work out all of the anticipated or actual costs of owning the house for the reference period including depreciation and the opportunity cost of the capital tied up in owning the dwelling). However, these three approaches to the treatment of owner occupied housing will usually give quite different numerical results in the short run. Since all three approaches have strong support, it would be reasonable for a statistical agency to pick one approach for their flagship index but make available the other two treatments as "analytical series" for interested users. A third example where multiple indexes would be useful occurs in the context of seasonal commodities. The usual CPI is a month to month index and it is implicitly assumed that all commodities are available in each month. But this assumption is not warranted: in most countries, some 5 to 10 % of all commodities are generally not available in all months. In this context, a month to month CPI will not be as "accurate" as a year over year CPI that compares the prices of commodities in this month with the corresponding commodities in the same month a year ago. Hence again, the need for multiple indexes emerges to cater to the needs of different users.[22]

I may be wrong, but I think that there is an emerging consensus that it is permissible to have more than one price index where the different indexes might serve different purposes. I see this as a positive development.

---

[20] Edgeworth (1888) (1925) also was a firm believer in multiple indexes for multiple purposes and opposed the views of Walsh (1901) (1921) and Fisher (1922) who thought that once the transactions domain of definition for the index was chosen, then there was a single best way of constructing the index for all purposes.

[21] See Cage, Greenlees and Jackman (2003).

[22] Diewert (1999), Diewert, Alterman and Feenstra (2004) and Diewert, Finkel and Artsev (2004) advocated the construction of at least 3 indexes in the seasonal context.

2.5 Problems in Constructing Elementary Indexes have been Recognized

When price statisticians construct a component of a CPI or PPI, they do not use Laspeyres price indexes at the elementary (or first) stages of aggregation, because the Laspeyres index requires quantity or expenditure weights, which are generally not available. Hence, at the first stage of aggregation, the Carli (1764) (arithmetic average of price relatives), Jevons (1865) (geometric average of price relatives) or Dutot (1738) (arithmetic average of current period prices divided by arithmetic average of base period prices) indexes are used. The Carli has a definite upward bias but all three indexes suffer from being unweighted indexes. Until relatively recently, when scanner data has become more readily available, it was thought that the biases that might result from the use of unweighted indexes were not particularly significant but recent evidence points to *a very significant bias problem* at lower levels of aggregation compared to results that are generated by the preferred target indexes mentioned above (i.e., the Fisher, Walsh and Törnqvist Theil price indexes). In any case, the standard statistical agency practice at lower levels of aggregation is simply not consistent with the Laspeyres index as a target index (since the Laspeyres index requires proper weighting at all levels of aggregation). I think that until recently, the problems with the construction of price indexes at the elementary level of aggregation were not generally recognized as being as serious as they appear to be.

Until fairly recently, it was not possible to determine how close an unweighted elementary index of the type noted above (the Carli, Jevons and Dutot indexes) is compared to an elementary aggregate that was constructed using a weighted superlative formula. However, with the availability of *scanner data* (i.e., of detailed data on the prices and quantities of individual items that are sold in retail outlets), it has been possible to compute ideal elementary aggregates for some item strata and compare the results with statistical agency estimates of price change for the same class of items. Of course, the statistical agency estimates of price change are usually based on the use of the Dutot, Jevons or Carli formulae. The following quotations summarize many of these scanner data studies:

"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 percentage 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; 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 (2003)). Yet it is worth following up on Diewert's (2002a) 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, 2001a), 10% for washing machines (Table 7 below), 7.5% for dishwashers, 15% for cameras and 5% for vacuum cleaners (Silver and Heravi (2001b). 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 (2002a; 25).

The above quotations summarize the results of many elementary aggregate index number studies that are based on the use of scanner data.[23] 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 often below the corresponding official statistical agency estimates of price change for that category. Sometimes the measures of price change based on the use of scanner data are *considerably below* the corresponding official measures. These results indicate that there may be large gains in the precision of elementary indexes if a *weighted* sampling framework is adopted.

2.6 Towards a Better Understanding of a Typical CPI

At the final stages of aggregation, the standard CPI index is *not* a true Laspeyres index since the expenditure weights used at higher levels of aggregation pertain to a base *year* which is different from the base *month* (or quarter) for prices. Thus the expenditure weights are chosen at an annual frequency whereas the prices are collected at a monthly frequency. To be a true Laspeyres index, the base period expenditures should *coincide* with the base period for the prices. In fact, the actual target index used by many statistical agencies at the last stage of aggregation is typically[24] a *Young* (1812) *index*, which is equal to a share weighted average of the monthly price ratios of period t prices to period 0 prices but instead of using month 0 expenditure shares as weights, the Young index uses the expenditure shares of a (possibly distant) base year. Thus the Young index depends on the prices of the base month, the prices of the current month and the expenditure shares of a base year. It can be seen that the Young index is a weighted

[23] See also Silver and Heravi (2003).

[24] Some statistical agencies use the Lowe (1823) as their target index for a CPI. This index is similar to a Laspeyres index in that it is the ratio of a fixed basket of commodities priced out at the prices of the current month (say month t) divided by the same basket priced out at the prices of a base month (say month 0). However, instead of the quantity basket being the basket of purchases at month 0, the quantity basket is the basket of a base year (say year a), which usually precedes the base month 0. For an analysis of the substitution bias in this index, see Balk and Diewert (2004).

version of the unweighted Carli index.[25] Both the Young and Carli indexes have definite upward biases compared to theoretical target indexes.[26]

For the most part, measurement economists have not fully recognized that the typical CPI is based on a Young index instead of a Laspeyres index. However, the new ILO (2004) CPI Manual deals more fully with the complexities of the Young index. This is also an important new development in index number theory.

2.7 Gaps with Respect to Service Prices have been Recognized

A troublesome problem with the system of price statistics in most countries is that the problems of measuring complex services are generally neglected by statistical agencies due to difficulties in funding the collection of services prices. For the most part, the economics profession has not strongly lobbied for more price collection in the services area, so we bear some responsibility for this situation.[27] At present, a typical CPI program will collect many more goods prices than services prices and will have many more commodity classes for goods rather than services. In a way, this just reflects the historical origins of existing CPI theory. Until fairly recently, CPI theory has essentially remained unchanged for 80 years but 80 years ago, goods were much more significant than services, and hence, there was not much focus on the problems involved in measuring services. It is only over the last 30 or 40 years that the shift to services has caused service expenditures to exceed those on goods in many countries. However, if one looks at published CPI categories, there will generally be many more goods categories than services categories.[28] In addition to inertia, there are some serious conceptual problems involved in measuring the prices of many services. Some examples of difficult to measure services are: expenditures on insurance, gambling, financial services, advertising services, telecommunication services (with complex plans), entertainment services and rental housing. In many cases, statistical agencies simply do not have appropriate methodologies to deal with these difficult conceptual measurement problems and so in many cases, these service sector outputs are either not measured at all or deflated with a very rough and ready deflator. We will have more to say about these problems in section x below but in the present section, we want to stress that at least in Canada and the U.S., the data deficiencies in the service sectors is being addressed.[29] I view this as a very positive development.

2.8 The Price of Capital Services may Make an Appearance in the Next SNA

---

[25] The Carli index is simply an evenly weighted average of the price ratios. Hence, if the base year expenditure shares were all equal, the Young index would collapse down to a Carli index.

[26] The usual theoretical target indexes would be either the Fisher, Walsh or Törnqvist Theil index.

[27] A notable exception is Zvi Griliches (1992) (1994), who brought the data deficiency in services to the attention of economists.

[28] Detailed consumer price indexes for approximately 160 commodities are available from Statistics Canada on a monthly basis. Of these 160 consumer price indexes, only about 40 are devoted to service prices.

[29] The Bureau of Labor Statistics has greatly expanded its service prices program and it is likely that Statistics Canada will also expand its program.

The current System of National Accounts does not have a proper decomposition of sources of income into price and quantity components (as is the case with outputs and intermediate inputs).[30] However, statistical agencies in many countries[31] produce Multifactor Productivity or Total Factor Productivity estimates following along the lines originally suggested by Jorgenson and Griliches (1967). In order to construct these productivity estimates, it is necessary to construct prices for the services of reproducible capital inputs (machinery and equipment and structures) as well as for land and inventories. Thus government statisticians have gradually warmed to the idea of providing prices or user costs for various capital services[32] and so there is a chance that the next international agency version of the System of National Accounts will have prices for capital services imbedded in the System.[33] I view this as a very positive recent development.

2.9 The Contribution of Entering and Exiting Firms to Productivity Growth

Another relatively recent development in economic measurement theory is the recognition of the role that entering and disappearing production units might play in contributing to the productivity growth of a country or an industry. John Haltiwanger (1997) (2000) has been a pioneer in bringing this factor to the attention of the profession.[34] I believe that this is an exciting and relatively new area of research in index number theory.

2.10 Growth Decompositions

A final area of index number theory where progress has been made in the past 20 years is the development of index number decompositions into explanatory factors that are tied to economic theory. Decompositions of a quantity index number measure of overall growth into individual component sources of growth is not new[35] but what is new is

---

[30] SNA 1993 does recommend a decomposition of compensation of employees into wage and quantity components; see Eurostat (1993).

[31] The official measurement of TFP started in the U.S. with the BLS program (see the Bureau of Labor Statistics (1983) and Dean and Kunze (1992) for a description). Canada and Australia also now have extensive TFP programs.

[32] The excellent productivity measurement manual written by Schreyer (2001) also helped popularize the user cost idea among national statisticians.

[33] Some countries, such as the U.S., may implement prices for capital services in their own accounts in any case; see Jorgenson and Landefeld (2004) and the other papers presented at this NBER/CRIW Conference. The Canberra Group on Capital Measurement is currently having meetings on the usefulness of having capital service prices (or their major components) in the SNA.

[34] Some of the more important papers in this area are Baldwin and Gorecki (1991), Baily, Hulten and Campbell (1992), Griliches and Regev (1995), Baldwin (1995), Haltiwanger (1997), Ahn (2001), Foster, Haltiwanger and Krizan (2001), Aw, Chen and Roberts (2001), Bartelsman and Doms (2002), Fox (2002), Baldwin and Gu (2002) (2003), Balk (2003), and Bartelsman, Haltiwanger and Scarpetta (2004). Balk (2003; 29) emphasized the importance of a symmetric treatment of time. A symmetric decomposition was proposed earlier by Griliches and Regev (1995) and a modification of it was used by Aw, Chen and Roberts (2001). Diewert and Fox (2005) noted the relationship of this literature with the literature on making multilateral comparisons.

[35] An interesting decomposition for the Fisher quantity index was obtained by Van Ijzeren (1987) and it has been independently derived by Dikhanov (1997). The Van Ijzeren decomposition is currently being used

decompositions that have an explicit economic interpretation. Diewert and Morrison (1986) obtained this type of economic decomposition for the Törnqvist quantity index and the same decomposition was independently derived by Kohli (1990).[36] Diewert (2002c) obtained an analogous economic decomposition for the Fisher formula. The full potential of these decompositions has perhaps not been generally recognized by economists and statisticians but it is likely that it will be in coming years.

We turn now to a discussion of future challenges.

3. Future Challenges in Measuring Price Change

Many of the future challenges are follow up items from the progress list set out in the previous section. Thus I think that although we have made some progress in recognizing the problems that make price measurement difficult, our future challenges will be to fix these problems as best we can.

3.1 The Problem of Quantity or Expenditure Weights at the Elementary Level

As was mentioned in the previous section, there seem to be problems with statistical agency methodology in collecting prices at the elementary index level. I believe that the problems are mostly due to the fact that statistical agencies generally do not collect quantities transacted to go with the prices that are collected and so unreliable, unweighted price indexes must be used at the elementary level. The future challenge will be to convince agencies to collect quantities along with prices.

3.2 The Problems Associated with Adjusting Prices for Quality Change Need to be Solved

There are two main classes of methods that can be used to address the problem of adjusting prices for quality change:

- Hedonic regression techniques and
- Econometric estimation of reservation prices.

Hedonic regression techniques date back to Waugh (1929) and Court (1939),[37] while the reservation price methodology for dealing with the introduction of new goods dates back to Hicks (1940, 114). Hausman (1997) (1999) has implemented the Hicksian methodology but it has not been adopted by any statistical agency as of this date.

---

by Bureau of Economic Analysis; see Moulton and Seskin (1999; 16) and Ehemann, Katz and Moulton (2002). For more on this decomposition and its relationship to "economic" decompositions, see Reinsdorf, Diewert and Ehemann (2002).

[36] For closely related materials, see Morrison and Diewert (1990) and Fox and Kohli (1998). The graphical presentation of the Törnqvist decomposition that is in the latter paper is particularly useful.

[37] The manual on methods for quality adjustment by Triplett (2002) gives the most comprehensive review of hedonic regression methods.

Hedonic regression methods also have recently been reviewed in Chapter 4 of Schultze and Mackie (2002), where a rather cautious approach to the use of hedonic regressions was advocated due to the fact that many issues had not yet been completely resolved. A recent paper by Heravi and Silver (2002b) also raised questions about the usefulness of hedonic regressions since this paper presented several alternative hedonic regression methodologies and obtained different empirical results using the alternative models.[38]

Some of the more important issues that need to be resolved before hedonic regressions can be routinely applied by statistical agencies include:

- Should the dependent variable be transformed or not?
- Should separate hedonic regressions be run for each of the comparison periods or should we use the dummy variable adjacent year regression technique initially suggested by Court (1939; 109-11) and used by Berndt, Griliches and Rappaport (1995; 260) and many others?
- Should regression coefficients be sign restricted or not?
- Should the hedonic regressions be weighted or unweighted? If they should be weighted, should quantity or expenditure weights be used?
- How should outliers in the regressions be treated? Can influence analysis be used?

Thus there is a bit of work to be done before a consensus on "best practice" hedonic regression techniques emerges.

3.3 How Should Seasonal Prices be Treated?

The problem of seasonal prices was mentioned in section 2.4 above. I personally feel that the existence of seasonal prices (particularly prices that are available in some seasons but not in all seasons) means that one month to month price index will not serve all purposes. For forecasting short term changes in general inflation, a month to month index is required. But for longer term purposes, year over year indexes for each month will generally be more accurate and thus there is a need for at least two indexes. However, once year over year monthly indexes are available, it is useful to combine these indexes into a rolling year annual index that compares the prices of the past 12 months with the corresponding seasonal prices in a base year.[39] However, a "best practice" methodology for dealing with seasonal prices has not yet emerged and so this topic remains on the list of challenges for the future.

3.4 The Development of Measurement Methodologies for Difficult to Measure Service Sector Outputs

---

[38] The observation that different variants of hedonic regression techniques can generate quite different answers empirically dates back to Triplett and McDonald (1977; 150) at least.

[39] See Diewert (1999) or the ILO (2004) CPI Manual for the details.

In the North American Industrial Classification System, there are some 926 NAICS 6-digit industries. Of these, 381 are goods industries. The remaining 545 service sector industries break down as follows:

- Public administration (29 industries).
- Religious, grant-making, civic and professional service industries (10 industries).
- Education, health and social assistance industries (49 industries).
- Wholesale and retail trade (147 industries).
- Transportation (51 industries).
- Services 1 (Communication Services consisting of 37 industries), including postal and courier services, warehousing, periodicals and books, software publishers, movies, music, radio and television, telecommunications, news and data processing.
- Services 2 (Business Services consisting of 98 industries) including property leasing, real estate management, car and other rental and leasing, lawyers, accountants, architectural engineering, drafting, design and similar business services, computer services, administrative services, consulting and R&D services, advertising, photography, veterinary services, head office services, employment agencies, telephone call centers, collection agencies, travel agencies, security services, janitorial and cleaning services, and waste collection and disposal services.
- Services 3, (Personal Services consisting of 79 industries), including performing arts, professional sports, museums, parks, zoos, gambling, sports facilities, hotels and other accommodation, food services, drinking places, auto repair, car washes, equipment maintenance and repair, barber shops and beauty salons, funeral homes, laundries, pet care, photo finishing and parking lots.
- Finance and insurance, (45 industries), including the Bank of Canada, banking and related services, brokerages, exchanges, investment advice, accident, property and life insurance agencies, brokerages and carriers, pension funds and other financial services.

Some of the above industries have outputs that seem fairly straightforward to price. Statistics Canada has very rough and ready price indexes for the wholesale and retail trade industries (147 industries)[40] and more accurate price indexes for the 51 transportation industries.[41] Statistics Canada also has approximately 60 indexes from the Consumer Price Index that it uses to deflate the outputs of some of the remaining service sector industries. This leaves about 290 industries for which we have no deflator at present.

Now it is not Statistics Canada's fault that many service industries do not have specific price indexes. The problem is partly the fault of academics! *The outputs of many service sector industries are extremely difficult to measure in a manner that will command general acceptance*. Hence, this would seem to be a natural area for academics to enter

[4040] A detailed methodology for pricing the outputs and intermediate inputs of a distribution firm can be found in Diewert and Smith (1994).

[41] The methodology for measuring the prices of transportation outputs is generally well developed.

and develop methodologies for measuring these difficult to measure outputs. However, with a few exceptions, this has not happened. Hopefully, the Vancouver Conference will stimulate a wider interest in these very complex measurement problems.

We conclude this section by commenting on some general categories of difficult to measure service products (the categories overlap).

- *Unique products*. That is, in different periods, different products are produced. This prevents routine matching of prices. This is a pervasive problem in the measurement of the prices of services.
- *Complex products*. Many service products are very complicated; e.g., telephone service plans.
- *Tied products*. Many service products are bundled together and offered as a single unit; e.g., newspapers, cablevision plans, banking services packages. In principle, hedonic regression techniques could be used to price out these first three types of service products.
- *Joint products*. For this type of product, the value depends partially on the characteristics of the purchaser; e.g., the value of a year of education depends not only on the characteristics of the school and its teachers but also on the social and genetic characteristics of the student population.
- *Marketing and advertising products*. This class of service sector outputs is dedicated to influencing or informing consumers about their tastes. A standard economic paradigm for this type of product has not yet emerged.
- *Heavily subsidized products*. In the limit, subsidized products can be supplied to consumers free of (explicit) charges. Is zero the "right" price for this type of product?
- *Financial products*. What is the "correct" real price of a household's monetary deposits? Somewhat surprisingly, this question has not yet been resolved in a definitive manner.
- *Products involving risk and uncertainty*. What is the correct pricing concept for gambling and insurance expenditures? What is the correct price for a movie or a record original when it is initially released?

Hopefully, this conference will help stimulate some practical suggestions for price statisticians on how to measure these complex products.

3.5 The Transfer Price Problem

A big problem facing statistical agencies trying to construct import and export price indexes is: how to deal with transfer prices. A transfer price is a border price set by a multinational firm that trades products between subsidiaries in different countries. Current transfer price theory[42] suggests that multinationals have incentives to choose transfer prices strategically in order to minimize their global tax burdens. National tax authorities have rules to prevent multinationals from choosing total junk as their transfer prices but it is likely that currently reported transfer prices represent "economic" prices

[42] See Eden (1998) and Diewert (1985).

that reflect the resource costs of the exports or imports. Thus it becomes an increasingly difficult challenge for statistical agencies to produce price indexes for exports and imports that are meaningful, given that the proportion of international trade that is conducted between subsidiaries is about 50 per cent.[43]

3.6 What is the Right Household Price of Time?

There is increasing interest in measuring welfare in a more comprehensive manner; i.e., many researchers would like to value household uses of time that are used to produce services that are either substitutes for commodities that could be purchased on the market (e.g., meals or cleaning services) or that are used to produce final consumption services (e.g., reading a book or watching television). For the first type of activity, one could use either the household opportunity cost wage or the cost of hiring the service. A recent paper on this valuation issue describes the problem as follows:[44]

> "How to measure the value of unpaid time devoted to nonmarket production is the central input valuation issue. One possible approach is to value this time at the opportunity cost of the person performing the nonmarket activity. Another approach employed in the literature has been to value this time at market substitute prices—the wage that would be paid to a person hired to perform the task in question. The two approaches may give quite different answers if higher wage individuals devote time to tasks for which the market wage is relatively low." Katherine G. Abraham and Christopher Mackie (2004; 24).

There is a need to work out what is the "best practice" treatment of this time valuation problem—a challenge for the future.

In the following section, we review some of the challenges that are involved in measuring Total Factor Productivity Growth.

4. Measuring Productivity Growth

In this section, we go through the major classes of inputs used and outputs produced by a "typical" industry and note some of the measurement problems that face statistical agencies when they attempt to produce TFP growth rates for that industry.[45]

4.1 Gross Outputs

In order to measure the productivity of a firm, industry or economy, we need information on the outputs produced by the production unit for each time period in the sample along with the average price received by the production unit in each period for each of the outputs. In practice, period by period information on revenues received by the industry for a list of output categories is required along with either an output index or a price

[43] A recent paper that looks at some of the practical measurement issues involving transfer prices is Diewert, Alterman and Eden (2004)
[44] Another recent paper that discusses these household valuation of time issues is Nordhaus (2004).
[45] See Balk (2003) and Diewert and Nakamura (2003) for recent surveys on the measurement of Total Factor Productivity. The modern literature on this topic starts with Jorgenson and Griliches (1967). This section draws heavily on Diewert (2001b).

index for each output. In principle, the revenues received should not include any commodity taxes imposed on the industry's outputs, since producers in the industry do not receive these tax revenues.[46] The above sentences sound very straightforward but many firms produce thousands of commodities so the aggregation difficulties are formidable. Moreover, as we have noted earlier, many outputs in service sector industries are difficult to measure conceptually: think of the proliferation of telephone service plans and the difficulties involved in measuring insurance, gambling, banking and options trading.

4.2 Intermediate Inputs

Again, in principle, we require information on all the intermediate inputs utilised by the production unit for each time period in the sample along with the average price paid for each of the inputs. In practice, period by period information on costs paid by the industry for a list of intermediate input categories is required along with either an intermediate input quantity index or a price index for each category. In principle, the intermediate input costs paid should include any commodity taxes imposed on the intermediate inputs, since these tax costs are actually paid by producers in the industry.

The major classes of intermediate inputs at the industry level are:

- materials
- business services
- leased capital.

The current input–output framework deals reasonably well in theory with the flows of materials but not with intersectoral flows of contracted labour services or rented capital equipment. The input-output system was designed long ago when the leasing of capital was not common and when firms had their own in house business services providers. Thus there is little or no provision for business service and leased capital intermediate inputs in the present system of accounts. With the exception of the manufacturing sector, even the intersectoral value flows of materials are largely incomplete in the industry statistics. There is also the problem of a lack of price surveys for intermediate inputs: typically, a price index for an intermediate input class into some industry is approximated by an output price index, which may have a different mix of detailed products in it.

4.3 Labor Inputs

Using the number of employees as a measure of labour input into an industry will not usually be a very accurate measure of labour input due to the possibility of changes in average hours worked per full time worker and possible changes in the use of part time workers. However, even total hours worked in an industry is not a satisfactory measure of labour input if the industry employs a mix of skilled and unskilled workers. Hours of

[46] A recent paper that discusses the treatment of indirect taxes in the context of productivity measurement in more detail is Diewert (2005).

work contributed by highly skilled workers generally contribute more to production than hours contributed by very unskilled workers. Hence, it is best to decompose aggregate labour compensation into its aggregate price and quantity components using index number theory.[47] The practical problem faced by statistical agencies is: how should the various categories of labour be defined? Dean and Harper (2001) provide an accessible summary of the literature in this area.

Another important problem associated with measuring real labour input is finding an appropriate allocation of the operating surplus of proprietors and the self employed into labour and capital components. There are two broad approaches to this problem:

- If demographic information on the self employed is available along with hours worked, then an imputed wage can be assigned to those hours worked based on the average wage earned by employees of similar skills and training. Then an imputed wage bill can be constructed and subtracted from the operating surplus of the self employed. The reduced amount of operating surplus can then be assigned to capital.
- If information on the capital stocks utilised by the self employed is available, then these capital stocks can be assigned user costs and then an aggregate imputed rental can be subtracted from operating surplus. The reduced amount of operating surplus can then be assigned to labour. These imputed labour earnings can then be divided by hours worked by proprietors to obtain an imputed wage rate.

The problems posed by allocating the operating surplus of the self employed are becoming increasingly more important as this type of employment grows. Fundamentally, the problem appears to be that the current System of National Accounts (SNA) does not address this problem adequately.

4.4 Reproducible Capital Inputs

When a firm purchases a durable capital input, it is not appropriate to allocate the entire purchase price as a cost to the initial period when the asset was purchased. It is necessary to distribute this initial purchase cost across the useful life of the asset. National income accountants recognize this and use depreciation accounts to do this distribution of the initial cost over the life of the asset. However, historically, national income accountants are reluctant to recognize the interest tied up in the purchase of the asset as a true economic cost. Rather, they tend to regard interest as a transfer payment. Thus the user cost of an asset (which recognizes the opportunity cost of capital as a valid economic cost) is not regarded as a valid approach to valuing the services provided by a durable capital input by many national income accountants. However, if a firm buys a durable capital input and leases or rents it to another sector, national income accountants regard the induced rental as a legitimate cost for the using industry. It seems very unlikely that the leasing price does not include an allowance for the capital tied up by the initial purchase of the asset; i.e., market rental prices include interest. Hence, it seems reasonable to include an imputed interest cost in the user cost of capital even when the

[47] The current SNA does recommend a wage index to deflate employee compensation; see Eurostat (1993).

asset is not leased. Put another way, interest is still not accepted as a cost of production in the SNA, since it is regarded as an unproductive transfer payment. But interest is productive; it is the cost of inducing savers to forego immediate consumption.

But, as mentioned in section 2, there is some hope that interest will be regarded as a cost of production in the next version of the System of National Accounts.

The treatment of capital gains on assets is even more controversial than the national accounts treatment of interest. In the national accounts, capital gains are not accepted as being productive[48] but if resources are transferred from a period where they are less valuable to a period where they are more highly valued, then a gain has occurred; i.e., capital gains are productive according to this view.

However, the treatment of interest and capital gains pose practical problems for statistical agencies. For example, which interest rate should be used?

- An ex post economy wide rate of return which is the alternative used by Christensen and Jorgenson (1969) (1970)?
- An ex post firm or sectoral rate of return? This method seems appropriate from the viewpoint of measuring ex post performance.
- An ex ante safe rate of return like a Federal Government one year bond rate? This method seems appropriate from the viewpoint of constructing ex ante user costs that could be used in econometric models.
- Or should the ex ante safe rate be adjusted for the risk of the firm or industry?

Since the ex ante user cost concept is not observable, the statistical agency will have to make somewhat arbitrary decisions in order to construct expected capital gains.[49] This is a strong disadvantage of the ex ante concept. On the other hand, the use of the ex post concept will lead to rather large fluctuations in user costs, which in some cases will lead to negative user costs, which in turn may be hard to explain to users. However, a negative user cost simply indicates that instead of the asset declining in value over the period of use, it rose in value to a sufficient extent to offset deterioration. Hence, instead of the asset being an input cost to the economy during the period, it becomes an intertemporal output. This makes sense from the ex post point of view but not from the ex ante point of view. Thus it is necessary to be clear whether an ex ante or ex post user cost should appear in the production accounts.[50]

A further complication is that our empirical information on the actual efficiency decline of assets is weak. We do not have good information on the useful lives of assets. The UK statistician assumes machinery and equipment in manufacturing lasts on average 26 years

[48] Capital gains do appear in the SNA 1993 in the Revaluation Accounts.

[49] Alternatively, the statistical agency could combine the nominal interest rate term in the user cost formula with the expected capital (or holding) gains term and simply assume a real interest rate.

[50] My own preference would be for the ex ante user cost concept to be used in the production accounts. Actual capital gains on assets should appear in the balance sheet accounts.

while the Japanese statistician assumes machinery and equipment in manufacturing lasts on average 11 years; see the OECD (1993; 13). The problems involved in measuring capital input are also being addressed by the Canberra Group on Capital Measurement, which is an informal working group of international statisticians dedicated to resolving some of these measurement problems.

A final set of problems associated with the construction of user costs is the treatment of business income taxes: should we assume firms are as clever as Hall and Jorgenson (1967) and can work out their rather complex tax–adjusted user costs of capital or should we go to the accounting literature and allocate capital taxes in the rather unsophisticated ways that are suggested there?

Our conclusion at this point is that there are significant problems to be addressed before user costs of capital can be introduced into the production accounts of the System of National Accounts.

4.5 Inventories

Because interest is not a cost of production in the national accounts and the depreciation rate for inventories is close to zero, many productivity studies neglect the user cost of inventories. This leads to misleading productivity statistics for industries where inventories are large relative to output, such as retailing and wholesaling. In particular, rates of return that are computed neglecting inventories will be too high since the opportunity cost of capital that is tied up in holding the beginning of the period stocks of inventories is neglected.

The problems involved in accounting for inventories are complicated by the way accountants and the tax authorities treat inventories. These accounting treatments of inventories are problematic in periods of high or moderate inflation. A treatment of inventories that is suitable for productivity measurement can be found in Diewert and Smith (1994). These inventory accounting problems seem to carry over to the national accounts in that for virtually all OECD countries, there are time periods where the real change in inventories has the opposite sign to the corresponding nominal change in inventories. This seems logically inconsistent. This is another area for future research.

4.6 Land

The production accounts in the current SNA has no role for land as a factor of production, perhaps because it is thought that the quantity of land in use remains roughly constant across time and hence it can be treated as a fixed, unchanging factor in the analysis of production. However, the quantity of land in use by any particular firm or industry does change over time. Moreover, the price of land can change dramatically over time and thus the user cost of land will also change over time and this changing user cost will, in general, affect correctly measured productivity.

Land ties up capital just like inventories (both are zero depreciation assets). Hence, when computing ex post rates of return earned by a production unit, it is important to account for the opportunity cost of capital tied up in land. Neglect of this factor can lead to biased rates of return on financial capital employed. Thus, industry rates of return and TFP estimates may not be accurate for sectors like agriculture which are land intensive.

Finally, property taxes that fall on land must be included as part of the user cost of land. In general, it may not be easy to separate the land part of property taxes from the structures part. In the national accounts, property taxes (which are input taxes) are lumped together with other indirect taxes that fall on outputs which is another shortcoming of the current SNA.

It should be noted that the TFP programs in the U.S., Canada and Australia all recognize the need to have land and inventories in their measures as inputs and they work out user costs for these inputs.

4.7 Resources

Examples of resource inputs include:

- depletion of fishing stocks, forests, mines and oil wells
- improvement of air, land or water environmental quality (these are resource "outputs" if improvements have taken place and are resource "inputs" if degradation has occurred).

The correct prices for resource depletion inputs are the gross rents (including resource taxes) that these factors of production earn. Resource rents are usually not linked up with the depletion of resource stocks in the national accounts although some countries, including the U.S. and Canada, are developing statistics for forest, mining and oil depletion; see Nordhaus and Kokkelenberg (1999).

The pricing of environmental inputs or outputs is much more difficult. From the viewpoint of traditional productivity analysis based on shifts in the production function, the 'correct' environmental quality prices are marginal rates of transformation while, from a consumer welfare point of view, the 'correct' prices are marginal rates of substitution; see Gollop and Swinand (2001). Needless to say, there are many difficulties involved in estimating these environmental prices.

The above seven major classes of inputs and outputs represent a minimal classification scheme for organizing information to measure TFP at the sectoral level. Unfortunately, no country has yet been able to provide satisfactory price and quantity information on all seven of these classes. To fill in the data gaps, it would be necessary for governments to expand the budget of the relevant statistical agencies considerably. This is one area of government expenditure that cannot be readily filled by the private sector. Given the importance of productivity improvements in improving standards of living, the accurate measurement of productivity seems necessary.

4.8 Other Statistical Agency Challenges in Measuring Productivity

One problem with current estimates of productivity growth is that every statistical agency uses *different* surveys to collect to collect information on the outputs of an industry and on the various input components. Furthermore, every statistical agency uses *different* surveys to collect information on prices and values.

These separate data collection surveys do not greatly impede the construction of reasonably accurate price and quantity aggregates for the components of final demand for the economy as a whole but they lead to extremely inaccurate estimates of prices and quantities for industries or smaller units such as firms or establishments. In particular, the firm or industry specific price indexes that are applied to the firm's or industry's value components (such as output, intermediate input, labor input, etc.) will typically be very inaccurate.

Many firms have taken advantage of the low cost of computing and have detailed data on all of their financial transactions (e.g., they have the value of each sale and the quantity sold by commodity). This opens up the possibility of the statistical agency replacing or supplementing their surveys on say, prices of outputs, by the electronic submission by firms to the statistical agency of their computerized transaction histories for a certain number of periods. This information would provide the industry/firm counterparts to the scanner data studies that have proved to be so useful in the context of the Consumer Price Index. This information would also lead to true microeconomic price and quantity indexes at the firm level and to accurate firm and industry productivity indexes.

We conclude our discussion of productivity measurement challenges by discussing another topic that has not received the attention it deserves.

4.9 System Wide Versus Sectoral Productivity Measurement

Individual firms or establishments could be operating efficiently (i.e., be on the frontiers of their production possibilities sets) but yet, the economy as a whole may not be operating efficiently. How can this be? The explanation for this phenomenon was given by Gerard Debreu (1951)[51]: there is a loss of system wide output (or *waste*, to use Debreu's term) due to the imperfection of economic organization; i.e., different production units, while technically efficient, face different prices for the same input or output, and this causes net outputs aggregated across production units to fall below what is attainable if the economic system as a whole, were efficient. In other words, a condition for system wide efficiency is that all production units face the same price for each separate input or output that is produced by the economy as a whole. Thus, if producers face different prices for the same commodity and if production functions

[51] Debreu (1951; 285) distinguished two other sources of waste in the allocation of resources: (a) waste due to the underemployment of available physical resources (e.g., unemployed workers) and (b) waste due to technical inefficiency in production. Obviously, the application of knowledge capital could be useful in diminishing waste (b). Waste (a) results from market imperfections between the aggregate production sector and the household sector of the economy.

exhibit some substitutability, then producers will be induced to jointly supply an inefficient economy wide net output vector. What are sources of system wide waste?

- industry specific taxes or subsidies that create differences in prices faced by production units for the same commodity; e.g., an industry specific subsidy for an output or a tax on the output of one industry where that output is used as an input by other industries (an example of the latter is a gasoline tax);
- tariffs on imports or subsidies or taxes on exports;
- union induced wage differentials across firms for the same type of labor service;
- monopolistic or monopsonistic markups on commodities by firms or any kind of price discrimination on the part of firms;
- imperfect regulation;
- a source of commodity price wedges that is related to the last source above is the difficulty that multiproduct firms have in pricing their outputs, particularly when there are large fixed costs involved in producing new (or old) products.

Diewert (1983) adapted Debreu's (1951) methodology and provided measures of the loss of efficiency due to the fact that different sectors in the economy may face different prices for the same input or output. However, Diewert used international prices as reference prices in his calculations and it is not clear that this choice of reference prices is the "right" one. Thus there is room for future research in this area.[52]

5. Conclusion

I have not covered all of the recent developments in index number theory and in the measurement of productivity but I hope that I have indicated that there has been a great deal of relevant research in the past 30 years.

During the past 30 years, there have been at least three groups who have kept up a focus on measurement problems:

- Wolfgang Eichhorn's series of meetings in Karlsruhe during the 1970's and 1980's that focused on index number theory;
- The Productivity Program in the National Bureau of Economic Research,[53] which started in the late 1970's under the direction of Zvi Griliches and is now headed up by Ernst Berndt and
- The International Working Group on Price Indices (the Ottawa Group), which had its first meeting in Ottawa in 1994.[54]

The Ottawa Group consists mostly of price statisticians and other researchers working in statistical agencies around the world. This group was formed so that "best practice"

[52] It is of some interest that macroeconomists are becoming interested in this aspect of productivity performance; see Basu and Fernald (2002).
[53] A closely related NBER group is the Conference for Research in Income and Wealth, which has been headed up by Charles Hulten for the past two decades. I should also mention the International CRIW group which has also kept measurement issues alive over the past 3 decades.
[54] The Ottawa Group was started by Bert Balk, Bohdan Schultz and Jacob Ryten.

developments in the measurement of prices could be more readily transmitted across countries. Another improvement that I have noticed in the last decade is that the old gulf between academics interested in measurement problems and "practical" statisticians has narrowed to the point where there is no obvious difference between the two groups. I regard this as a very positive development.

Another very positive development has been the production of the new ILO (2004) CPI and IMF (2004) PPI Manuals. These Manuals fill a gap in that until these Manuals were produced, there was no comprehensive single source of materials on index number theory and the measurement problems that are associated with the production of a CPI and PPI.[55]

A not so positive development that has occurred in the past 30 years is that national income accounting and the associated measurement problems are no longer taught in graduate macroeconomics courses. However, I believe that academic interest in measurement problems is increasing and so perhaps in coming years, index number theory[56] and the associated measurement problems will creep back into graduate economics programs.

References

Abraham, K.G. and C. Mackie (2004), "A Framework for Nonmarket Accounting", paper presented at the CRIW Conference, New Architecture of the U.S. National Accounts, Washington D.C., April 16-17, 2004.

Aw, B.Y., X. Chen and M.J. Roberts (2001), "Firm Level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing", *Journal of Development Economics* 66, 51-86.

Ahn, S. (2001), "Firm Dynamics and Productivity Growth: A Review of Micro Evidence from OECD Countries", OECD Economics Department Working Paper No. 297, OECD, Paris.

Baily, M.N., C. Hulten, D. Campbell, (1992), "Productivity Dynamics in Manufacturing Establishments", *Brookings Papers on Economic Activity: Microeconomics 1992*, 187-249.

Baldwin, J.R. (1995), *The Dynamics of Industrial Competition: A North American Perspective*, Cambridge: Cambridge University Press.

[55] Hill (1993) provided an excellent exposition of index number theory in SNA 1993, but he was not able to go into the level of detail that is present in the new Manuals and there were no references to the literature in his exposition.

[56] A good sign of increased general interest in measurement problems on the part of academics is the appendix on price and productivity indexes in a graduate textbook on international trade; see Feenstra (2004; 410-428.

Baldwin, J.R. and P.K. Gorecki (1991), "Entry, Exit, and Productivity Growth", in: P.A. Geroski and J. Schwalbach (eds.), *Entry and Market Contestability: An International Comparison*, Oxford: Blackwell.

Baldwin, J.R. and W. Gu (2002), "Plant Turnover and Productivity Growth in Canadian Manufacturing", OECD STI Working Paper 2002/1, Paris: OECD. Available at www.statcan.ca/english/studies/prod.htm

Baldwin, J. and W. Gu (2003), "Industrial Competition, Shifts in Market Share and Productivity Growth", Working Paper, Micro Economic Analysis Division, Statistics Canada, Ottawa.

Balk, B.M. (1995), "Axiomatic Price Index Theory: A Survey", *International Statistical Review* 63, 69-93.

Balk, B.M. (2000), "Divisia Price and Quantity Indexes 75 Years After", Department of Statistical Methods, Statistics Netherlands, P.O. Box 4000, 2270 JM Voorburg, The Netherlands.

Balk, B.M. (2003), "The Residual: On Monitoring and Benchmarking Firms, Industries and Economies with Respect to Productivity", *Journal of Productivity Analysis* 20, 5-74.

Balk, B.M. and W.E. Diewert (2004), "The Lowe Consumer Price Index and Its Substitution Bias", Discussion Paper 04-07, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Bartelsman, E.J. and M. Doms (2000), "Understanding Productivity: Lessons from Longitudinal Microdata", *Journal of Economic Literature* 38, 569-595.

Bartelsman, E., J. Haltiwanger and J. Scarpetta (2004), "Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries", unpublished manuscript.

Basu, S. and G.J. Fernald (2002), "Aggregate Productivity and Aggregate Technology", *European Economic Review* 46, 963-991.

Berndt, E. R., Z. Griliches and N. J. Rappaport (1995), "Econometric Estimates of Price Indexes for Personal Computers in the 1990's", *Journal of Econometrics* 68, 243-268.

Boskin, M.J. (Chair), E.R. Dullberger, R.J. Gordon, Z. Griliches and D.W. Jorgenson (1996), *Final Report of the Commission to Study the Consumer Price Index*, U.S. Senate, Committee on Finance, Washington DC: US Government Printing Office.

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, Voorburg, April 16-18.

Bureau of Labor Statistics (1983), *Trends in Multifactor Productivity*, 1948-81, Bulletin 2178, U.S. Government Printing Office, Washington, D.C.

Cage, R., J.S. Greenlees, and P. Jackman (2003): "Introducing the Chained CPI," paper presented at the Seventh Meeting of the International Working Group on Price Indices, (Ottawa Group), Paris, France, May 27.

Carli, Gian-Rinaldo, (1804), "Del valore e della proporzione de' metalli monetati", pp. 297-366 in *Scrittori classici italiani di economia politica*, Volume 13, Milano: G.G. Destefanis (originally published in 1764).

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): 293-320.

Christensen, L.R. and D.W. Jorgenson (1970), "U.S. Real Product and Real Factor Input, 1929-1967," *Review of Income and Wealth* 16(1): 19-50.

Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 98-117 in *The Dynamics of Automobile Demand*, New York: General Motors Corporation.

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

Dean, E. and M. Harper (2001), "The BLS Productivity Measurement Program", pp. 55-84 in *New Developments in Productivity Analysis*, C.R. Hulten, E.R. Dean and M.J. Harper (eds.), Chicago: University of Chicago Press.

Dean, E.R. and K. Kunze (1992), "Productivity Measurement in the Service Industries", pp. 73-101 in *Output Measurement in the Service Sectors*, Studies in Income and Wealth 56, Chicago: University of Chicago Press.

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, Statistics Netherlands, Voorburg, April 16-18.

Debreu, G. (1951), "The Coefficient of Resource Utilization", *Econometrica* 19, 273-292.

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

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

Diewert, W.E. (1980), "Aggregation Problems in the Measurement of Capital", pp. 433-528 in *The Measurement of Capital*, Dan Usher (ed.), University of Chicago Press, Chicago.

Diewert, W.E. (1983), "The Measurement of Waste Within the Production Sector of an Open Economy", *Scandinavian Journal of Economics* 85:2, 159-179.

Diewert, W. E. (1985), "Transfer Pricing and Economic Efficiency", pp. 47-81 in *Multinationals and Transfer Pricing*, A. M. Rugman and L. Eden (eds.), London: Croom Helm.

Diewert, W.E. (1997), "Commentary on Mathew D. Shapiro and David W. Wilcox: Alternative Strategies for Aggregating Prices in the CPI", *The Federal Reserve Bank of St. Louis Review*, Vol. 79:3, 127-138.

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

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

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

Diewert, W. E. (2001b), "Which (Old) Ideas on Productivity Measurement are Ready to Use?" pp. 85-101 in *New Developments in Productivity Analysis*, C.R. Hulten, E.R. Dean and M.J. Harper (eds.), Chicago: University of Chicago Press.

Diewert, W. E. (2001c), "The Consumer Price Index and Index Number Purpose", *Journal of Economic and Social Measurement* 27, 167-248.

Diewert, W.E. (2002a), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations", *Swiss Journal of Economics and Statistics*, 138, 547-637.

Diewert, W.E. (2002b), "Similarity and Dissimilarity Indexes: An Axiomatic Approach", Discussion Paper 02-10, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E. (2002c), "The Quadratic Approximation Lemma and Decompositions of Superlative Indexes", *Journal of Economic and Social Measurement* 28, 63-88.

Diewert, W.E. (2003), "Services and the New Economy: Data Needs and Challenges", Discussion Paper No. 03-09, Department of Economics, University of British Columbia, Vancouver Canada. http://www.econ.ubc.ca/diewert/disc.htm#2003

Diewert, W.E. (2005), "The Treatment of Indirect Taxes and Margins and the Reconciliation of Industry with National Productivity Measures", Discussion Paper 05-06, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E., W.F. Alterman and L. Eden (2004), "Transfer Prices and Import and Export Price Indexes: Theory and Practice", paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, June 30-July 3, 2004.

Diewert, W.E., W.F. Alterman and R.C. Feenstra (2004), "Time Series versus Index Number Methods of Seasonal Adjustment", Discussion Paper 04-06, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E., Y. Finkel and Y. Artsev (2004), "On the Treatment of Seasonal Commodities in the CPI: the Israeli Experience", paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, June 30-July 3, 2004.

Diewert, W.E. and K.J. Fox (2005), "On Measuring the Contribution of Entering and Exiting Firms to Aggregate Productivity Growth", Discussion Paper 05-02, Department of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1.

Diewert, W.E. and C.J. Morrison (1986), "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade", *The Economic Journal* 96, 659-679.

Diewert, W.E. and A.O. Nakamura (2003), "Index Number Concepts, Measures and Decompositions of Productivity Growth", *Journal of Productivity Analysis* 19, 127-159.

Diewert, W.E. and A.M. Smith (1994), "Productivity Measurement for a Distribution Firm", *The Journal of Productivity Analysis* 5, 335-347.

Dikhanov, Y. (1997), "The Sensitivity of PPP-Based Income Estimates to Choice of Aggregation Procedures", mimeo, International Economics Department, The World Bank, Washington D.C., January.

Divisia, F. (1926), *L'indice monetaire et la theorie de la monnaie*, Paris: Societe anonyme du Recueil Sirey.

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

Eden, L. (1998), *Taxing Multinationals: Transfer Pricing and Corporate Income Taxation in North America*, Toronto: University of Toronto Press.

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

Edgeworth, F.Y. (1925), "The Plurality of Index Numbers", *The Economic Journal* 35, 379-388.

Ehemann, C., A.J. Katz and B.R. Moulton (2002), "The Chain-Additivity Issue and the U.S. National Accounts", *Journal of Economic and Social Measurement* 28, 37-49.

Eichhorn, W. (1978), *Functional Equations in Economics*, Reading, MA: Addison-Wesley Publishing Company.

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

Eurostat (1993), *System of National Accounts 1993*, Eurostat, IMF, OECD, UN and World Bank, Luxembourg, Washington, D.C., Paris, New York, and Washington, D.C.

Feenstra, R.C. (2004), *Advanced International Trade: Theory and Practice*, Princeton: Princeton University Press.

Fenwick, D., A. Ball, M. Silver and P. H. Morgan (2003), "Price Collection and Quality Assurance of Item Sampling in the Retail Price Index: How Can Scanner Data Help?", pp. 67-87 in *Scanner Data and Price Indexes*, Mathew Shapiro and Rob Feenstra (eds.), National Bureau of Economic Research, Studies in Income and Wealth, Volume 61, Chicago: University of Chicago Press.

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

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

Foster, L., J. Haltiwanger and C.J. Krizan (2001), "Aggregate Productivity Growth: Lessons from Microeconomic Evidence", pp. 303-372 in *New Developments in Productivity Analysis*, C.R. Hulten, E.R. Dean and M.J. Harper (eds.), NBER Studies in Income and Wealth Volume 63, Chicago: University of Chicago Press.

Fox, K.J. (2003), "Problems with (Dis) Aggregating Productivity and Another Productivity Paradox", Discussion Paper 2003/1, School of Economics, The University of New South Wales, Australia.

Fox, K.J. and U. Kohli (1998), "GDP Growth, Terms of Trade Effects and Total Factor Productivity", *The Journal of International Trade and Economic Development* 7:1, 87-110.

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

Gollop, F.M. and G.P. Swinand (2001), "Total Resource Productivity: Accounting for Changing Environmental Quality", pp. 587-605 in *New Developments in Productivity Analysis*, C.R. Hulten, E.R. Dean and M.J. Harper (eds.), Chicago: University of Chicago Press.

Griliches, Z. (1971a), "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change", pp. 55-87 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

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

Griliches, Z. (1992), "Introduction", pp. 1-22 in *Output Measurement in the Service Sectors*, Studies in Income and Wealth 56, Chicago: University of Chicago Press.

Griliches, Z. (1994), "Productivity, R&D and the Data Constraint", *American Economic Review* 84, 1-20.

Griliches, Z. and H. Regev (1995), "Firm Productivity in Israeli Industry: 1979-1988", *Journal of Econometrics* 65, 175-203.

Hall, R.E. and D.W. Jorgenson, (1967), "Tax Policy and Investment Behavior", *American Economic Review* 57, 391-414.

Haltiwanger, J. (1997), "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence", *Federal Reserve Bank of St. Louis Economic Review* 79:3, 55-77.

Haltiwanger, J. (2000), "Aggregate Growth: What have we Learned from Microeconomic Evidence?", Working Paper No. 267, Economics Department, OECD, Paris.

Hausman, J.A. (1997):, "Valuation of New Goods Under Perfect and Imperfect Competition," pp. 209-237 in Bresnahan and Gordon (eds.), *The Economics of*

New Goods, NBER Studies in Income and Wealth 58, Chicago: University of Chicago Press.

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

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, Statistics Netherlands, Voorburg, April 16-18.

Hicks, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, 105-140.

Hill, R.J. (1995), Purchasing Power Methods of Making International Comparisons, Ph. D. dissertation, Vancouver: The University of British Columbia.

Hill, R.J. (1999a), "Comparing Price Levels across Countries Using Minimum Spanning Trees", *The Review of Economics and Statistics* 81, 135-142.

Hill, R.J. (1999b), "International Comparisons using Spanning Trees", pp. 109-120 in *International and Interarea Comparisons of Income, Output and Prices*, A. Heston and R.E. Lipsey (eds.), Studies in Income and Wealth Volume 61, NBER, Chicago: The University of Chicago Press.

Hill, R.J. (2001), "Measuring Inflation and Growth Using Spanning Trees", International Economic Review 42, 167-185.

Hill, R.J. (2004), "Constructing Price Indexes Across Space and Time: The Case of the European Union", *American Economic Review*, forthcoming.

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

Hill, T.P. (1993), "Price and Volume Measures", pp. 379-406 (Chapter 16) in *System of National Accounts 1993*, Eurostat, IMF, OECD, UN and World Bank, Luxembourg, Washington, D.C., Paris, New York, and Washington, D.C.

International Labour Organization (2004), *Consumer Price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: The ILO, forthcoming.

International Monetary Fund (2004), *Producer Price Index Manual: Theory and Practice*, Paul Armknecht (ed.), Washington: IMF, forthcoming.

Jevons, W.S., (1865), "The Variation of Prices and the Value of the Currency since 1782", *Journal of the Statistical Society of London* 28, 294-320; reprinted in

Investigations in Currency and Finance (1884), London: Macmillan and Co., 119-150.

Jorgenson, D.W. and Z. Griliches (1967), "The Explanation of Productivity Change", *Review of Economic Studies* 34, 249-283.

Jorgenson, D.W. and S. Landefeld (2004), "Blueprint for an Expanded and Integrated Set of Accounts for the United States", paper presented at the CRIW Conference, New Architecture of the U.S. National Accounts, Washington D.C., April 16-17, 2004.

Kohli, U. (1990), "Growth Accounting in the Open Economy: Parametric and Nonparametric Estimates", *Journal of Economic and Social Measurement* 16, 125-136.

Lehr, J. (1885), *Beitrage zur Statistik der Preise*, Frankfurt: J.D. Sauerlander.

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

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

Morrison, C.J. and W.E. Diewert (1990), "Productivity Growth and Changes in the Terms of Trade in Japan and the United States", pp. 201-227 in *Productivity Growth in Japan and the United States*, C.R. Hulten (eds.), NBER Studies in Income and Wealth Vol. 53, Chicago: The University of Chicago Press.

Moulton, B.R., and E.P. Seskin (1999), "A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts", *Survey of Current Business* 79 (October), 6-17.

Nordhaus, W.D. (2004), "Principles of National Accounting for Non-Market Accounts", paper presented at the CRIW Conference, New Architecture of the U.S. National Accounts, Washington D.C., April 16-17, 2004.

Nordhaus, W.D. and E.D. Kokkelenberg (1999), *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment*, Washington D.C.: National Academy Press.

OECD (1993), *Methods Used by OECD Countries to Measure Stocks of Fixed Capital, National Accounts: Sources and Methods No. 2*, Publications Service, Paris France: OECD.

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, Washington D.C., April.

Reinsdorf, M., W.E. Diewert and C. Ehemann (2002), "Additive Decompositions for Fisher, Törnqvist and Geometric Mean Indexes", *Journal of Economic and Social Measurement* 28, 51-61.

Schreyer, P. (2001), *OECD Productivity Manual: A Guide to the Measurement of Industry-Level and Aggregate Productivity Growth*, Paris: OECD.

Schultze, C.L. and C. Mackie (eds.) (2002), *At What Price? Conceptualizing and Measuring Cost-of-Living Indexes*, Washington D. C.: National Academy Press.

Shiratsuka, S. (1995a), "Effects of Quality Changes on the Price Index: A Hedonic Approach to the Estimation of a Quality Adjusted Price Index for Personal Computers in Japan", *Monetary and Economic Studies*, Bank of Japan, 13:1, 17-52.

Shiratsuka, S. (1995b), "Automobile Prices and Quality Changes: A Hedonic Price Analysis of the Japanese Automobile Market", *Monetary and Economic Studies*, Bank of Japan, 13:2, 1-44.

Shiratsuka, S. (1999a), "Measurement Errors and Quality Adjustment Methodology: Lessons from the Japanese CPI", *Economic Perspectives*, Federal Reserve Bank of Chicago, 23:2, 2-13.

Shiratsuka, S. (1999b), "Measurement Errors in the Japanese Consumer Price Index", *Monetary and Economic Studies*, Bank of Japan, 17:3, 69-102.

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

Silver, M.S. and Heravi S. (2001a), "Scanner Data and the Measurement of Inflation", *The Economic Journal,* 111 June, F384-F405

Silver, M.S. and Heravi S. (2001b), "Hedonic Price Indices and the Matched Models Approach", Mimeo, Cardiff Business School, Cardiff, U.K.

Silver, M.S. and S. Heravi (2002a), "Why the CPI Matched Models Method May Fail Us", European Central Bank Working Paper 144, Frankfurt.

Silver, M.S. and S. Heravi (2002b), "On the Stability of Hedonic Coefficients and their Implications for Quality Adjusted Price Change Measurement", paper presented at the Third Annual Joint NBER and CRIW Measurement Workshop, Cambridge MA, July 29-31, 2002.

Silver, M. and S. Heravi (2003), "The Measurement of Quality Adjusted Price Changes", pp. 277-316 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Stigler, G. (ed.) (1961), *The Price Statistics of the Federal Government*, New York: National Bureau of Economic Research.

Szulc, B.J. (1983), "Linking Price Index Numbers," pp. 537-566 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

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

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

Triplett, J.E. (1983), "Escalation Measures: What is the Answer? What is the Question?", pp. 457-482 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Triplett J.E. (2004), *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products*, OECD Directorate for Science, Technology and Industry, Paris: OECD.

Triplett, J. E. and R. J. McDonald (1977), "Assessing the Quality Error in Output Measures: The Case of Refrigerators", *The Review of Income and Wealth* 23:2, 137-156.

Van Ijzeren (1987), *Bias in International Index Numbers: A Mathematical Elucidation*, Dissertation for the Hungarian Academy of Sciences, Den Haag: Koninklijke Bibliotheek.

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

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

Waugh, F.V. (1929), *Quality as a Determinant of Vegetable Prices*, New York: AMS Press.

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