

INDEX NUMBER THEORY AND MEASUREMENT ECONOMICS

W. Erwin Diewert.

October 15, 2014.

CHAPTER 19: Recent Developments in Consumer Price Index Theory

1. Introduction

There have been some important new developments in Consumer Price Index theory during the past few years. In this chapter, we will outline many of these new developments.¹

In section 2, the *chain drift problem* will be defined. Sections 3 and 4 will discuss two possible solutions to the chain drift problem.

Section 5 will discuss various problems associated with the construction of *elementary indexes*. These indexes are constructed using price information only. When value or quantity information is not available to the price statistician, then it is only possible to construct an elementary index.

Section 6 will briefly review recent developments on alternative approaches to *quality adjustment*.

Section 7 will discuss additional *problem areas* associated with the construction of Consumer Price Indexes where further research is required.

2. The Chain Drift Problem

The *Consumer Price Index Manual*² recommended that the Fisher or Törnqvist-Theil price index (or other superlative index) be used as a *target month to month index*, in a Consumer Price Index, provided that monthly price and expenditure data for the class of expenditures in scope were available. Recently, grocery chains in some countries (e.g., Australia, Japan, the Netherlands and Norway) have been willing to donate their sales value and quantity information by detailed product to their national statistical agencies so it has become possible to calculate month to month superlative indexes for at least some strata of the country's Consumer Price Index. However, the following issue arises: should the indexes fix a base month (for 13 months) and calculate Fisher or Theil indexes as chained indexes or as fixed base indexes? The *CPI Manual* offered the following advice on this choice in the chapter on seasonal commodities:

“22.63 A reasonable method for dealing with seasonal commodities in the context of picking a target index for a month to month CPI is the following one:³

¹ This chapter is based on Diewert (2013a).

² See the ILO, Eurostat, IMF, OECD, UN and the World Bank (2004).

³ For more on the economic approach and the assumptions on consumer preferences that can justify month to month maximum overlap indexes, see Diewert (1999a; 51-56).

- Determine the set of commodities that are present in the marketplace in both months of the comparison.
- For this maximum overlap set of commodities, calculate one of the three indices recommended in previous chapters; i.e., calculate the Fisher, Walsh or Törnqvist Theil index.

Thus the bilateral index number formula is applied only to the subset of commodities that are present in both periods.

22.64 The question now arises: should the comparison month and the base month be adjacent months (thus leading to chained indices) or should the base month be fixed (leading to fixed base indices)? It seems reasonable to prefer chained indices over fixed base indices for two reasons:

- The set of seasonal commodities which overlaps during two consecutive months is likely to be much larger than the set obtained by comparing the prices of any given month with a fixed base month (like January of a base year). Hence the comparisons made using chained indices will be more comprehensive and accurate than those made using a fixed base.
- In many economies, on average 2 or 3 percent of price quotes disappear each month due to the introduction of new commodities and the disappearance of older ones. This rapid sample attrition means that fixed base indices rapidly become unrepresentative and hence it seems preferable to use chained indices which can more closely follow marketplace developments.” ILO (2004; 407)

Thus the *Manual* recommended chained Fisher or Törnqvist-Theil indexes as a target index concepts. But as we shall see in the following paragraphs, this advice does not always work out too well.

We will first review some material on fixed base versus chained indexes. Suppose that we have decided on a “best” price index formula that compares the prices of period 0 with those of period 1, say $P(p^0, p^1, q^0, q^1)$. Suppose further that we have price and quantity data for 3 periods. There are at least two ways that a sequence of price levels for the three periods could be formed using the given index number formula:

- Fixed base indexes or
- Chained indexes.

The sequence of the price levels for the three periods under consideration, P^0 , P^1 and P^2 , using *fixed base* (or direct) *indexes* can be constructed as follows:

$$(1) P^0 \equiv 1; P^1 \equiv P(p^0, p^1, q^0, q^1); P^2 \equiv P(p^0, p^2, q^0, q^2).$$

Thus the prices in period 2, p^2 , are compared *directly* with the prices in period 0, p^0 .

The sequence of the three price levels, P^0 , P^1 and P^2 , using *chained indexes* can be constructed as follows:

$$(2) P^0 \equiv 1; P^1 \equiv P(p^0, p^1, q^0, q^1); P^2 \equiv P(p^0, p^1, q^0, q^1)P(p^1, p^2, q^1, q^2).$$

Thus fixed base and chained price levels coincide for the first two periods but in subsequent periods t , the fixed base indexes compare the prices in period t *directly* to the prices in period 0 whereas the chained indexes simply update the price level in the previous period by multiplying by the period over period *chain link index* $P(p^{t-1}, p^t, q^{t-1}, q^t)$.

The two methods of index construction will coincide if the bilateral price index formula $P(p^0, p^1, q^0, q^1)$ satisfies the following test:

$$(3) \text{Circularity Test: } P(p^0, p^1, q^0, q^1) P(p^1, p^2, q^1, q^2) = P(p^0, p^2, q^0, q^2).$$

If there is only one commodity in the aggregate, then the price index $P(p^0, p^1, q^0, q^1)$ just becomes the single price ratio, p_1^1/p_1^0 , and the circularity test becomes the equation $[p_1^1/p_1^0][p_1^2/p_1^1] = [p_1^2/p_1^0]$, which is obviously satisfied. The equation in the circularity test illustrates the difference between chained index numbers and fixed base index numbers. The left hand side of (3) uses the *chain principle* to construct the overall inflation between periods 0 and 2 whereas the right hand side uses the *fixed base principle* to construct an estimate of the overall price change between periods 0 and 2.⁴

It would be good if our preferred index number formulae, the Fisher and Törnqvist indexes (P_F and P_T), satisfied the circularity test but unfortunately, they do not satisfy (3).⁵ Hence, a statistical agency compiling a CPI has to choose between the two methods of index construction. As indicated above, the *Manual* favoured the use of chained superlative indexes for the reasons indicated above.

A main advantage of using chained indexes is that if prices and quantities are trending relatively smoothly, chaining will reduce the spread between the Paasche and Laspeyres indexes.⁶ These two indexes each provide an asymmetric perspective on the amount of price change that has occurred between the two periods under consideration and it could be expected that a single point estimate of the aggregate price change should lie between these two estimates. Thus the use of either a chained Paasche or Laspeyres index will usually lead to a smaller difference between the two and hence to estimates that are closer to the “truth”. Since annual data generally has smooth trends, the use of chained indexes is generally appropriate at this level of aggregation; see Hill (1993; 136-137).

However, the story is different at subannual levels; i.e., if the index is to be produced at monthly or quarterly frequencies. Hill (1993; 388), drawing on the earlier research of

⁴ Recall that Fisher (1911; 203) introduced this fixed base and chain terminology. The concept of chaining is due to Lehr (1885) and Marshall (1887; 373).

⁵ Alterman, Diewert and Feenstra (1999; 61-65) showed that if the logarithmic price ratios $\ln(p_n^t/p_n^{t-1})$ trend linearly with time t and the expenditure shares s_i^t also trend linearly with time, then the Törnqvist index P_T will satisfy the circularity test *exactly*. They extended this exactness result to cover the case when there are monthly proportional variations in prices and the expenditure shares have constant seasonal effects in addition to linear trends. However, when sales of products at irregular intervals occur, P_T will no longer satisfy the circularity test. Recall that section 10 in Chapter 3 looked at axioms on $P(p^0, p^1, q^0, q^1)$ that ensured that the circularity test would be satisfied. However, all known indexes that satisfy the circularity test will be subject to a certain amount of substitution bias.

⁶ See Diewert (1978; 895) and Hill (1988) (1993; 387-388). Chaining under these conditions will also reduce the spread between fixed base and chained indexes using P_F or P_T as the basic bilateral formula.

Szulc (1983) and Hill (1988; 136-137), noted that it is not appropriate to use the chain system when prices oscillate or “bounce” to use Szulc’s (1983; 548) term. This phenomenon can occur in the context of regular seasonal fluctuations or in the context of sales. The extent of the *price bouncing problem* or the problem of *chain drift* can be measured if we make use of the following test due to Walsh (1901; 389), (1921; 540):⁷

$$(4) \text{ Multiperiod Identity Test: } P(p^0, p^1, q^0, q^1)P(p^1, p^2, q^1, q^2)P(p^2, p^0, q^2, q^0) = 1.$$

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

The *Manual* did not take into account how severe the chain drift problem could be in practice.⁸ The problem is mostly caused by periodic temporary promotional *sales* of products.⁹ An example will illustrate the problem.

Suppose that we are given the price and quantity data for 2 commodities for 4 periods. The data are listed in Table 1 below.¹⁰

Table 1: Price and Quantity Data for Two Products for Four Periods

| Period t | p_1^t | p_2^t | q_1^t | q_2^t |
|----------|---------|---------|---------|---------|
| 1 | 1.0 | 1.0 | 10 | 100 |
| 2 | 0.5 | 1.0 | 5000 | 100 |
| 3 | 1.0 | 1.0 | 1 | 100 |
| 4 | 1.0 | 1.0 | 10 | 100 |

The first commodity is subject to periodic sales (in period 2), when the price drops to $\frac{1}{2}$ of its normal level of 1. In period 1, we have a “normal” off sale demand for commodity 1 which is equal to 10 units. In period 2, the sale takes place and demand explodes to 5000 units.¹¹ In period 3, the commodity is off sale and the price is back to 1 but most

⁷ This is Diewert’s (1993; 40) term for the test.

⁸ Szulc (1983) (1987) demonstrated how big the chain drift problem could be with chained Laspeyres indexes but the authors of the *Manual* did not realize that chain drift could also be a problem with chained *superlative* indexes.

⁹ Pronounced fluctuations in the prices of seasonal commodities can also cause chain drift.

¹⁰ This example is taken from Diewert (2012).

¹¹ This example is based on an actual example that used Dutch scanner data. When the price of a detergent product went on sale at approximately one half of the regular price, the volume sold shot up approximately one thousand fold; see de Haan (2008). This paper brought home the magnitude of volume fluctuations due to sales.

shoppers have stocked up in the previous period so demand falls to only 1 unit.¹² Finally in period 4, the commodity is off sale but we are back to the “normal” demand of 10 units. Commodity 2 is dull: its price is 1 in all periods and the quantity sold is 100 units in each period. Note that the only thing that has happened going from period 3 to 4 is that the demand for commodity one has picked up from 1 unit to the “normal” level of 10 units. Also note that, conveniently, the period 4 data are exactly equal to the period 1 data so that for Walsh’s test to be satisfied, the product of the period to period chain links must equal one.

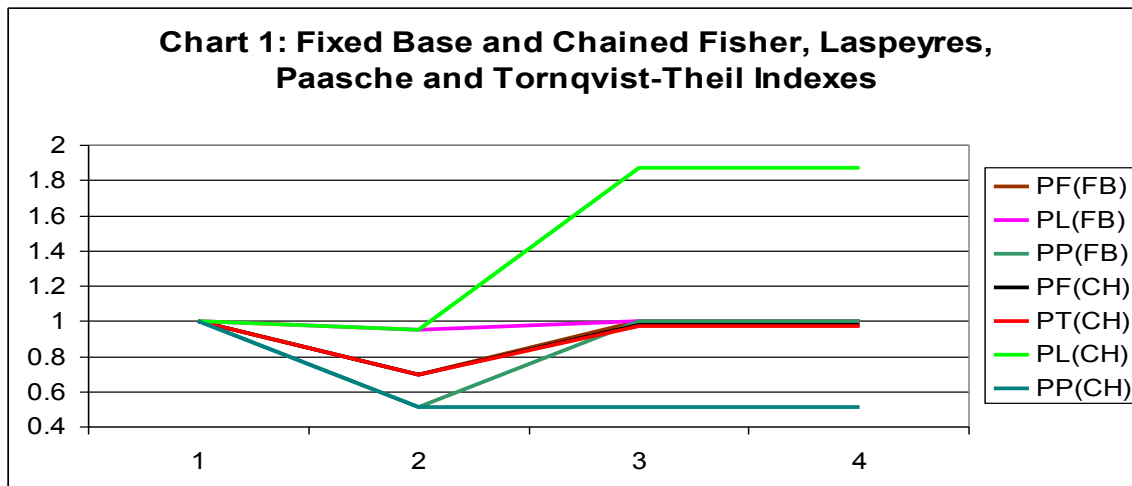
Table 2 lists the fixed base Fisher, Laspeyres and Paasche price indexes, $P_{F(FB)}$, $P_{L(FB)}$ and $P_{P(FB)}$ and as expected, they behave perfectly in period 4, returning to the period 1 level of 1. Then the chained Fisher, Törnqvist-Theil, Laspeyres and Paasche price indexes, $P_{F(CH)}$, $P_{T(CH)}$, $P_{L(CH)}$ and $P_{P(CH)}$ are listed. Obviously, the chained Laspeyres and Paasche indexes have chain link bias that is extraordinary but what is interesting is that the chained Fisher has a 2% downward bias and the chained Törnqvist has a close to 3% downward bias.

Table 2: Fixed Base and Chained Fisher, Törnqvist-Theil, Laspeyres and Paasche Indexes

| Period | $P_{F(FB)}$ | $P_{L(FB)}$ | $P_{P(FB)}$ | $P_{F(CH)}$ | $P_{T(CH)}$ | $P_{L(CH)}$ | $P_{P(CH)}$ |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 2 | 0.698 | 0.955 | 0.510 | 0.698 | 0.694 | 0.955 | 0.510 |
| 3 | 1.000 | 1.000 | 1.000 | 0.979 | 0.972 | 1.872 | 0.512 |
| 4 | 1.000 | 1.000 | 1.000 | 0.979 | 0.972 | 1.872 | 0.512 |

The above indexes are plotted in Chart 1 below. Because of the wide spreads between the chained Laspeyres and Paasche indexes, it is difficult to distinguish the small bias in the chained Fisher and Theil indexes. Nevertheless, these small biases are significant when they cumulate over long periods of time.

¹² Feenstra and Shapiro (2003) also looked at the chain drift problem that was caused by sales and restocking dynamics. Their suggested solution to the chain drift problem was to use fixed base indexes.



If the above data were monthly, and they repeated themselves 3 times over the year, the overall chain link bias would build up to the 6 to 8% range, which is significant.

What explains the results in the above table? The problem is this: when commodity one comes off sale and goes back to its regular price in period 3, *the corresponding quantity does not return to the level it had in period 1*: the period 3 demand is only 1 unit whereas the “normal” period 1 demand for commodity 1 was 10 units. It is only in period 4, that demand for commodity one recovers to the period 1 level. However, since prices are the same in periods 3 and 4, all of the chain links show no change (even though quantities are changing) and this is what causes the difficulties. If demand for commodity one in period 3 had immediately recovered to its “normal” period 1 level of 10, then there would be no chain drift problem.

There are at least three possible solutions to the chain drift problem that is associated with the use of a superlative index in a situation where monthly scanner data is available to the statistical agency for components of the CPI:¹³

- Stick to the usual annual basket Lowe (1823) index that uses annual expenditure weights from a past year;
- Use Rolling Year GEKS to control for chain drift or
- Use the Weighted Time Dummy Product method to control for chain drift.

¹³ There is a possible fourth method to avoid chain drift within a year when using a superlative index and that is to simply compute a sequence of 12 year over year monthly indexes so that say January prices in the previous year would be compared with January prices in the current year and so on. Handbury, Watanabe and Weinstein (2013) use this methodological approach for the construction of year over year monthly superlative Japanese consumer price indexes using the Nikkei point of sale data base. This data base has monthly price and expenditure data covering the years 1988 to 2010 and contains 4.82 billion price and quantity observations. This type of index number was recommended in the ILO (2004; chapter 22) as a valid year over year index that would avoid seasonality problems. However, central banks and other users require month to month CPIs in addition to year over year monthly CPIs and so the approach of Handbury, Watanabe and Weinstein does not solve the problems associated with the construction of superlative month to month indexes.

The last two methods will be explained below along with a discussion of their relative merits. The problem with the first method is that the Lowe index is subject to a small amount of upper level substitution bias, usually in the range of 0.15 to 0.40 percentage points per year.¹⁴ Note that none of the four main approaches to index number theory that were described in previous chapters endorsed the Lowe index as a target index. The widespread use of the Lowe index is due to its practical nature and the fact that the amount of substitution bias is generally not all that large.¹⁵

3. The Rolling Year GEKS Approach to Index Number Theory

We turn now to an explanation of the Rolling Year GEKS method. The GEKS method for making international index number comparisons between countries is due to Gini (1931; 12). It was derived in a different fashion by Eltetö and Köves (1964) and Szulc (1964) and thus the method is known as either the GEKS or EKS method for making multilateral comparisons. Of course, it can also be adapted to making comparisons between multiple time periods. Basically, the GEKS method in the time series context works as follows. Suppose we have price and quantity information for a component of the CPI on a monthly basis for a sequence of 13 consecutive months. Now pick one month (say month k) in this augmented year as the base month and construct Fisher price indexes for all 13 months relative to this base month. Denote the resulting sequence of Fisher indexes as $P_F(1/k)$, $P_F(2/k)$, ..., $P_F(13/k)$.¹⁶ The final set of GEKS indexes for the 13 months is simply geometric mean of all 13 of the specific month indexes; i.e., the final set of *GEKS indexes for the months in the augmented year* is any normalization of the following sequence of indexes:¹⁷

$$(5) \left[\prod_{k=1}^{13} P_F(1/k) \right]^{1/13}, \left[\prod_{k=1}^{13} P_F(2/k) \right]^{1/13}, \dots, \left[\prod_{k=1}^{13} P_F(13/k) \right]^{1/13}.$$

The above GEKS indexes have a number of important properties:¹⁸

¹⁴ For recent retrospective studies on upper level substitution bias for national CPIs, see Armknecht and Silver (2014), Diewert, Huwiler and Kohli (2009) and Huang, Wimalaratne and Pollard (2013). For studies of lower level substitution bias for a Lowe index, see Diewert, Finkel and Artsev (2009) and Diewert (2013A).

¹⁵ Recent Canadian research has indicated that the substitution bias can be reduced substantially by more frequent updating of the annual basket; see Huang, Wimalaratne and Pollard (2013).

¹⁶ Using scanner data, it is not trivial to construct these Fisher indexes. The problem is that for each pair of months, it is necessary to determine the list of products that sold in both months so that the relevant Fisher index between those two months can be constructed; see Nakamura and Steinsson (2008) and Nakamura, Nakamura and Nakamura (2011) on these difficulties.

¹⁷ Balk (1981; 74) derived the GEKS parities using this type of argument rather than the usual least squares derivation of the GEKS parities; see Balk (1996) (2008) and Diewert (1999b) for these alternative derivations.

¹⁸ The basic idea of adapting a multilateral method to the time series context is due to Balk (1981) who set up a framework that is very similar to the one explained here (which follows Ivancic, Diewert and Fox (2011) more closely). Balk (1981) used an index number formula due to Vartia (1976) in place of maximum overlap bilateral Fisher indexes as his basic building blocks and he considered augmented years of varying length instead of a 13 month augmented year but the basic idea of adapting multilateral methods to the time series context is certainly due to him.

- They satisfy Walsh's multiperiod identity test so that if any two months in the augmented year have exactly the same price and quantity vectors, then the above index values will coincide for those two months; i.e., *the above indexes are free from chain drift*.
- The above indexes do not asymmetrically single out any single month to play the role of a base period; all possible base months contribute to the overall index values.¹⁹
- The above indexes make use of all possible bilateral matches of the price data between any two months in the augmented year.
- Strongly seasonal commodities make a contribution to the overall index values.

The last property explains why the augmented year should include at least 13 consecutive months, so that *strongly seasonal commodities*²⁰ can make a contribution to the overall index.

The major problem with the GEKS indexes defined by (5) is that the indexes change as the data for a new month becomes available. A headline CPI cannot be revised from month to month due to the fact that many contracts are indexed to a country's headline consumer price index. A solution to this problem was proposed by Ivancic, Diewert and Fox (2011). Their method added the price and quantity data for the most recent month to the augmented year and dropped the oldest month from the old augmented year in order to obtain a new augmented year. The GEKS indexes for the new augmented year are calculated in the usual way and the ratio of the index values for the last month in the new augmented year to the index value for the previous month in the new augmented year is used as an *update factor* for the value of the index for the last month in the previous augmented year. The resulting indexes are called *Rolling Year GEKS indexes*.

Numerical experiments with Australian and Dutch scanner data from grocery chains show that the Rolling Year GEKS indexes work well when up to date price and quantity data are made available to the statistical agency; see Ivancic, Diewert and Fox (2011), de Haan and van der Grient (2011), Johansen and Nygaard (2011), van der Grient and de Haan (2011) and Krsinich (2011). In particular, adding and dropping a month of data and recomputing the GEKS indexes does not seem to change past index values very much.²¹ Basically, the method seems to control chain drift quite well.²² More research on the method needs to be done²³ but it looks quite promising.

¹⁹ Thus the above GEKS procedure seems to be an improvement over the suggestion of Feenstra and Shapiro (2003) who chose only a single base month.

²⁰ A strongly seasonal commodity is one that is not present in the marketplace for all months of the year.

²¹ Balk (1981; 77) also observed the same phenomenon as he computed his GEKS indexes using successively larger data sets. Diewert (2013a) also found that Rolling Year GEKS estimates were quite close to their GEKS counterparts for his small data set on Israeli seasonal commodities.

²² The Australian Bureau of Statistics plans to use RYGEKS for some components of its Consumer Price Index. Statistics Netherlands also computed RYGEKS indexes for some components of its CPI on an experimental basis with good results but they did not implement the method officially; see de Haan and van der Grient (2011). A problem is that the method is difficult to explain to users.

²³ An issue that requires further research is the effects of having different window lengths on the estimates.

We turn to an alternative method that could be used to control the chain drift problem.

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

The *Rolling Year Weighted Time Product Dummy* (RYWTPD) method for constructing indexes that are largely free of chain drift had its origins in the international comparisons literature, just as GEKS also had its origins in that literature. The *Country Product Dummy* (CPD) method²⁴ is due to Summers (1973) and a version of it (adapted to the time series context) will now be explained.²⁵

Suppose that we are attempting to make a comparison of prices over T consecutive months over a reasonably homogeneous group of say N items. Suppose initially no expenditure weights are available for the collected prices and that exactly K outlets are sampled for each of the N items in each time period. Thus there are TNK price quotes collected across all of the time periods.

Let p_{tnk} denote the price of item n in outlet k in time period t for $t = 1, \dots, T$; $n = 1, \dots, N$; $k = 1, \dots, K$. The basic statistical model that is assumed is the following one:

$$(6) \quad p_{tnk} = a_t b_n u_{tnk} ; \quad t = 1, \dots, T; n = 1, \dots, N; k = 1, \dots, K$$

where the a_t and b_n are unknown parameters to be estimated and the u_{tnk} are independently distributed error terms with means 1 and constant variances. The parameter a_t is to be interpreted as the *average level of prices* (over all items in this group of items) in time period t and the parameter b_n is to be interpreted as *multiplicative units of measurement factor* that is specific to product n .²⁶ If the error terms are all unity, then it can be seen that the N item prices move in a proportional manner over time and thus weighting is not important since all reasonable price index formula will generate the a_t as the overall price levels up to a factor of proportionality. Thus the a_t are the period t price levels that we want to determine while the b_n are product effects. The basic hypothesis is that the price of product n in outlet k in time period t is equal to a price level a_t times an item commodity adjustment factor b_n times a random error that fluctuates around 1. Taking logarithms of both sides of (6) leads to the following model:

$$(7) \quad y_{tnk} = \alpha_t + \beta_n + \varepsilon_{tnk} ; \quad t = 1, \dots, T; n = 1, \dots, N; k = 1, \dots, K$$

where $y_{tnk} \equiv \ln p_{tnk}$, $\alpha_t \equiv \ln a_t$, $\beta_n \equiv \ln b_n$ and $\varepsilon_{tnk} \equiv \ln u_{tnk}$.

²⁴ This method can be viewed as a simple type of hedonic regression model or alternatively, as a descriptive statistics method that summarizes price movements into simple indexes along the lines pioneered by Theil (1967; 136-138).

²⁵ This material on the TPD method is adapted from Rao (2004) (2005) and Diewert (2004) (2005). The term TPD is due to de Haan and Krsinich (2012) (2014). The adaptation of the CPD method to the time series context is due to Aizcorbe, Corrado and Doms (2003).

²⁶ The model assumes that the quality of the outlet k is the same for each product n . If this is not the case, then each product in each outlet should be considered a separate commodity and the k index disappears from the model.

The model defined by (7) is a linear regression model where the independent variables are dummy variables. The least squares estimators for the α_t and β_n can be obtained by solving the following least squares minimization problem:

$$(8) \min_{\alpha_t, \beta_n} \left\{ \sum_{t=1}^T \sum_{n=1}^N \sum_{k=1}^K [y_{tnk} - \alpha_t - \beta_n]^2 \right\}.$$

We also require a normalization on the α_t and β_n such as $\alpha_1 = 0$.²⁷ Solve (8) for the least squares solution parameters α_t^* and β_n^* and let $a_t \equiv \exp[\alpha_t^*]$ for $t = 2, 3, \dots, T$ and $b_n \equiv \exp[\beta_n^*]$ for $n = 1, \dots, N$. It turns out that the price level for period t , a_t , is the following expression:

$$(9) a_t = \prod_{n=1}^N \prod_{k=1}^K p_{tnk}^{1/NK} / \prod_{n=1}^N \prod_{k=1}^K p_{1nk}^{1/NK}; \quad t = 1, \dots, T.$$

Thus the TPD price level for period t (using the balanced sample of prices) is equal to the geometric mean of all of the period t prices divided by the geometric mean of all of the period 1 prices; i.e., it is a Jevons (1865) index! However, the solution is much more complicated when some outlet prices are missing from some period or when the number of outlets varies from period to period. We will deal with these more complicated situations below when we introduce weighting.

Now introduce weighting into the picture. Thus for product n in time period t , we assume that there are $K(t, n)$ outlets that have transactions in product n ²⁸ and that the *unit value price* for the k th such transaction is p_{tnk} and the associated *quantity transacted* is q_{tnk} for $k = 1, 2, \dots, K(t, n)$. Again, $y_{tnk} \equiv \ln p_{tnk}$ is the logarithm of the price p_{tnk} . For each time period t , we use the prices and quantities p_{tnk} and q_{tnk} in order to form the following *period t expenditure shares* across all products n and all outlets k :

$$(10) s_{tnk} \equiv p_{tnk} q_{tnk} / \sum_{i=1}^N \sum_{j=1}^{K(t, n)} p_{tij} q_{tij}; \quad t = 1, \dots, T; \quad n = 1, \dots, N; \quad k = 1, \dots, K(t, n).$$

For each time period t , these expenditure shares sum up to 1:

$$(11) \sum_{n=1}^N \sum_{k=1}^{K(t, n)} s_{tnk} = 1; \quad t = 1, \dots, T.$$

The *Weighted Time Product Dummy* (WTPD) counterpart to the unweighted least squares minimization problem (8) above is:

$$(12) \min_{\alpha_t, \beta_n} \left\{ \sum_{t=1}^T \sum_{n=1}^N \sum_{k=1}^{K(t, n)} s_{tnk} [y_{tnk} - \alpha_t - \beta_n]^2 \right\}.$$

²⁷ This normalization implies that $a_1 = 1$; i.e., the aggregate price level is set equal to unity in the first period. Thus the price levels for subsequent periods a_t become price indexes (relative to the level of prices in period 1).

²⁸ We allow $K(t, n)$ to be zero; i.e., it can be the case that for some time periods t , there are no price quotes collected for product n .

Again, the parameters α_t and β_n cannot be uniquely identified so we will choose to set the price level in period 1, $a_1 \equiv \exp[\alpha_1]$, equal to 1, which implies the following normalization on the parameters appearing in (12):

$$(13) \alpha_1 = 0.$$

In order to obtain a classical regression model that has a solution consistent with the weighted least squares minimization problem (12) subject to the constraint (13), we need to multiply each y_{tnk} by the square root of the associated expenditure share s_{tnk} defined by (10); i.e., the counterparts to our linear regression equations (7) are now the following equations:

$$(14) s_{tnk}^{1/2} y_{tnk} = s_{tnk}^{1/2} \alpha_t + s_{tnk}^{1/2} \beta_n + \varepsilon_{tnk}; \quad t = 1, \dots, T; n = 1, \dots, N; k = 1, \dots, K(c, n)$$

where $y_{tnk} \equiv \ln p_{tnk}$ and the α_t for $t = 2, \dots, T$ and β_n for $n = 1, \dots, N$ are parameters to be estimated (α_1 is set equal to 0) and the ε_{tnk} are assumed to be independently distributed error terms with means 0 and variances σ^2 . If for any t and n , $K(t, n) = 0$ so that there are no item n prices collected in time period t , then the corresponding equations in (14) are dropped.

In order to rigorously justify the linear regression model (14) from an econometric point of view, we need to assume that the variance of y_{tnk} is proportional to σ^2/s_{tnk} for $t = 1, \dots, T; n = 1, \dots, N; k = 1, \dots, K(t, n)$.²⁹ This means that the smaller is the expenditure share s_{tnk} , the bigger will be the variance of y_{tnk} . This assumption is not likely to be precisely justified from a statistical point of view but solving the weighted least squares problem (12) leads to very reasonable estimates for the period t price levels, $a_t \equiv \exp[\alpha_t^*]$ for $t = 2, 3, \dots, T$ where the α_t^* are the least squares estimates of the α_t for the linear regression model defined by (14). These estimates are reasonable from the viewpoint of classical index number theory, where weighting by economic importance is regarded as being extremely important. It is worth quoting Irving Fisher on the importance of weighting:

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

“But on what principle shall we weight the terms? Arthur Young’s guess and other guesses at weighting represent, consciously or unconsciously, the idea that relative *money values* of the various commodities should determine their weights. A value is, of course, the product of a price

²⁹ An alternative way for justifying the weighted model (12) is to argue that each logarithmic price $\ln p_{tnk}$ should be repeated according to its economic importance; i.e., if consumers are spending e_{tnk} dollars on commodity n during time period t , then $\ln p_{tnk}$ should appear e_{tnk} times in the regression instead of only once. In order to standardize these weights across time periods, we change the e_{tnk} weight to s_{tnk} . This type of argument was used by Diewert (2005) (2006).

per unit, multiplied by the number of units taken. Such values afford the only common measure for comparing the streams of commodities produced, exchanged, or consumed, and afford almost the only basis of weighting which has ever been seriously proposed.” Irving Fisher (1922; 45).

Thus it can be argued that solving (12) leads to index numbers that are reasonable from a *descriptive statistics point of view*; i.e., the resulting price levels are a reasonable way of summarizing overall price trends in the data, where the relative economic importance of each unit value price is taken into account in the model.³⁰

It can be verified that if the expenditure and price data are exactly the same for any two periods, then the WTDP method will generate price levels for the two periods that are also identical. Thus the WTPD estimates satisfy an identity test and hence are free of chain drift over the T periods in the sample.³¹

The WTPD price level estimates suffer from the same problem that the GEKS estimates suffer from: the addition of one more period to the sample will change all of the estimates. Thus Ivancic, Diewert and Fox (2009) proposed a Rolling Year approach to the Weighted Time Product Dummy (RYTPD) estimation procedure; i.e., set $T = 13$ and as a new month's data is added, delete the data for the oldest month in the sample, obtain new WTPD estimates and use the month over month movement in the estimated price levels for the last two months to update the previous estimates.³²

When we move from WTPD estimates to RYWTPD estimates, the Multiperiod Identity Test is no longer satisfied by the price level estimates and so the rolling year variant of the method is subject to possible chain drift. However, as was the case with the move from GEKS to RYGEKS, empirically very little difference is found between the rolling year indexes and their fixed sample counterparts.³³ Thus both the RYGEKS and RYWTPD methods seem to be largely free of chain drift.

How do the RYGEKS estimates compare with the corresponding RYWTPD estimates when using the same data set? Empirical experience is limited but in the studies that have

³⁰ There is another way of proceeding and that is to solve the weighted least squares problem but instead of assuming the stochastic specification given below (14), assume that $y_{cnk} = \alpha_c + \beta_n + \epsilon_{cnk}$ where the ϵ_{cnk} are independently distributed and have mean zero and variance σ^2 . We still solve (12) for the weighted least squares α_t^* and β_n^* but the resulting parameter estimates are no longer minimum variance unbiased for the new stochastic specification. However, the resulting estimates are still unbiased under the new stochastic specification and they are *representative* from the viewpoint of index number theory. Hill and Timmer (2006) take this point of view. Note also that Diewert (2005) derived an explicit index number formula for a_2 using the weighted least squares model defined by (12) for the two period case; i.e., the case where $T = 2$. Diewert also showed that the resulting index number formula approximated the Törnqvist-Theil index to the second order around an equal price and quantity point.

³¹ De Haan and Krsinich (2012) noted this property of the method.

³² Ivancic, Diewert and Fox (2009) is essentially the same as Ivancic, Diewert and Fox (2011) except the former paper had an extra section in it which compared the RYWTPD method to the RYGEKS method using Australian scanner data.

³³ See Ivancic, Diewert and Fox (2009) and de Haan and Krsinich (2012) (2014).

compared the two methods, there is a general tendency for the RYWTPD estimates to be slightly less than their RYGEKS counterparts.³⁴

A possible explanation for the differences in the indexes generated by the two methods may be due to the *democratic weighting* that is inherent in the GEKS method. Thus the GEKS estimates are formed by averaging a series of 13 separate sets of index numbers where the data of each month in the augmented rolling year are used as the base price and quantities in each of the bilateral indexes. If the data for one month is sparse so that the value of transactions in that month is unusually low and perhaps not “typical”, then these atypical indexes are averaged with all of the other 12 sets of indexes and given an equal weight in the averaging process.³⁵ On the other hand, the WTPD method would automatically give a much lower weight to the possibly atypical prices in the low volume month. *The WTPD method works on a principle that tries to fit heterogeneous price movements over the sample into a simpler framework where all price movements are approximated by proportional movements in prices over time, taking into account the economic importance of the prices.*

A simple (extreme) example may help to illustrate possible problems with the GEKS methodology. Suppose we have price and expenditure share data for 3 products for 3 periods but each product is present in only 2 of the 3 periods. Suppose the first product is present in periods 1 and 2 with prices p_1^1 , p_1^2 , the second product is present in periods 2 and 3 with prices p_2^2 , p_2^3 and the third product is present in periods 1 and 3 with prices p_3^1 , p_3^3 . The period 1 expenditure shares for products 1 and 3 are s_1^1 and s_3^1 , the period 2 expenditure shares for products 1 and 2 are s_1^2 and s_2^2 and the period 3 expenditure shares for products 2 and 3 are s_2^3 and s_3^3 . The expenditure shares for each period sum to one. Because of the missing data, we can only calculate 3 matched product bilateral indexes across the 3 periods. The Fisher index for period 2 relative to period 1, $P_F(2/1)$, turns out to equal the price ratio p_1^2/p_1^1 ; the Fisher index for period 3 relative to period 2, $P_F(3/2)$, turns out to equal the price ratio p_2^3/p_2^2 and the Fisher index for period 3 relative to period 1, $P_F(3/1)$, turns out to equal the price ratio p_3^3/p_3^1 . We can compute three separate set of price levels using different combinations of the available bilateral indexes.

The first set of parities uses the index $P_F(2/1)$ to determine the period 2 price level relative to the period 1 level and the period 3 price level relative to the period 1 level is determined as $P_F(3/1)$. The resulting price levels are the following ones:

$$(15) P^1 = 1 ; P^2 = (p_1^2/p_1^1) ; P^3 = (p_3^3/p_3^1).$$

The second set of parities uses the index $P_F(2/1)$ to determine the period 2 price level relative to the period 1 level and the period 3 price level relative to the period 1 level is

³⁴ See Ivancic, Diewert and Fox (2009) and de Haan and Krsinich (2012) (2014). The latter authors used the Törnqvist-Theil index formula as their basic bilateral formula in their RYGEKS estimates instead of the Fisher index but it is unlikely that this formula difference would affect the results.

³⁵ This situation occurs frequently in the context of making international comparisons of prices using the GEKS method.

determined as the product $P_F(2/1)$ times $P_F(3/2)$.³⁶ The resulting price levels are the following ones:

$$(16) P^1 = 1 ; P^2 = (p_1^2/p_1^1) ; P^3 = (p_1^2/p_1^1)(p_2^3/p_2^2).$$

The third set of parities uses the index $P_F(3/1)$ to determine the period 3 price level relative to the period 1 level and the period 2 price level relative to the period 1 level is determined as the product $P_F(3/1)$ times $P_F(2/3)$. The resulting price levels are the following ones:

$$(17) P^1 = 1 ; P^2 = (p_3^3/p_3^1)(p_2^2/p_2^3) ; P^3 = (p_3^3/p_3^1).$$

The price levels defined by (15) are a normalization of the Fisher parities generated by using period 1 as the base period, while the price levels defined by (16) and (17) are normalizations of the Fisher parities that use periods 2 and 3 as the base periods respectively.

Taking the geometric mean of the above price levels leads to the following GEKS price levels:

$$(18) P^1 = 1 ; P^2 = [(p_1^2/p_1^1)^2(p_3^3/p_3^1)(p_2^2/p_2^3)]^{1/3} ; P^3 = [(p_1^2/p_1^1)(p_2^3/p_2^2)(p_3^3/p_3^1)^2]^{1/3}.$$

Note that the GEKS price levels do not depend on the expenditure shares. However, the Weighted Time Product Dummy price levels for this example *will* depend on the expenditure shares. The exact formula for these price levels is too complicated to be exhibited here but we know that the WTPD price levels will be weighted according to the size of the expenditure shares in each period.³⁷ In particular, suppose the commodity 3 expenditure shares, s_3^1 and s_3^3 , are tiny. Then the WTPD price levels will be close to the price levels defined by (16) (which do not involve the prices p_3^1 and p_3^3). On the other hand, suppose the commodity 2 expenditure shares, s_2^2 , s_2^3 , are close to zero. Then the WTPD price levels will be close to the price levels defined by (15) (which do not involve the prices p_2^2 and p_2^3). Finally, suppose the commodity 1 expenditure shares, s_1^1 , s_1^2 , are close to zero. Then the WTPD price levels will be close to the price levels defined by (17) (which do not involve the prices p_1^1 and p_1^2). In each of the three cases just considered, the WTPD price levels are very reasonable; the unimportant commodity is given a low weighting in the overall index but this is not the case for the GEKS price levels: the GEKS price levels remain the same under all three scenarios! Thus if price movements are far from proportional over time, so that the price levels defined by (15)-(17) are very different, then the GEKS indexes may be rather far removed from their WTPD counterparts, which will be much more reasonable in each of the three cases considered above. These possible problems with the GEKS indexes carry over to Rolling Year GEKS indexes.

³⁶ Because there is only one commodity whose price is compared in each bilateral Fisher index, these bilateral Fisher (and Laspeyres and Paasche) indexes all collapse down to simple price ratios.

³⁷ In this example, $K(t,n)$ is always equal to 1; i.e., we have only one outlet for each product.

More research is required on pinning down the differences between the GEKS estimates and their WTPD counterparts but at this stage, we tentatively conclude that in the case where the period to period data is sparse and there is a lack of product matching for each pair of periods under consideration, the WTPD estimates may be preferable to the corresponding GEKS estimates (and the RYWTPD estimates may be preferable to the corresponding RYGEKS estimates).

5. Elementary Indexes: New Developments

The ILO *Manual* basically recommended the Jevons formula for elementary indexes.³⁸ This advice was based on the axiomatic approach to elementary indexes; see Diewert (1995). In the case of complete data on a sample of products with no sample attrition, we end up with the formula (9) in the previous section. But real life does not generate complete samples with no attrition; products disappear and then reappear³⁹ and some products disappear permanently. If a product disappearance is thought to be temporary, price statisticians typically impute the missing prices. But there are many imputation methods⁴⁰ and this creates a certain amount of uncertainty about the accuracy of the index at any give time period.

There is another problem associated with elementary indexes that the *Manual* did not deal with and that is the fact that many statistical agencies do not chain their elementary indexes every month: they choose a reference month and then calculate item prices relative to the item price in the base month for 13 months.⁴¹ The problem with this strategy is that the procedure depends asymmetrically on the choice of the base month. Some items will not be available in the base month and so how are we to treat these items which appear in subsequent months?

Thus chaining elementary price quotes is problematic (due to the necessity of imputing prices for temporarily disappearing items and for strongly seasonal items) and using a fixed base methodology for elementary indexes is also problematic (due to the fact that some products may not be available in the base month and more generally, due to the asymmetry of choosing one month out of 12 months as the base month).

A solution to these problems was suggested by de Haan and Krsinich (2012) and Diewert (2012): use the Time Product Dummy methodology in order to construct elementary indexes. These TPD indexes are generated by solving the least squared minimization

³⁸ Elementary indexes are constructed using item prices only (due to the unavailability of quantity information).

³⁹ This can happen with strongly seasonal products or the temporary disappearance may be due to the fact that retailers sometimes rotate the brand items that they sell in order to generate price discounts from manufacturers.

⁴⁰ See Feenstra and Diewert (2001) for a review of alternative imputation methods and references to the literature.

⁴¹ Thus the Retail Prices Index in the UK uses January as its base month whereas the Harmonized Index of Consumer Prices used as the official Eurostat measure of European household inflation uses December as its base month for a sequence of 13 months; see Diewert (2012) for more discussion on these short term fixed base methods. In many other countries, month to month chained elementary indexes are used.

problem (12) above, except the expenditure shares s_{tnk} are all set equal to one. The resulting elementary indexes, $a_1 = 1$, $a_2 = \exp[\alpha_2]$, ..., $a_N = \exp[\alpha_N]$ have a large number of good axiomatic properties.⁴² This new approach to the construction of elementary indexes seems promising. De Haan and Krsinich (2012) and Diewert (2012) suggested that the TPD methodology could be generalized into a Rolling Year Time Product Dummy (RYTPD) method where a moving sample of 13 consecutive months of item price data is used to generate TPD price levels and then the movement in the index for the last two months is used to update the previous index.

The RYTPD method for constructing elementary indexes seems to be very promising. It is relatively easy to implement, there are no imputations required for the method and it treats the price data for each period in a symmetric manner.

More research into the method is required. It would also be useful to look at variants of the method that allowed for longer windows; e.g., instead of using 13 consecutive months of data, perhaps more stable estimates may be obtained using 25 consecutive months of data in each RYTPD regression.⁴³

6. New Approaches to Quality Adjustment

A problem with the RYGEKS and RYWTPD methods described above is that these methods do not deal adequately with the introduction of new products. Thus if a new product enters the marketplace during the last period in the Rolling Year, it will have no effect on the index for the current period and all previous periods. De Haan and Krsinich (2012) (2014) invented a method that deals with this problem. The basic building block in their method is a time dummy hedonic regression model that uses the data for two periods. The dependent variable in the model is the logarithm of the item price and a time dummy and various characteristics of the product enter the regression as independent variables. The time dummy coefficient and the characteristic “prices” are the result of a weighted least squares minimization problem. If an item appears in both periods under consideration, the weights in the weighted regression are the (arithmetic) average of the expenditure shares for the item in the two periods; if the item appears in only one of the two periods, one half of the expenditure share on the item for that period is used as the weight. The resulting bilateral price index turns out to equal the usual Törnqvist index if all items are present in both periods but for unmatched items, an imputed price for the missing price enters the index number formula and this imputed price is obtained as a predicted price using hedonic regression. Thus in the general case when there are unmatched items in the two periods under consideration, we obtain a generalization of the usual Törnqvist index that makes use of imputed prices from the hedonic regression and hence de Haan and Krsinich (2012) call the resulting bilateral index number formula the *Imputation Törnqvist index*.⁴⁴ De Haan and Krsinich (2012) (2014) proposed the

⁴² See de Haan and Krsinich (2012) (2014). If approximate share weights are available, then these approximate weights should be used and the Weighted TPD method should be used.

⁴³ If the time period is longer than an (augmented) year, then the term Rolling *Window* TPD method is more appropriate.

⁴⁴ This index is derived in de Haan and Krsinich (2012)(2014) and draws on earlier contributions by Diewert (2003) and de Haan (2003) (2004).

following variation of the Rolling Year GEKS method: instead of using bilateral Fisher indexes as the basic building blocks, the Fisher indexes are replaced by bilateral Imputation Törnqvist indexes. They call the resulting indexes ITRYGEKS indexes.⁴⁵

Which of the three methods discussed above is “best”? Methods 1 and 2 (RYGEKS and RYWTPD) have the disadvantage that unmatched items in any bilateral index used as building blocks in these methods have no impact on the resulting indexes. But these methods have the advantage that no information on product characteristics is required in order to implement these indexes. Method 3 (ITRYGEKS) has the advantage that it is likely to have the least amount of bias due to the introduction of new models and the disappearance of old models but of course, it has the disadvantage that product characteristics information is required in order to implement the method. The bottom line is that ITRYGEKS may be the best method that can deal with chain drift and quality change in the context of using scanner data.⁴⁶

De Haan and Krsinich (2012) (2014) and Krsinich (2013) showed that for electronic products in New Zealand, the RYWTPD indexes were closer to their “gold standard” ITRYGEKS indexes than their RYGEKS counterparts. This is a somewhat surprising result since it is known that in the two period case where all products are present in both periods, RYGEKS and RYWTPD approximate each other closely.⁴⁷ However, the results presented by Krsinich indicate that this close correspondence does not necessarily hold in more realistic environments when not all products are present in all periods. The implication of the results presented by de Haan and Krsinich is that when information on product characteristics is not available, the RYWTPD method may be preferred to the RYGEKS method. This is an important result but more research on this is required.⁴⁸

The results derived by de Haan and Krsinich required the availability of scanner data on sales and the prices of various electronic products. Suppose the statistical agency does not have access to data on sales and prices. What is the “best” approach to quality adjustment when only price data is collected? This is an open question.

The work by de Haan and Krsinich is perhaps the most important work on the theory and practice of quality adjustment that has appeared since the publication of the *Manual*. Of lesser importance is the research by Diewert, Heravi and Silver (2009) and de Haan (2010) that examines more closely the differences between the time dummy approach and the hedonic imputation approach to hedonic regressions.

7. Long Time Problems with the CPI that Still Need to be Addressed

⁴⁵ It might be more appropriate to call these indexes ITRYCCD indexes since multilateral Caves, Christensen and Diewert (1982) indexes are used in place of multilateral GEKS indexes in the case where all items are matched.

⁴⁶ However, it should be possible to adapt the WTPD method to deal with quality change.

⁴⁷ See Diewert (2005; 564).

⁴⁸ This result reinforces the earlier misgivings about the democratic nature of the GEKS indexes in the context of sparse data.

There are several long standing problems associated with the construction of a CPI that have troubled national statisticians over the years. The author's list of vexing problems is the following list:

- Should the CPI be compiled on a domestic, national or household inflation basis?
- What is the appropriate treatment of Owner Occupied Housing (OOH) in the CPI?
- How exactly should financial services be treated in the CPI?
- How can strongly seasonal commodities make a contribution to the month to month CPI?
- Is it possible to construct real time CPIs, using current month information on prices and older information on household expenditure shares, that will approximate a superlative CPI that is constructed later when additional data on expenditures shares become available?

These problems will be discussed in the following subsections.

7.1. National versus Domestic Versus Household Inflation

If the CPI is to be used to deflate the consumption expenditures of domestic households in a nation, then the national CPI should be calculated on a domestic basis; i.e., the expenditures of households that primarily reside in the nation should be used as weights in the national CPI. Many economists use the domestic CPI to deflate nominal consumer expenditures to look at the welfare of residents in the country. Thus from the viewpoint of welfare economics, nations should provide a national CPI. However, when we look at the production accounts of a country (and the associated Multifactor Productivity or Total Factor Productivity accounts of a nation), we require a “domestic” CPI which is usually labelled as the Domestic Consumption Deflator. This deflator is used to convert all sales of consumer goods generated by producers domiciled in the country into real domestic consumption. Thus we require *both* the national and domestic consumption deflators to fill in the deflator cells in the System of National Accounts.⁴⁹

The Harmonized Index of Consumer Prices was introduced by Eurostat to measure consumer inflation across countries belonging to the European Union in a comparable manner. The HICP is widely used by central banks use to gauge relative inflation rates across member countries. This is a valid household inflation index but it does not provide a substitute for the national and domestic consumer price indexes described in the above paragraph.⁵⁰

7.2 The Problem of Owner Occupied Housing in the CPI

⁴⁹ The main difference between the two indexes is the expenditures of nationals abroad (this is in scope for the national concept) and the expenditures of tourists in the home country (this is in scope for the domestic concept).

⁵⁰ The Eurostat HICP index was originally introduced as an index that would make absolutely no imputations. However, over time, it was recognized that quality change required imputations and eventually the HICP allowed imputations for quality change. Diewert (2002) criticized the HICP from the viewpoint that it did not exactly fit into the System of National Accounts. However, if central banks want a minimal imputation index that measures consumer inflation, then the HICP fills a very useful function.

There is no consensus on how Owner Occupied Housing (OOH) should be treated in the CPI. The main approaches to the treatment of OOH are as follows:

- The *rental equivalence approach*. In this approach, the value of the services of OOH is the rent that the owned unit could accrue if it were rented.
- The *monetary expenditures approach*. In this approach, the out of pocket costs of home ownership are totalled to provide an imputed “rent”.⁵¹
- The *acquisitions approach*. In this approach, the ownership of previously purchased housing properties is ignored; only newly constructed housing units are in scope.⁵²
- The *user cost approach*. In this approach, the imputed value of housing services is set equal to the financial cost of tying up owner’s capital in the house.
- The *opportunity cost approach*. In this approach, the maximum of the user cost and rental equivalent price is used as the valuation of the services of the housing unit.⁵³

Since no consensus on the appropriate approach to the valuation of the services of Owner Occupied Housing has been achieved in the literature, it seems reasonable to ask national statistical agencies to provide analytical series for all five approaches to the valuation of housing services. To date, these analytical series have not been forthcoming!

7.3 The Measurement of Financial Services in the CPI

Financial services are an important component of GDP and a somewhat important component of household consumption. What is amazing is that there is absolutely no agreement on how to measure these services. The main components of financial services for households are the services provided by their monetary deposits and insurance services for property and life. There are other financial services that are easier to measure such as stock trading (this is basically a margin industry and can be treated in a manner similar to wholesaling and retailing).⁵⁴

Reviewing all of the alternative treatments of financial services that have been suggested in the literature would be a major task. Suffice it to say that there is extreme

⁵¹ The problem with this approach is that it does not list all of the (opportunity) costs and benefits of ownership. The main missing costs are depreciation and the financial capital tied up in the equity of the housing unit and the main benefit that is missing is the expected capital gains (or losses) on the property. This is my least preferred alternative treatment of owner occupied housing.

⁵² This is the variant that the HICP has chosen to implement. There are two variants of the method: include only the structure portion of the new building or include the structure and land components together.

⁵³ See Diewert (2011) and Diewert, Nakamura and Nakamura (2009) for a description of the alternative approaches.

⁵⁴ However, in wholesaling and retailing, the price of the service is the margin times the price of the products being purchased. In the case of stock trading, the appropriate price is not completely clear.

heterogeneity in these treatments.⁵⁵ It would be good if academics could turn their attention to these basic measurement problems in the area of financial services in the near future.

7.4 How Can Strongly Seasonal Commodities Make a Contribution to the Month to Month CPI?

The answer to the above question is relatively straightforward in the light of the analysis that we have done above in looking at the properties of the GEKS and WTPD methods. For both of these methods, strongly seasonal commodities play a role in the overall index construction. Similarly, if we look at elementary indexes, the TPD method explained above deals adequately with strongly seasonal commodities.

The practical question to be resolved in the coming years is whether the RYGEKS is better than the RYWTPD method. This is an open question.

At the elementary level, the RYTPD method seems to be a much superior option to other methods for constructing elementary indexes.

7.5 Predicting Superlative CPIs Using Current Prices and Past Expenditure Shares

Is it possible to construct a real time CPI using current month information on prices and older information on household expenditure shares that will approximate a superlative CPI that is constructed later when additional data on expenditures shares become available? It seems unlikely that we will be able to approximate a superlative CPI perfectly using currently available data but recent research has indicated that it is possible to obtain pretty good approximations to a superlative CPI using current data; see Armknecht and Silver (2014) and Huang, Wimalaratne and Pollard (2013). This is an promising area which requires more research.

8. Conclusion

In the decade since the *Consumer Price Index Manual* appeared, there have been some significant new developments in the theory and practice of CPI construction. A new development is the fact that some supermarket firms in some countries are willing to share their price and quantity data with national statistical agencies. Hopefully, this spirit of cooperation will spread to other countries.⁵⁶

With the advent of scanner data availability, it becomes possible to compute the type of indexes that have been recommended by index number theorists over the past century.

⁵⁵ For materials on variants of the user cost approach to measuring financial sector inputs and outputs, see Barnett (1980), Hancock (1985) (1991), Wang, Basu and Fernald (2009), Colangelo and Inklaar (2012), Diewert, Fixler and Zieschang (2013a) (2013b), Inklaar and Wang (2013) and Diewert (2014).

⁵⁶ Many retailers have the information in files; the costs of providing this information to national statistical agencies is trivial.

But new problems have emerged; in particular the problem of chain drift for superlative indexes has emerged.

This chapter has indicated methods for overcoming the chain drift problem. When price and quantity information is available, the Rolling Year GEKS or the Rolling Year Weighted Time Product Dummy method is recommended. For high tech products that are undergoing rapid technological change, the methods for quality adjustment developed by de Haan and Krsinich are recommended. At the level of elementary indexes, the Rolling Year Time Product Dummy method is recommended.

However, the above methods have not been thoroughly tested and so perhaps some caution is in order. Hopefully, further research in the coming years will demonstrate whether the suggested methods are definitely preferred.

Finally, in section 7 above, some long standing problem areas with respect to CPI construction have been highlighted. It would be good if some progress could be made on resolving these problems in the coming decade.

References

- Alterman, W.F., W.E. Diewert and R.C. Feenstra (1999), *International Trade Price Indexes and Seasonal Commodities*, Bureau of Labor Statistics, Washington D.C.
- Armknrecht, P. and M. Silver (2014), "Post-Laspeyres: The Case for a New Formula for Compiling Consumer Price Indexes", *Review of Income and Wealth* 60:2, 225-244.
- Aizcorbe, A., C. Corrado and M. Doms (2003), "When Do Matched-Model and Hedonic Techniques Yield Similar Price Measures?", Working Paper 2003-14, Federal Reserve Bank of San Francisco.
- Balk, B.M. (1981), "A Simple Method for Constructing Price Indices for Seasonal Commodities", *Statistische Hefte* 22 (1), 1-8.
- Balk, B.M. (1996), "A Comparison of Ten Methods for Multilateral International Price and Volume Comparisons", *Journal of Official Statistics* 12, 199-222.
- Balk, B.M. (2008), *Price and Quantity Index Numbers*, New York: Cambridge University Press.
- Barnett, W.A. (1980), "Economic Monetary Aggregates: An Application of Index Number and Aggregation Theory", *Journal of Econometrics* 14, 11-48.
- Caves D.W., Christensen, L.R. and Diewert, W.E. (1982), "Multilateral Comparisons of Output, Input, and Productivity using Superlative Index Numbers", *Economic Journal* 92, 73-86.

- Colangelo, A. and R. Inklaar (2012), "Bank Output Measurement in the Euro Area: A Modified Approach", *Review of Income and Wealth* 58:1, 142-165.
- de Haan, J. (2003), "Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement", Paper presented at the Seventh Meeting of the Ottawa Group, Paris, France, May 27-29.
- de Haan, J. (2004), "The Time Dummy Index as a Special Case of the Imputation Törnqvist Index", Paper presented at the Eighth Meeting of the Ottawa Group, Helsinki, Finland, August 23-25.
- de Haan, J. (2008), "Reducing Drift in Chained Superlative Price Indexes for Highly Disaggregated Data", paper presented at the EMG Workshop 2008, December 10-12, Crowne Plaza Hotel, Coogee Beach, Sydney.
- de Haan, J. (2010), "Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and 'Re-pricing' Methods", *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)* 230, 772-791.
- de Haan, J. and H.A. van der Grient (2011), "Eliminating Chain drift in Price Indexes Based on Scanner Data", *Journal of Econometrics* 161, 36-46.
- de Haan, J. and F. Krsinich (2012), "The Treatment of Unmatched Items in Rolling Year GEKS Price Indexes: Evidence from New Zealand Scanner Data", paper presented at the Meeting of Groups of Experts on Consumer Price Indices Organized jointly by UNECE and ILO at the United Nations Palais des Nations, Geneva Switzerland, May 30-June 1, 2012.
- de Haan, J. and F. Krsinich (2014), "Scanner Data and the Treatment of Quality Change in Rolling Year GEKS Indexes", *Journal of Business and Economic Statistics* 32:3, 341-358.
- Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica* 46, 883-900.
- Diewert, W.E. (1993), "The Early History of Price Index Research", pp. 33-65 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.
- Diewert, W.E. (1995), "Axiomatic and Economic Approaches to Elementary Price Indexes", Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.
- Diewert, W.E. (1999a), "Index Number Approaches to Seasonal Adjustment", *Macroeconomic Dynamics* 3, 1-21.

- Diewert, W.E. (1999b), "Axiomatic and Economic Approaches to International Comparisons", pp. 13-87 in *International and Interarea Comparisons of Income, Output and Prices*, A. Heston and R.E. Lipsey (eds.), Studies in Income and Wealth, Volume 61, Chicago: The University of Chicago Press.
- Diewert, W.E. (2002), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations", *Swiss Journal of Economics and Statistics* 138:4, 547-637.
- Diewert, W.E. (2003), "Hedonic Regressions: A Review of Some Unresolved Issues", Paper presented at the Seventh Meeting of the Ottawa Group, Paris, France, May 27-29.
- Diewert, W.E. (2004), "On the Stochastic Approach to Linking the Regions in the ICP", Department of Economics, Discussion Paper 04-16, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1.
- Diewert, W.E. (2005), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", *The Review of Income and Wealth* 51:4, 561-571.
- Diewert, W.E. (2006), "Adjacent Period Dummy Variable Hedonic Regressions and Bilateral Index Number Theory", *Annales d'économie et de statistique*, No. 79/80, 1-28.
- Diewert, W.E. (2011), "The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Further Directions", pp. 87-116 in *Price and Productivity Measurement, Volume 1, Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Publishing.
- Diewert, W.E. (2012), *Consumer Price Statistics in the UK*, Government Buildings, Cardiff Road, Newport, UK, NP10 8XG: Office for National Statistics.
<http://www.ons.gov.uk/ons/guide-method/userguidance/prices/cpi-and-rpi/index.html>
- Diewert, W.E. (2013a), "An Empirical Illustration of Index Construction using Israeli Data on Vegetables" paper presented at the 13th Meeting of the Ottawa Group On Prices, May 1-3, Copenhagen, Denmark.
- Diewert, W.E. (2013b), "The Consumer Price Index: Recent Developments", Discussion Paper 13-12, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada.
- Diewert, W.E. (2014), "The Treatment of Financial Transactions in the SNA: A User Cost Approach", *Eurostat Review on National Accounts and Macroeconomic Indicators* 1, 73-89.

- Diewert, W.E., Y. Finkel and Y. Artsev (2009), “Empirical Evidence on the Treatment of Seasonal Products: The Israeli Experience”, pp. 53-78 in *Price and Productivity Measurement: Volume 2: Seasonality*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Victoria, Canada: Trafford Press.
- Diewert, W.E., D. Fixler and K. Zieschang (2013a), “The Measurement of Banking Services and the System of National Accounts”, forthcoming in *Price and Productivity Measurement: Volume 3—Services*, W. Erwin Diewert, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Trafford Press.
- Diewert, W.E., D. Fixler and K. Zieschang (2013b), “Problems with the Measurement of Banking Services in a National Accounting Framework”, forthcoming in *Price and Productivity Measurement: Volume 3—Services*, W. Erwin Diewert, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Trafford Press.
- Diewert, W.E., S. Heravi and M. Silver (2009), “Hedonic Imputation versus Time Dummy Hedonic Indexes”, pp. 161-196 in *Price Index Concepts and Measurement*, Studies in Income and Wealth, Volume 20, W.E. Diewert, J. Greenlees and C. Hulten (eds.), Chicago: University of Chicago Press.
- Diewert, W.E., M. Huwiler and U. Kohli (2009), *Swiss Journal of Economics and Statistics* 145(2), 127-135.
- Diewert, W.E., A.O. Nakamura and L.I. Nakamura (2009), “The Housing Bubble and a New Approach to Accounting for Housing in a CPI”, *Journal of Housing Economics* 18:3, 156-171.
- Eltető, Ö., and Köves, P. (1964), “On a Problem of Index Number Computation Relating to International Comparisons”, (in Hungarian), *Statisztikai Szemle* 42, 507-518.
- Feenstra, R.C. and W.E. Diewert (2001), “Imputation and Price Indexes: Theory and Evidence from the International Price Program”, Bureau of Labor Statistics Working Paper 335, U.S. Department of Labor, Washington D.C.; <http://www.bls.gov/ore/pdf/ec010030.pdf>
- Feenstra, Robert C. and Matthew D. Shapiro (2003), “High-Frequency Substitution and the Measurement of Price Indexes”, pp. 123-146 in *Scanner Data and Price Indexes*, Robert C. Feenstra and Matthew D. Shapiro (eds.), Studies in Income and Wealth Volume 64, Chicago: The University of Chicago Press.
- Fisher, I. (1911), *The Purchasing Power of Money*, London: Macmillan.
- Fisher, I. (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.
- Gini, C. (1931), “On the Circular Test of Index Numbers”, *Metron* 9:9, 3-24.

- Hancock, D. (1985), “The Financial Firm: Production with Monetary and Non-Monetary Goods”, *Journal of Political Economy* 93, 859-880.
- Hancock, D. (1991), *A Theory of Production for the Financial Firm*, Boston: Kluwer Academic Press.
- Handbury, J., T. Watanabe and D.E. Weinstein (2013), “How Much do Official Price Indexes Tell us about Inflation”, paper presented at the Economic Measurement Group Workshop Asia 2013, Tokyo, Japan, October 16.
- Hill, T.P. (1988), “Recent Developments in Index Number Theory and Practice”, *OECD Economic Studies* 10, 123-148.
- Hill, T.P. (1993), “Price and Volume Measures”, pp. 379-406 in *System of National Accounts 1993*, Eurostat, IMF, OECD, UN and World Bank, Luxembourg, Washington, D.C., Paris, New York, and Washington, D.C.
- Hill, R.J., and M.P. Timmer (2006), “Standard Errors and Weights in Multilateral Price Indexes”, *Journal of Business and Economic Statistics* 24:3, 366-377.
- Huang, N., W. Wimalaratne and B. Pollard (2013), “Investigation on Reducing the Substitution Bias in the Canadian CPI”, paper presented at the Meeting of the Statistics Canada Price Measurement Advisory Committee, October 2, Ottawa.
- ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004), *Consumer Price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: International Labour Office.
- Inklaar, R. and J.C. Wang (2013), “Real Output of Bank Services: What Counts is What Banks Do, Not What They Own”, *Economica* 80, 96-118.
- Ivancic, L., W.E. Diewert and K.J. Fox (2009), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, Discussion Paper 09-09, Department of Economics, University of British Columbia, Vancouver, Canada.
- Ivancic, L., W.E. Diewert and K.J. Fox (2011), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, *Journal of Econometrics* 161, 24-35.
- Jevons, W.S., (1865), “The Variation of Prices and the Value of the Currency since 1782”, *Journal of the Statistical Society of London* 28, 294-320; reprinted in *Investigations in Currency and Finance* (1884), London: Macmillan and Co., 119-150.
- Johansen, I. and R. Nygaard (2011), “Dealing with Bias in the Norwegian Superlative Price Index of Food and Non-alcoholic Beverages”, Paper presented at the Twelfth meeting of the Ottawa Group, May 4-6, Wellington, New Zealand.

- Krsinich, F. (2011), "Price Indexes from Scanner Data: A Comparison of Different Methods", Paper presented at the twelfth meeting of the International Working Group on Price Indices (Ottawa Group), 4-6 May 2011, Wellington, New Zealand.
- Krsinich, F. (2013), "Using the Rolling Year Time Product Dummy Method for Quality Adjustment in the Case of Unobserved Characteristics", Ottawa Group 2013 Meeting, Copenhagen, Denmark, May 1-3.
- 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.
- Nakamura, A.O., E. Nakamura and L.I. Nakamura (2011), "Price Dynamics, Retail Chains and Measurement," *Journal of Econometrics*, 151;1, 47-55.
- Nakamura, E., and J. Steinsson (2008), "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics* 123, 1415-1464.
- Rao, D.S. Prasada (2004), "The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power parities in the ICP", paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, June 30-July 3, 2004, Vancouver, Canada.
- Rao, D.S. Prasada (2005), "On the Equivalence of Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", *Review of Income and Wealth* 51:4, 571-580.
- Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.
- Szulc, B.J. (1964), "Indices for Multiregional Comparisons", (in Polish), *Przegląd Statystyczny* 3, 239-254.
- 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.
- Szulc, B.J. (1987), "Price Indices below the Basic Aggregation Level", *Bulletin of Labour Statistics* 2, 9-16.

- Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland Publishing.
- van der Grient, H.A. and J. de Haan (2011), “Scanner Data Price Indexes: The ‘Dutch Method’ versus Rolling Year GEKS”, paper presented at the 12th Meeting of the Ottawa Group, May 4-6, Wellington, New Zealand.
- Vartia, Y.O. (1976), “Ideal Log-Change Index Numbers”, *Scandinavian Journal of Statistics* 3, 121-126.
- Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.
- Walsh, C. M. (1921), “Discussion”, *Journal of the American Statistical Association* 17, 537-544.
- Wang, J.C., S. Basu and J.G. Fernald (2009), “A General Equilibrium Asset Approach to the Measurement of Nominal and Real Bank Output ”, pp. 273-328 in *Price Index Concepts and Measurement*, W.E. Diewert, J. Greenlees and C. Hulten (eds.), Studies in Income and Wealth, Volume 70, Chicago: University of Chicago Press.