Journal of Official Statistics, Vol. 18, No. 1, 2002, pp. 61–85

Generalised Fisher Price Indexes and the Use of Scanner Data in the Consumer Price Index (CPI)

Jan de Haan[1]

Statistics Netherlands intends to use scanner data provided by retailers in compiling the Consumer Price Index (CPI). This has two important advantages. First, taking a sample of items to estimate the commodity group price index for a particular type of outlet becomes unnecessary. Second, the Laspeyres-type index formula currently applied can be replaced by an index formula that is better suited for handling dynamic aspects such as commodity substitution and the introduction of new goods. This article suggests the use of a so-called generalised Fisher price index, based on a set of goods that is variable but overlapping over time. It contains prices of new and disappearing goods that cannot be observed directly and that should be imputed. The relation with quality adjustment procedures is addressed as well.

Key words: Consumer price index; imputation; new goods; quality adjustment; substitution.

1. Introduction

In the past, empirical research on the Consumer Price Index (CPI), as well as the actual CPI calculation, had to rely on survey data. During the last couple of years the increasing availability of bar-code scanning data has provided the opportunity to exploit the entire set of commodities (goods for short) belonging to a commodity group. This has stimulated the discussion on the preferred treatment of new and disappearing goods, which is a highly important aspect in the compilation of a CPI. Statistics Netherlands intends to use scanner data provided by nation-wide retailers in the production process of its CPI. The basic idea is to use the data on all goods at the commodity group level instead of taking samples, and to switch over the currently used Laspeyres formula to the Fisher formula.[2]

Scanner data provide for each product a description and information on sales (the number of transactions), expenditures and times of purchase. Average transaction prices (unit values) are easily computed. For the outlets considered, the unit value covers all transactions in a certain period, not just the display price on a given day in a month, which is what statistical agencies usually collect. Unlike those from market research companies like GfK, the scanner data obtained directly from retailers often do not include detailed

[1] Statistics Netherlands, Division of Macro-economic Statistics and Dissemination, Support and Development Department, P.O. Box 4000, 2270 JM Voorburg, The Netherlands. Email: jhhn@cbs.nl.

[2] For fast-moving consumer goods sold in supermarkets various problems were encountered, most of which will be discussed in this article. As a short-run solution Statistics Netherlands chose to implement a Laspeyres price index with annual reweighting. In the longer run, and especially for durable goods, the Fisher formula is still preferred.

Acknowledgments: The views expressed in this article are those of the author and do not necessarily reflect the policies of Statistics Netherlands. Any errors or omissions are also the responsibility of the author. He would like to thank Bert Balk, Jörgen Dalén, Leendert Hoven, Mick Silver and two anonymous referees for helpful comments and Eddy Opperdoes for research assistance.

product characteristics. Diewert (1993) was the first to propose the use of scanner data for price measurement. A growing body of empirical literature is emerging in this field, e.g., Silver (1995), Reinsdorf (1996), Bradley et al. (1998), Dalén (1998), Lowe (1998), Moulton et al. (1998), De Haan et al. (1999), and Hawkes and Smith (1999). Recently, Fenwick et al. (2001), Richardson (2001), and Lowe and Ruscher (2001) looked into the relative merits of scanner data for official CPI compilation in the U.K., the U.S.A. and Canada, respectively. They all caution that such data are not without problems. One of the problems could be the presence of (some) business-to-business sales in scanner data, which is obviously undesirable. Another problem might be the classification of the goods according to international standards like COICOP (Classification Of Individual Consumption by Purpose). Though Schut (2001) does not report major deficiencies in scanner data obtained directly from Dutch supermarkets, she also recognises the need for careful checking, and possibly cleaning, of the data. The classification problem has been resolved with the help of a branch organisation.

This article suggests the use of a Fisher price index that explicitly accounts for new and disappearing goods, which will be called a generalised Fisher price index. The aim is not so much to develop any new ideas about CPI construction but rather to place various old ideas in the context of the generalised Fisher price index and to find out what we can learn from these ideas when we have scanner data at our disposal for an entire commodity group. Owing to the work of Diewert (1976) it has become generally accepted that a superlative price index like Fisher's reasonably approximates a cost of living index. In a challenging contribution, Balk (2000a) addresses the new goods bias and the substitution bias of the Laspeyres-CPI with respect to a cost of living index. Unfortunately there is no international consensus among statistical agencies on the question whether the theory of the cost of living index should be the underlying conceptual framework for the CPI (Triplett 2001). Statistics Netherlands has always been, and still remains, one of the advocates of using this framework. However, the present article takes a more pragmatic point of view. Hence it may appeal to those agencies that feel somewhat uncomfortable with the cost of living index methodology.

The remainder of this article is organised as follows. Section 2 describes current practices used to estimate Laspeyres commodity group price indexes, and reviews some well-known drawbacks. Section 3 outlines the general idea behind the generalised Fisher price index. This index is based on a variable set of goods. It contains prices of new and disappearing goods that cannot be observed directly, which should be imputed. Section 4 discusses the main features of the generalised Fisher price index. Section 5 compares Balk's approach with ours and shows that the two methods are compatible. Section 6 makes a case for using chained indexes. Section 7 presents empirical evidence using supermarket scanner data. Section 8 dicusses the findings and points to possible future empirical work in this field. In the appendix it is shown how quality-adjusted unit values can link a disappearing good to a newly introduced one when both goods are close substitutes.

2. The Laspeyres Price Index Estimator

The population Laspeyres price index for some commodity group I describes how the cost of purchasing the fixed set I^0 of goods belonging to I in the base period 0 evolves over time. For the current (comparison) period t this index is expressed as

$$P_{I,L}^{t} = \frac{\sum_{i \in I^0} p_i^t q_i^0}{\sum_{i \in I^0} p_i^0 q_i^0} \tag{1}$$

where p_i^s is the price of good i in period s ($s = 0, t$) and q_i^0 the quantity sold in the base period. Note that p_i^s will generally be some kind of average. With i deemed homogeneous, the unit value taken over (a sub-set of) all outlets that sell i–that is, the total expenditure on i divided by the total quantity bought–is the relevant average transaction price concept. See also Balk (1998).

In order to estimate $P_{I,L}^t$ the statistical agency draws a sample of items $\hat{I}^0$ from I^0. The usual estimator becomes

$$\hat{P}_{I,L}^{t} = \frac{\sum_{i \in \hat{I}^0} p_i^t q_i^0}{\sum_{i \in \hat{I}^0} p_i^0 q_i^0} = \sum_{i \in \hat{I}^0} \hat{w}_i^0 (p_i^t / p_i^0) \tag{2}$$

where the weight $\hat{w}_i^0 = p_i^0 q_i^0 / \Sigma_{i \in \hat{I}^0} p_i^0 q_i^0$ denotes the expenditure share of i with respect to the sample. In practice the weights are often proxies. If the sample size is small, the sampling variance of $\hat{P}_{I,L}^t$ may turn out to be rather large.[3] Normally a sampled item cannot be called a homogeneous good: the item description leaves enough room for the price collectors to select different varieties of the item in different outlets. The unit value taken over outlets will not then be a meaningful price concept, and the unit value index should not be used as an indicator of the item price index p_i^t / p_i^0. Hence, another type of item price index is called for. Since quantity and/or expenditure data at such detailed levels of aggregation are generally lacking, statistical agencies calculate item price index numbers from price data only.[4]

One of the problems associated with using the Laspeyres index (1) is that, by holding the quantities fixed at base period levels, substitution that takes place when consumers adjust their consumption behaviour in reaction to relative price changes is not taken into account. The question then arises how statistical offices treat new and disappearing goods. Some agencies regularly update their samples at least partially to account for new goods, like the BLS through annual sample rotation. Most agencies, however, re-sample only at base year revisions. Although this might not be good practice, it is consistent with the Laspeyres price index (1), which ignores new goods altogether. Suppose that a statistical agency decides to adhere strictly to the Laspeyres principle. If it wants to keep the sample size fixed over time, the agency is forced to act when a sampled good disappears from the market. In that case it selects another item that replaces the ''old'' one, probably the most similar item. The newly sampled item does not have to be completely new; it may already

[3] Depending on the specific sampling design used, the estimator may exhibit small-sample bias (De Haan et al. 1999). Under sampling proportional to base period expenditure, the weights are left out of the estimation formula; they are implicitly reflected by the inclusion probabilities.
[4] The U.S. Bureau of Labor Statistics (BLS), for example, uses the ratio of geometric means of the prices observed in (a sample of) outlets to estimate most item price indexes. Statistics Netherlands still uses the ratio of arithmetic means but is considering changing to geometric means; see De Haan and Opperdoes (2001a).

have been sold during the base period. To make the replaced item and its successor comparable, so that a quality change does not affect the price index, a quality adjustment must be made.

Quality adjustment seems a suitable name in the context of a survey-based Laspeyres price index estimator. But what if the entire set of goods is taken into consideration instead of a sample of items? There may not be an obvious one-to-one relation between a disappearing good and a successor. It may even be the case that the number of goods belonging to commodity group I decreases over time. The important thing to recognise is that quality adjustment methods are essentially imputation methods. Let $I^{0(D)}$ denote the sub-set of I^0 that disappeared in period t. The prices p_i^t for all $i \in I^{0(D)}$ in the Laspeyres price index (1) are ''fictitious'' in the sense of being unobservable directly since there are no (monetary) transactions involved. These prices must be imputed. In a critical review of the harmonised index of consumer prices (HICP) constructed by many European statistical offices, Diewert (2001) also notes that imputations of this kind cannot be avoided but claims that HICP regulations prohibit the use of imputations for non-monetary transactions. It is a matter of terminology, though. Quality adjustments do form part of the HICP methodology. Apparently the term imputations has been avoided to ensure that quality adjustments are allowed.

3. The Generalised Fisher Price Index

Price indexes are usually defined on sets of goods that are fixed over time. In real life most sets are not fixed at all: apart from existing goods that disappear from the market, new goods enter also. Before turning to variable sets, we take a look at the Paasche price index. Let I^t be the set of goods belonging to commodity group I in period t and q_i^t the quantity sold of good i. The Paasche price index reads

$$P_{I,P}^t = \frac{\sum_{i \in I^t} p_i^t q_i^t}{\sum_{i \in I^t} p_i^0 q_i^t} \tag{3}$$

The set of goods that are new in period t will be denoted by $I^{t(N)}$. Base period prices p_i^0 cannot be observed directly for all $i \in I^{t(N)}$. These prices are fictitious–just like p_i^t for $i \in I^{0(D)}$ in the Laspeyres price index (1)–and must be imputed. The Paasche index excludes goods that disappeared after the base period, whereas the Laspeyres index ignores new goods. Both formulas also neglect substitution effects. To handle such dynamic changes we should look for an alternative price index formula. Using a symmetry argument, the squared root of the product of the Laspeyres and Paasche price indexes seems a sensible candidate. This leads to the *generalised Fisher price index*, which is thus defined on a variable set of goods:

$$P_{I,F}^t = \left[\frac{\sum_{i \in I^0} p_i^t q_i^0}{\sum_{i \in I^0} p_i^0 q_i^0} \frac{\sum_{i \in I^t} p_i^t q_i^t}{\sum_{i \in I^t} p_i^0 q_i^t} \right]^{1/2} \tag{4}$$

in which $p_i^t = \hat{p}_i^t$ for $i \in I^{0(D)}$ and $p_i^0 = \hat{p}_i^0$ for $i \in I^{t(N)}$ are imputed prices.

The choice for the Fisher (ideal) price index is sometimes justified on other grounds, particularly with reference to the test or axiomatic approach to measuring aggregate consumer price change. Diewert (1992) showed that the Fisher price index satisfies 20 ‘‘reasonable’’ tests, which is more than its competitors satisfy. These tests are based on fixed sets of goods; it is not necessarily true that the same would hold if similar tests were developed for variable sets. The symmetry argument, on the other hand, applies both to the generalised and to the ordinary fixed-set Fisher price index.

Let $I^{0t} = I^0 \cap I^t$ denote the set of goods common to period 0 and period t. This set corresponds to Dalén’s (1998b) intersection universe. It is assumed that $I^{0t} \neq \emptyset$. The Fisher price index defined on the set I^{0t} is

$$P^t_{I^{0t},F} = \left[\frac{\sum_{i \in I^{0t}} p_i^t q_i^0}{\sum_{i \in I^{0t}} p_i^0 q_i^0} \frac{\sum_{i \in I^{0t}} p_i^t q_i^t}{\sum_{i \in I^{0t}} p_i^0 q_i^t} \right]^{1/2} \tag{5}$$

Using (5), the generalised Fisher index (4) can be decomposed into three factors:

$$P^t_{I,F} = P^t_{I^{0t},F} \left[\frac{\sum_{i \in I^t} p_i^t q_i^t \Big/ \sum_{i \in I^{0t}} p_i^t q_i^t}{\sum_{i \in I^0} p_i^0 q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^0} \right]^{1/2} \left[\frac{\sum_{i \in I^0} p_i^t q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^t q_i^0}{\sum_{i \in I^t} p_i^0 q_i^t \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^t} \right]^{1/2} \tag{6}$$

The second factor of (6) re-scales $P^t_{I^{0t},F}$ for the fact that the expenditures of new and disapppearing goods have not been taken into account. $P^t_{I^{0t},F}$ and the re-scaling factor only contain variables that can–at least in principle–be observed directly, for instance through scanner data. The third factor, which contains imputed prices, is needed to handle new and disappearing goods in the correct way.

The third factor of decomposition (6) deserves special attention. It can be rewritten as

$$\left[\frac{1 + \sum_{i \in I^{0(D)}} \hat{p}_i^t q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^t q_i^0}{1 + \sum_{i \in I^{t(N)}} \hat{p}_i^0 q_i^t \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^t} \right]^{1/2} \tag{7}$$

The imputed prices $\hat{p}_i^t$ in the numerator of (7) might be obtained by making use of a quality adjustment method. In particular a hedonic regression could be run on data pertaining to the goods sold in period t, i.e., on (a sample of) the set I^t, apart from practical problems that may arise due to the lack of data on the goods’ price-determining characteristics. With respect to the goods belonging to the set $I^{t(N)}$ in the denominator of (7) it is worthwhile distinguishing between ‘‘more or less’’ new goods having technical characteristics of already existing goods, and completely new goods representing a new technology. The imputed prices $\hat{p}_i^0$ of the first sub-set of new goods could again in principle be estimated using hedonics, this time performed on data from the base period set I^0. Fisher price indexes combined with the use of hedonic regression belong to the class of superlative hedonic price indexes; see e.g., Ioannidis and Silver (1999). For the second sub-set of new goods, the completely new ones, the conceptual and practical problems are much bigger. These problems will be touched upon in Section 6.

4. Biases of the Laspeyres Price Index Estimator

The second and third factors of decomposition (6) both contain price and quantity data of new as well as disappearing goods. To gain further insight it can be helpful to separate new goods from disappearing goods. Rewriting the second factor of (6) as

$$\left[\frac{1+\sum_{i\in I^{t(N)}} p_i^t q_i^t \Big/ \sum_{i\in I^{0t}} p_i^t q_i^t}{1+\sum_{i\in I^{0(D)}} p_i^0 q_i^0 \Big/ \sum_{i\in I^{0t}} p_i^0 q_i^0}\right]^{1/2} \tag{8}$$

and subsequently rearranging the product of (8) and (7) yields

$$P_{I,F}^t = P_{I^{0t},F}^t \mu_I^t \lambda_I^t \tag{9}$$

as an alternative decomposition, in which

$$\mu_I^t = \left[\frac{1+\sum_{i\in I^{t(N)}} p_i^t q_i^t \Big/ \sum_{i\in I^{0t}} p_i^t q_i^t}{1+\sum_{i\in I^{t(N)}} \hat{p}_i^0 q_i^t \Big/ \sum_{i\in I^{0t}} p_i^0 q_i^t}\right]^{1/2};\lambda_I^t = \left[\frac{1+\sum_{i\in I^{0(D)}} \hat{p}_i^t q_i^0 \Big/ \sum_{i\in I^{0t}} p_i^t q_i^0}{1+\sum_{i\in I^{0(D)}} p_i^0 q_i^0 \Big/ \sum_{i\in I^{0t}} p_i^0 q_i^0}\right]^{1/2}$$

The factors μ_I^t and λ_I^t can be seen as the effects of new and disappearing goods, respectively, on the generalised Fisher price index $P_{I,F}^t$. If the set I^{0t} of ongoing goods is very large compared to the sets $I^{t(N)}$ and $I^{0(D)}$, then μ_I^t and λ_I^t will both exhibit values close to 1, and $P_{I^{0t},F}^t$ might be an acceptable proxy for $P_{I,F}^t$. But we cannot expect this to be the case a priori.

Notice that $\mu_I^t < 1$ if and only if

$$\frac{\sum_{i\in I^{t(N)}} p_i^t q_i^t}{\sum_{i\in I^{0t}} p_i^t q_i^t} < \frac{\sum_{i\in I^{t(N)}} \hat{p}_i^0 q_i^t}{\sum_{i\in I^{0t}} p_i^0 q_i^t}$$

which is equivalent to requiring

$$P_{I^{t(N)},P}^t = \frac{\sum_{i\in I^{t(N)}} p_i^t q_i^t}{\sum_{i\in I^{t(N)}} \hat{p}_i^0 q_i^t} < \frac{\sum_{i\in I^{0t}} p_i^t q_i^t}{\sum_{i\in I^{0t}} p_i^0 q_i^t} = P_{I^{0t},P}^t \tag{10}$$

provided that there is at least one new good. Both the left-hand side and the right-hand side of (10) are Paasche-type price indexes, defined on the sets of new and ongoing goods, respectively. Similarly, $\lambda_I^t > 1$ if and only if

$$P_{I^{0(D)},L}^t = \frac{\sum_{i\in I^{0(D)}} \hat{p}_i^t q_i^0}{\sum_{i\in I^{0(D)}} p_i^0 q_i^0} > \frac{\sum_{i\in I^{0t}} p_i^t q_i^0}{\sum_{i\in I^{0t}} p_i^0 q_i^0} = P_{I^{0t},L}^t \tag{11}$$

provided that at least one good has disappeared. The left-hand and right-hand side of (11) are Laspeyres-type price indexes, defined on the sets of disappearing and ongoing goods, respectively. Why would inequality (11) hold? Suppose no excess demand exists, i.e., there is no supply rationing, so that consumers can freely buy any goods they want in any quantities. Suppose also that markets are transparent and consumers are well informed. Hence, the disappearance of goods implies that demand has shifted away from goods with obsolete characteristics towards similar goods (either already existing goods or ''more or less new'' ones), or that demand simply fell to zero. This must mean that the imputed prices $\hat{p}_i^t$ of the obsolete goods are too high relative to the prices of the ongoing goods. Durable goods experiencing rapid technological change, such as personal computers, might be an example. The same line of reasoning suggests that new goods (varieties) would have been bound to exhibit relatively high base period prices had they been sold during that period, which is consistent with the economics of (completely) new goods. Thus in ''purely competitive markets'' we expect (10) to hold too.

The demand-oriented view taken above fails if consumers cannot freely choose between substitutes, particularly in times of rationed supply. Suppose the manufacturer and/or retailer of some durable good decides not to sell a specific model of a specific brand any longer. At the same time a new model is introduced. The price is increased, but by more than what could have been expected from the quality improvement. If consumers cannot easily switch over to other brands which may have similar models (for instance because markets do not work well), they face a hidden price increase. Such a case is likely to occur in monopolistic-type markets, since the producer and/or retailer must have a considerable degree of market power to behave like this. It is now unclear whether the imputed prices $\hat{p}_i^t$ of disappearing goods will be relatively high or low. Whether λ_I^t will lie above or below unity thus appears to be an empirical matter. The same goes for μ_I^t.

Recapitulating, by using the Fisher price index based on the matched set of goods I^{0t} only, it is uncertain whether the generalised Fisher price index (4) will be overstated or understated. The direction of the error depends on the prevailing market conditions and marketing pricing policies, and its size partly depends on the size of the matched set. Whatever the answer might be, the discussion suggests that incorporating a proxy variable for the degree of competition into hedonic models that are used to impute unobservable prices is a promising approach. This was proposed earlier by Silver and Heravi (2001).

The phenomenon of new/disappearing goods is strongly related to that of substitution; it is merely a specific type of substitution.[5] Substitution effects and the effects of new and disappearing goods are usually treated separately, though. Substitution among ongoing goods (i.e., within the set I^{0t}) is probably what most people have in mind when speaking loosely of substitution. This may be called substitution in a narrow sense. The first factor of decomposition (9), the Fisher price index $P^t_{I^{0t},F}$, takes account of this. The second and third factors, μ_I^t and λ_I^t, capture all other forms of substitution. Suppose that the statistical agency aims at estimating the Laspeyres price index (1). It is easy to check that

[5] Balk (2000a) argues along the following lines. In each time period the (representative) consumer is confronted with a set of available goods and a corresponding set of prices and makes his or her choice. In a later period some (new) goods are purchased that were not bought before; some (disappearing) goods are no longer bought; and some goods continue to be purchased albeit perhaps in different quantities. So there are four types of substitution: among ongoing goods, between new and ongoing goods, between ongoing and disappearing goods, and between new and disappearing goods.

$P^t_{I,L} = P^t_{I^{0t},L}(\lambda^t_I)^2$. Where $P^t_{I,F} = P^t_{I^{0t},F}\mu^t_I\lambda^t_I$, the difference between the expectation of $\hat{P}^t_{I,L}$ and $P^t_{i,F}$ describes the bias of the Laspeyres price index estimator with respect to the generalised Fisher price index.

A straightforward way to split the bias into three additive terms is:

$$E\hat{P}^t_{I,L} - P^t_{I,F} = [E\hat{P}^t_{I,L} - P^t_{I,L}] + (\lambda^t_I)^2[P^t_{I^{0t},L} - P^t_{I^{0t},F}] + P^t_{I,F}\left[\frac{\lambda^t_I}{\mu^t_I} - 1\right] \tag{12}$$

The first term on the right-hand side of (12) represents ''statistical bias,'' which mainly depends on the imputations for disappearing goods made in practice. The sign of this term, albeit unknown a priori, will most likely be positive if the statistical agency substantially overstates the imputed prices $\hat{p}^t_i$ or, less precisely expressed, if it undervalues quality improvements. Boskin et al. (1996) suspected this to be the case for the U.S.A. CPI. Notice that when there is no item sampling involved, as will be the case with scanner data, the first term becomes $P^t_{I^{0t},L}[(\hat{\lambda}^t_I)^2 - (\lambda^t_I)^2]$, where $\hat{\lambda}^t_I$ is a non-survey estimate of λ^t_I. The second term of Equation (12) can be referred to as substitution bias in a narrow sense. We expect it to be positive. With $\lambda^t_I \geq 1$ ($\lambda^t_I < 1$) the difference $P^t_{I^{0t},L} - P^t_{I^{0t},F}$ yields a lower (upper) bound to this type of bias that can easily be calculated from scanner data.

The third term of decomposition (12) might be called dynamic universe bias of the Laspeyres price index. This term is in fact made up of two separate parts:

$$P^t_{I,F}\left[\frac{\lambda^t_I}{\mu^t_I} - 1\right] = \lambda^t_I P^t_{I,F}\left[\frac{1}{\mu^t_I} - 1\right] + P^t_{I,F}[\lambda^t_I - 1] \tag{13}$$

The first term of Equation (13) measures new goods bias, whereas the second term measures ''disappearing goods bias.'' In competitive markets, where we expect to find $\lambda^t_I > 1$ and $\mu^t_I < 1$, both terms will most likely be positive. The second term needs further explanation since price statisticians (and economists) will not associate the Laspeyres price index with disappearing goods bias. Suppose for convenience that no new goods have been introduced, so that the first term of (13) would be zero, but some existing ones disappeared. Thus, the population size–the number of goods in the universe–decreased. If the period t prices for disappearing goods in the population Laspeyres price index (1) were imputed correctly, then the first component of (12) would be (approximately) zero too. Unlike the generalised Fisher index, however, the (fixed-set) Laspeyres index ignores the diminishing size of the population. The Laspeyres price index must therefore be biased, unless of course $\lambda^t_I = 1$. That is what the second part of (13) relates to.

5. A Comparison with Balk's Approach

Balk (2000a) addresses the CPI's substitution and new goods bias from a cost of living index perspective. He starts by making two assumptions: i) the preference structure of the representative consumer exhibits homotheticity, and ii) the unit cost (expenditure) function is of the CES (constant elasticity of substitution) type. Assumption ii) states that for any pair of goods i, j $(i \neq j)$ the demand elasticity of substitution $\sigma = -d \ln(q^t_i/q^t_j)/d \ln(p^t_i/p^t_j)$ is the same ($\sigma \geq 0$; $\sigma \neq 1$) and that σ is also time-invariant. Assumption i) says that the optimal expenditure shares are independent of the utility level.

Furthermore, it is assumed that iii) the actual expenditure shares in both the base period and the comparison period are equal to the optimal shares.

A feature of Balk's (2000a) approach–which makes it quite interesting to compare it with ours–is that the set of goods under consideration is variable but overlapping over time. He proceeds by assuming a two-level structure in the consumer's preferences, which consists of unchanging commodity groups (at the upper level) and changing ranges of goods within these groups (at the lower level). Within each group the elasticity of substitution is assumed to be constant, but between groups it may differ. It appears that requiring within-group substitution elasticities to be larger than 1 is consistent with the inter-group elasticity being smaller than 1.

It is shown that the cost of living index, or rather the subindex for a certain commodity group I, can be expressed in a number of ways as the product of a conventional price index, defined on the set of goods common to the base period and the comparison period, and a factor depending on the change of the range of goods. Recast in our notation, the cost of living subindex can be expressed as

$$P^t_{I^{0t},(.)} \left[\frac{\sum_{i \in I^t} p_i^t q_i^t \Big/ \sum_{i \in I^{0t}} p_i^t q_i^t}{\sum_{i \in I^0} p_i^0 q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^0} \right]^{1/1-\sigma} \tag{14}$$

where $P^t_{I^{0t},(.)}$ denotes the price index defined on the intersection I^{0t} according to some conventional formula (.). For reasons of simplicity, a subscript for the commodity group has not been added to σ. Although in theory the index formula to be used depends on the value of σ, we can safely choose the Fisher formula because all relevant types of index numbers will be much alike in practice. Thus

$$P^t_{I^{0t},F} \left[\frac{\sum_{i \in I^t} p_i^t q_i^t \Big/ \sum_{i \in I^{0t}} p_i^t q_i^t}{\sum_{i \in I^0} p_i^0 q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^0} \right]^{1/1-\sigma} \tag{15}$$

can be used to approximate the cost of living subindex. This result is similar to that of Aizcorbe et al. (2000), who used a matched-item Törnqvist index instead of a matched-item Fisher index.

Comparing (15) with (6) suggests that, if the generalised Fisher price index (4) is meant to approximate a cost of living subindex, the third factor of (6)–that is, expression (7)–must be

$$\left[\frac{1 + \sum_{i \in I^{0(D)}} \hat{p}_i^t q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^t q_i^0}{1 + \sum_{i \in I^{t(N)}} \hat{p}_i^0 q_i^t \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^t} \right]^{1/2} = \left[\frac{\sum_{i \in I^t} p_i^t q_i^t \Big/ \sum_{i \in I^{0t}} p_i^t q_i^t}{\sum_{i \in I^0} p_i^0 q_i^0 \Big/ \sum_{i \in I^{0t}} p_i^0 q_i^0} \right]^{\frac{1+\sigma}{2(1-\sigma)}} \tag{16}$$

Now suppose that some new goods have been introduced, but no existing goods have disappeared. Expression (16) then reduces to

$$\left[1+\sum_{i\in I^{t(N)}}\hat{p}_i^0 q_i^t \Big/ \sum_{i\in I^{0t}} p_i^0 q_i^t\right]^{-1/2} = \left[1+\sum_{i\in I^{t(N)}} p_i^t q_i^t \Big/ \sum_{i\in I^{0t}} p_i^t q_i^t\right]^{\frac{1+\sigma}{2(1-\sigma)}} \tag{17}$$

Since the left-hand side is smaller than 1, Expression (17) can only hold for $\sigma > 1$. Balk (2000a) arrives at the same conclusion about the value of σ, albeit in a rather different way. The following relation can be derived from (17):

$$\frac{P^t_{I^{t(N)},P}}{P^t_{I^{0t},P}} = \frac{\eta^t}{(1+\eta^t)^{\frac{\sigma+1}{\sigma-1}}-1} \tag{18}$$

where $\eta^t = \sum_{i\in I^{t(N)}} p_i^t q_i^t / \sum_{i\in I^{0t}} p_i^t q_i^t$ is the ratio of the period t expenditures on new and ongoing goods. Note that the period t expenditure share of new goods is $w_N^t = \eta^t/(1+\eta^t)$. Equation (18) describes the ratio of a Paasche-type price index of new goods and a Paasche price index of the ongoing goods. Since $\sigma > 1$ and $\eta^t > 0$, this relative price change must be smaller than 1, which is in agreement with our intuitive finding (10) for ''purely competitive markets.'' It is an increasing function of η^t for given σ and a decreasing function of σ for given η^t.

As an illustration, Table 1 contains percentage relative Paasche-type price changes, that is $[(P^t_{I^{t(N)},P}/P^t_{I^{0t},P}) - 1]100\%$, evaluated at various values of σ and η^t. For example, Table 1 says that a substitution elasticity of 1.5 and a current market share of 4.8% of new goods (i.e., $\eta^t = 0.05$) corresponds to an (imputed) Paasche-type price change of new goods which lies 81.9% below the Paasche price change of ongoing goods. Notice the sensitivity of the relative price change of new goods to changes in the value of σ. As can be seen from Expression (18), $P^t_{I^{t(N)},P} \approx P^t_{I^{0t},P}$ for extremely large values of σ. Take for instance $\eta^t = 0.05$. For $\sigma = 40$ the price change of new goods is 5% below that of ongoing goods, while $\sigma = 100$ is necessary to bring it down to 2%. In this case all goods belonging to commodity group I become almost identical from an economic point of view, and we would expect to find equal price trends for all goods.

Suppose next that some goods disappeared from the market, but no new goods were introduced. Expression (16) now reduces to

$$\left[1+\sum_{i\in I^{0(D)}}\hat{p}_i^t q_i^0 \Big/ \sum_{i\in I^{0t}} p_i^t q_i^0\right]^{1/2} = \left[1+\sum_{i\in I^{0(D)}} p_i^0 q_i^0 \Big/ \sum_{i\in I^{0t}} p_i^0 q_i^0\right]^{\frac{\sigma+1}{2(\sigma-1)}} \tag{19}$$

Table 1. Percentage relative (Paasche-type) price change of new goods

	$\eta^t = 0.025$ $w_N^t = 0.024$	$\eta^t = 0.05$ $w_N^t = 0.048$	$\eta^t = 0.10$ $w_N^t = 0.091$	$\eta^t = 0.20$ $w_N^t = 0.167$	$\eta^t = 0.40$ $w_N^t = 0.286$
$\sigma = 1.2$	−92.0	−93.0	−94.6	−96.9	−99.0
$\sigma = 1.5$	−81.0	−81.9	−83.6	−86.6	−90.9
$\sigma = 2.0$	−67.5	−68.3	−69.8	−72.5	−77.1
$\sigma = 5.0$	−33.7	−34.1	−34.9	−36.4	−39.1

From (19) it follows that

$$\frac{P^t_{I^{0(D)},L}}{P^t_{I^{0t},L}} = \frac{(1+\xi^0)^{\frac{\sigma+1}{\sigma-1}} - 1}{\xi^0} \tag{20}$$

where $\xi^0 = \sum_{i \in I^{0(D)}} p_i^0 q_i^0 / \sum_{i \in I^{0t}} p_i^0 q_i^0$ is a shorthand notation for the ratio of the base period expenditures on disappearing goods and ongoing goods. Since $\sigma > 1$ and $\xi^0 > 0$, $P^t_{I^{0(D)},L} / P^t_{I^{0t},L}$ must be larger than 1. This result is in agreement with our earlier intuitive finding (11), provided that markets work well and consumers are free to choose any goods they like.

Although our approach does not rest on the theory of the cost of living index, it seems that Balk's (2000a) method does not conflict with ours–at least not when focusing entirely on the demand side (as Balk is in fact doing). He also proposes some simple methods for estimating σ. So perhaps his approach enables us to approximate the generalised Fisher price index from observable variables only, for instance according to (15), without having to rely on imputed prices. The ongoing goods should be classified according to the economic criterion of equal within-group substitution elasticities. But how should new goods be dealt with? Balk proposes to start by allocating goods which are new in period t to certain groups to the best of our knowledge. Once information about period $t+1$ becomes available their position can be reconsidered, since at that time they belong to the set of ongoing goods. Eventually this procedure might lead to a recomputation of the cost of living index going from period t to period $t-1$. Hence, he proposes to impute unobservable substitution elasticities for new goods instead of imputing unobservable base period prices.

6. Chained Indexes

New goods should be incorporated into the CPI as soon as possible. The natural method for doing this is to use chained indexes. Suppose that the direct index (4) is replaced by the product of the month-to-month generalised Fisher price indexes $P^{\tau/\tau-1}_{I,F}$:

$$P^t_{I,cF} = \prod_{\tau=1}^{t} P^{\tau/\tau-1}_{I,F} \tag{21}$$

Chaining causes i) the set of ongoing goods not to shrink too much over time and ii) the sets of new and disappearing goods not to grow too large. Point ii) has obvious advantages for estimating the imputed prices $\hat{p}$. For example, it makes little sense to estimate in period t a base period price for a new good when the base period lies in the far past.

The chained index $P^t_{I,cF}$ is path dependent, whether or not the set of goods changes over time. That is, the index number in period t does not only depend on the prices and quantities in the base period 0 and the comparison period t, but also on the prices and quantities of all time periods $\tau = 2, \ldots, t-1$ in between. Path dependency is not so problematic if the CPI is primarily viewed as a short-term indicator. Most empirical studies, based on fixed sets of goods, tend to show no large systematic differences between direct and chained Fisher indexes anyway.

According to Balk (2000b), one might view any chained price index as an approximation to the line integral Divisia price index. He argues that the Divisia price index can be given a meaningful interpretation using micro-economic theory of consumer behaviour if the usual assumption of a static preference ordering is relaxed. This assumption becomes less realistic when the time interval to which the index relates grows. Balk concludes that the Divisia price and quantity indexes can conceptually be viewed as the ''ultimate economic price and quantity indices.''

The imputed prices $\hat{p}_i^{\tau-1}$ for completely new goods introduced in period τ are called (Hicksian) reservation prices. The reservation price is the fictitious price that would reduce the demand for the product to zero had it been available in the period prior to its introduction. Some economists argue that reservation prices can really be computed in practice (see e.g., Hausman 1999), while others criticise the concept altogether (Hill 1999). There appear to be no statistical agencies that are planning to apply the concept of reservation prices to their CPIs. One may doubt whether it would make much difference in practice provided that completely new goods enter the index shortly after their introduction on the market; neglecting those goods in (21) will in general have a minor impact. Though reservation prices are typically high, the quantities sold during the introduction period will be small. Moreover, the introduction of completely new goods does not happen frequently, at least not compared to that ''more or less'' new goods, which are merely new varieties of existing products.

Although through chaining new goods will be incorporated into the CPI as quickly as possible, calculating the commodity group price index from data of the set of goods common to period τ and period $\tau - 1$ only is not a perfect solution, even without the introduction of completely new goods. An obvious imperfection is the loss of information incurred. For example, the prices and quantities of new goods in period τ, while observable, are omitted from the computation. A related issue is that exact matching and chaining does not necessarily yield an accurate approximation to the chained generalised Fisher price index due to neglecting the effects of new and disappearing goods, similarly to the analysis in Section 5 for the direct index.

7. Some Empirical Evidence

Statistics Netherlands receives scanner data directly from two of the largest Dutch supermarket chains, which are going to be used (according to current plans) in the computation of the CPI, starting from May 2002. Here, a selection has been made from the provisional CPI database, covering expenditure and quantity data on nine commodity groups from week 26, 1999, up to and including week 49, 2000. The weekly data were aggregated into data pertaining to 19 four-week periods (instead of calendar months, which would be required for the CPI). The European Article Number (EAN) identifies scannable products. Different EAN codes will be treated as separate goods, notwithstanding that different codes may represent items that are identical from the consumer's perspective; Section 8 addresses this problem.

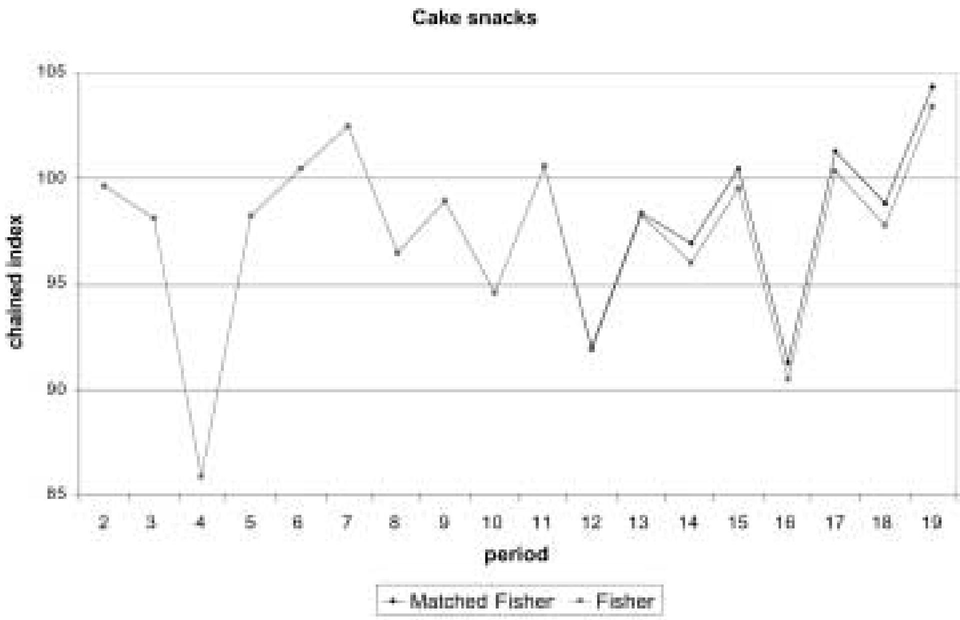
There are some 80 outlets in the sample. To ease computation the sample of outlets is held constant over time; it is thus a panel. In the CPI database, on the other hand, the sample is allowed to change slightly. For each EAN code sold in two adjacent periods a period-to-period unit value index has been calculated over all outlets as the ratio of the period-to-period expenditure index and the quantity index (being the index of the number of scans).

Next, commodity group period-to-period Laspeyres, Paasche, and Fisher price indexes were computed using the unit value indexes of all matched EAN codes belonging to the group in question. No attempt was made to impute fictitious prices of new and disappearing EAN codes. That would have been an impossible task in the case of fast-moving consumer goods. What was done instead was to compute the period-to-period counterparts of Expression (15), which approximates Balk's (2000a) CES-based price index and should thus be an approximation of a cost of living subindex. This index will be referred to as the Fisher price index. The within-group elasticities of substitution were estimated for each time period and according to three different methods following Balk's suggestions; see Opperdoes (2001) for details. Finally, chained indexes for the nine commodity groups were calculated. Figure 1 presents the results. Note that the chained Laspeyres and Paasche price indexes are not shown. As is well known, these indexes exhibit strong upward and downward drift, respectively. To give an example: the chained matched Laspeyres index numbers of baby's napkins and detergents are even well above 200 at the end of the time interval considered.

The difference between the matched Fisher price index and the Fisher price index differs among the nine commodity groups. In eight cases, the Fisher price lies below the matched version. While for some groups the difference is rather small (paticularly for cake snacks, cereals and crisps), for some other commodity groups it cannot be neglected. As a matter of fact, for tea, scents and yoghurt with additives the matched Fisher price index measures a price increase during the entire time interval, whereas the Fisher price index measures a price decrease.

Notice furthermore the volatility of the indexes. At first sight this erratic behaviour may come as a surprise, especially since there are such large amounts of data involved. Discounts, particularly those given to customer cardholders, which are incorporated into the scanner turnover data are the probable cause. Consumers react strongly to discounts. Since discounts rest generally on popular products and because the use of customer cards is extensive (although varying across outlets), the impact of discounts on the commodity group price index numbers can indeed be quite large. Perhaps statistical agencies will be reluctant to accept such volatile price indexes. However, the erratic pattern reflects the use of average transaction prices at the level of individual goods coupled with the use of a superlative index formula to aggregate the goods' price indexes. Hence, it reflects actual consumer behaviour and describes a real phenomenon. The concluding Section 8 contains a short discussion of this topic.

Figure 2 shows the aggregate price change of the nine commodity groups. The period-to-period matched Fisher and Fisher commodity group indexes have been aggregated in two ways: using the Laspeyres formula (leaving expenditure weights fixed at period $t-1$ levels) as well as the Fisher formula. Time series were again obtained through chaining. As might be expected, the volatility of scanner data-based price index numbers diminishes at a higher level of commodity aggregation. Notice that the use of the chained Fisher formula to aggregate the period-to-period matched Fisher commodity group price indexes points to a 2% price decrease between period 19 and period 1, whereas the use of the chained Laspeyres formula points to a 7% price increase. This difference illustrates the upward drift that is found using the latter formula, due to both the (upper level) Laspeyres formula as such and the chaining principle.



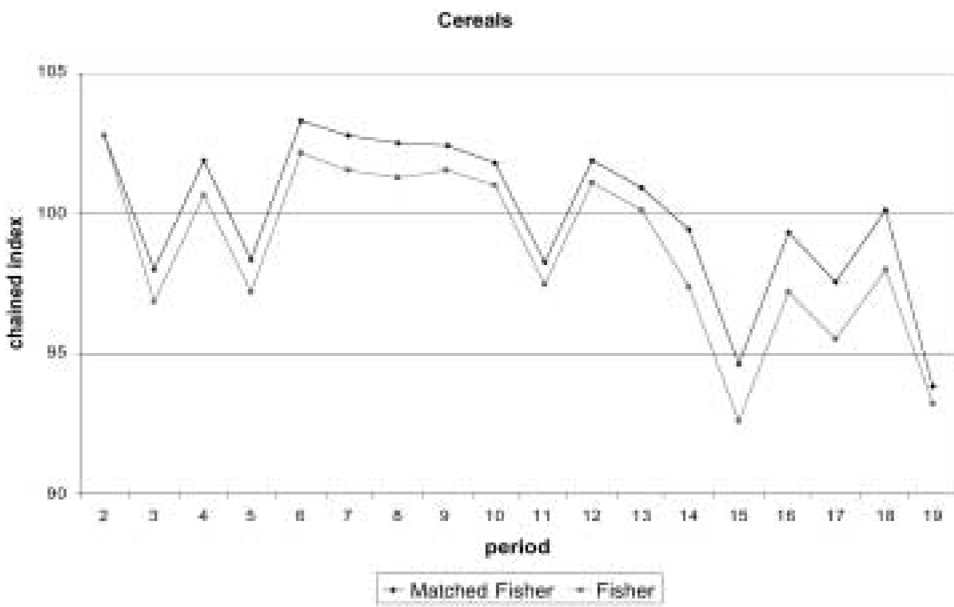
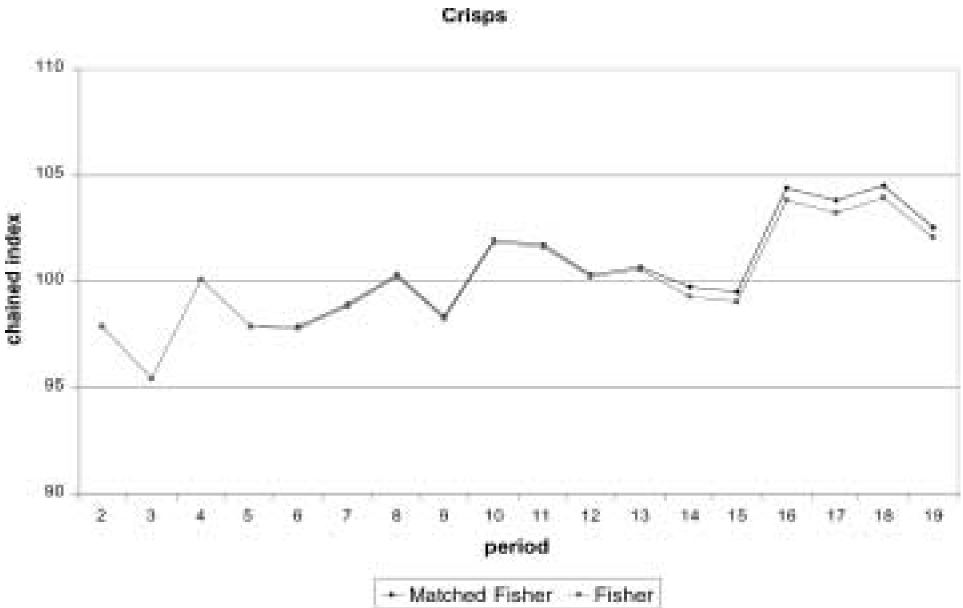


Fig. 1. Chained price indexes for nine commodity groups

Let us now take a closer look at Balk's CES method. As was mentioned in Section 5, he points to the fact that the goods should be classified according to the economic criterion of equal within-group substitution elasticities. That procedure has not been followed here. The individual EAN codes were classified according to a conventional classification scheme. Table 2 contains some statistics about the substitution elasticities and the number of EAN codes per commodity group. The 10% trimmed arithmetic mean of the estimated



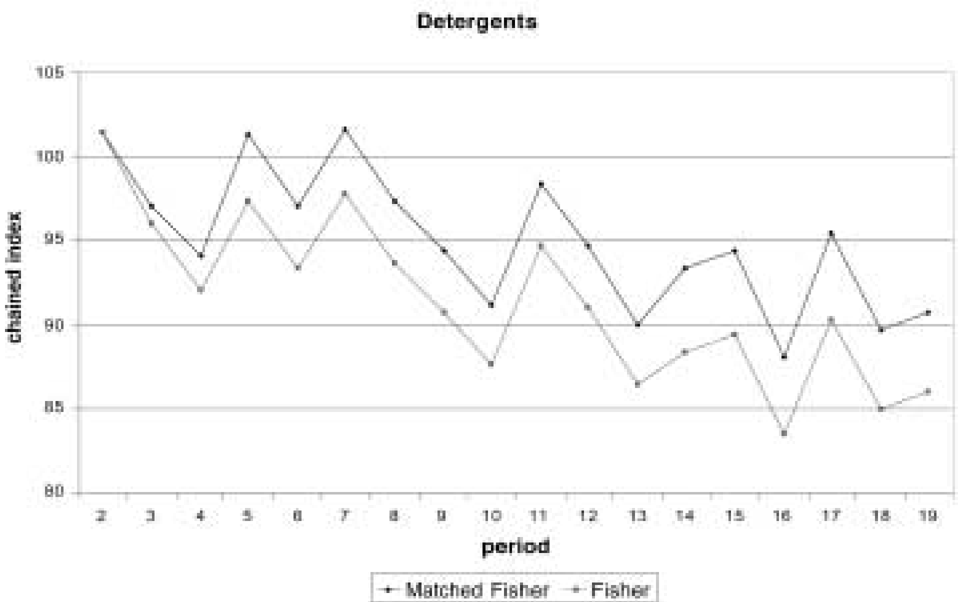
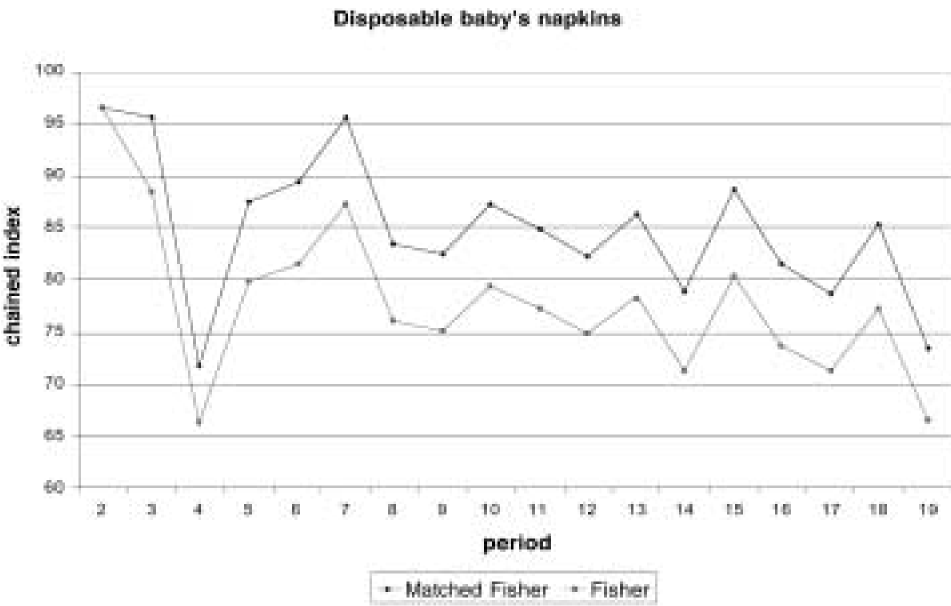


Fig. 1. Continued

substitution elasticities has been used to compute the Fisher price indexes for each commodity group. In accordance with Balk's theory the trimmed means exceed unity, ranging from 2.1 (scents) to 6.4 (baby's napkins). However, for some commodity groups specific values below 1 and even negative values (cereals and scents) were found during certain periods. The large standard deviations suggest that the elasticities are not time-invariant at all. For all nine commodity groups the number of products (EAN codes) varies over time and generally exhibits an upward trend.



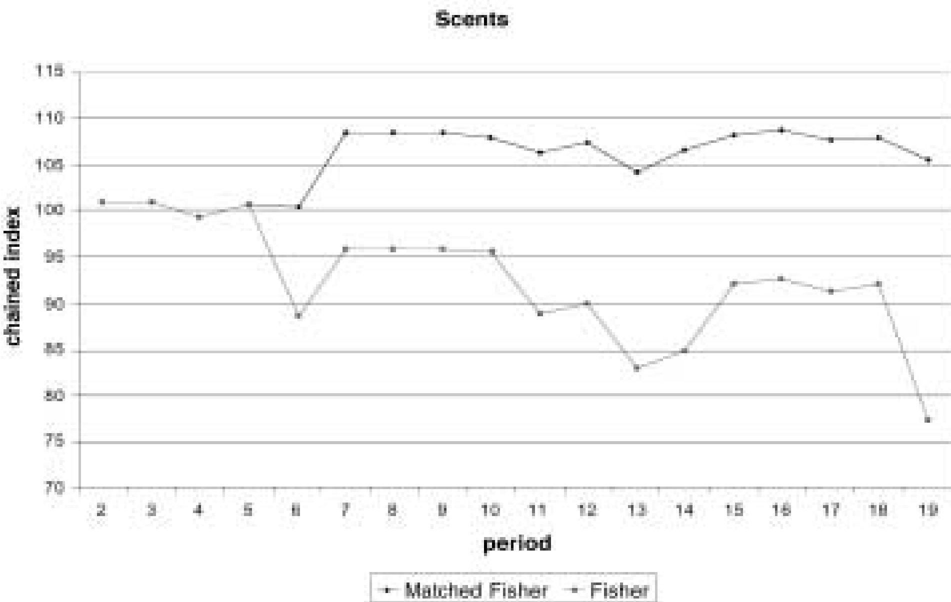


Fig. 1. Continued

The uncertainty about the true value of the elasticities can be particularly a nuisance taking into account the sensitivity of the implicit (''fictitious'') relative price change of new goods to changes in the value of the substitution elasticity shown in the example of Table 1. This raises another question. Is it realistic to assume, as Balk is doing, that a good's substitution elasticity is the same during its introduction and in later time periods? Moreover, it is difficult to understand how Balk's procedure of classifying goods should be carried out in practice. In any case, the procedure would probably bring about a lot of work

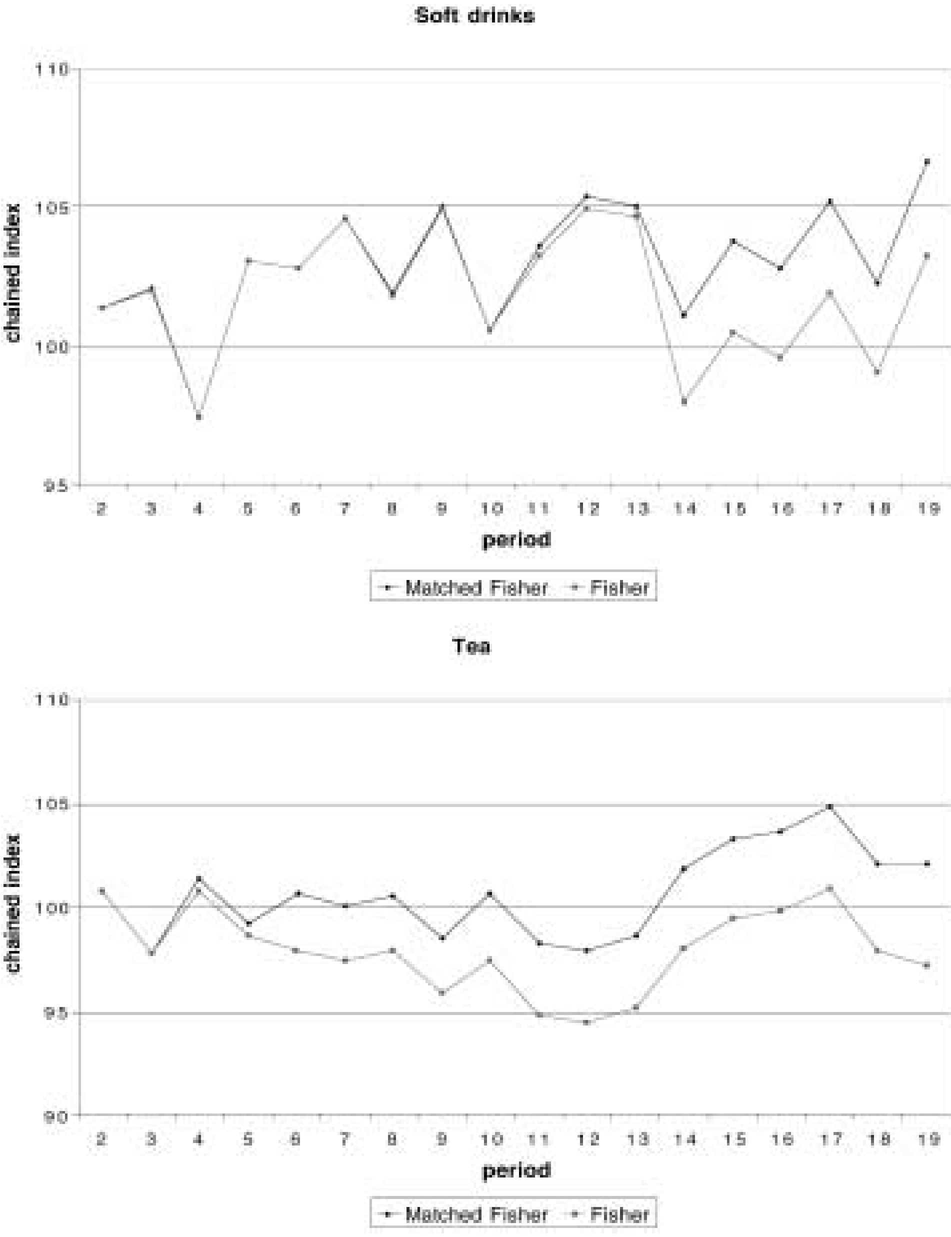


Fig. 1. Continued

for a statistical agency, and its practical feasibility can be doubted. Nevertheless, Figures 1 and 2 remind us that the use of the chained Fisher formula restricted to matched EAN codes can lead to biased figures.

8. Summary and Discussion

CPI statisticians may sometimes get the feeling that they are faced with a paradox. On the one hand they should separate price from quantity changes, and to this end they try to keep

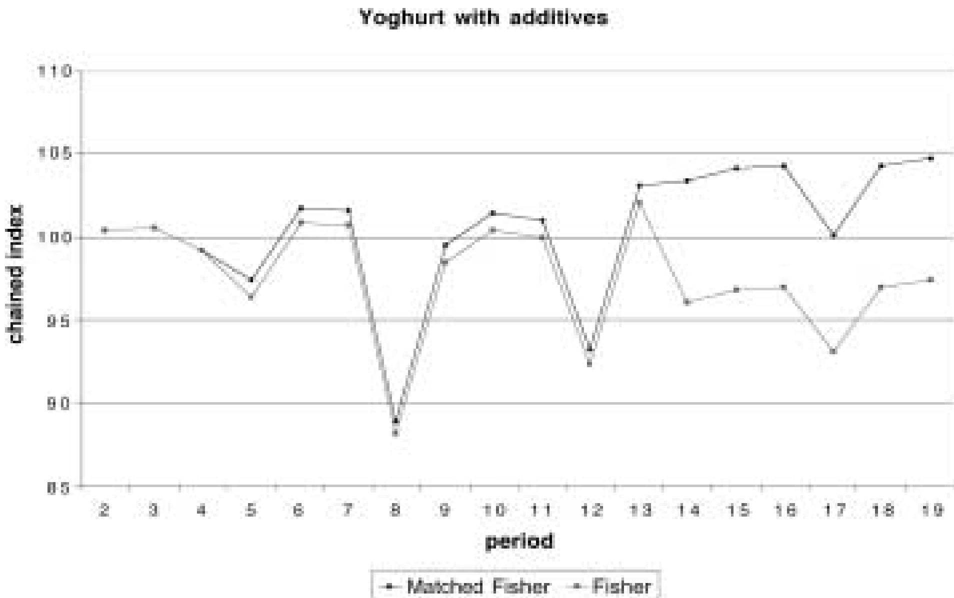


Fig. 1. Continued

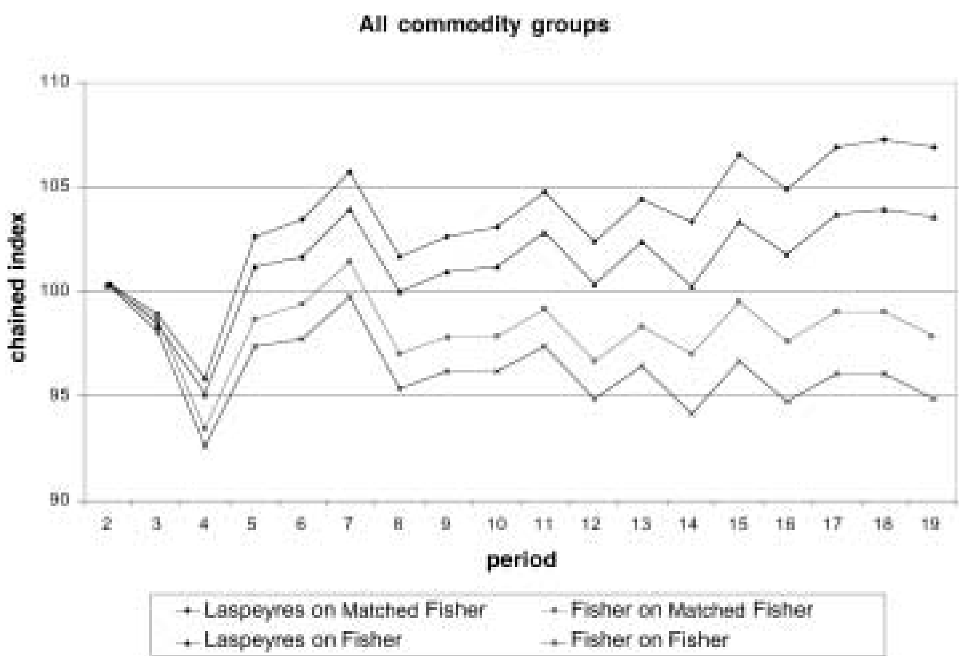


Fig. 2. Aggregate chained price indexes

Table 2. Substitution elasticities and number of EAN codes

Commodity group	Substitution elasticities				Number of EAN codes			
	Mean*)	Stand. dev.	Min.	Max.	Mean	Stand. dev.	Min.	Max.
Cake snacks	4.5	1.2	2.6	6.6	73.5	2.7	70	79
Cereals	2.4	1.9	−5.4	4.0	18.4	0.7	17	19
Crisps	4.4	1.0	2.6	6.7	37.6	3.9	33	43
Detergents	5.0	1.4	2.3	8.4	50.3	3.2	45	56
Disposable baby's napkins	6.4	1.9	1.3	9.2	38.8	6.6	21	45
Scents	2.1	3.0	−3.5	7.3	22.4	1.9	20	26
Soft drinks	3.5	0.6	2.4	4.7	101.3	2.6	98	107
Tea	5.0	1.1	2.6	7.5	138.7	10.6	115	150
Yoghurt with additives	3.9	1.1	1.8	5.4	59.8	5.7	54	70

*) 10% trimmed mean

many things fixed, for example the basket of goods and services. On the other hand they should account for a changing consumption pattern of the representative household. The statisticians are aware of the discrepancy between a static Laspeyres-type CPI and the dynamic world they live in. To overcome this problem, pragmatic choices are being made. When a good of which the prices are collected disappears from the market, another good will be selected and a quality adjustment made to compute the desired ''pure'' price change. When a completely new good appears it will usually be incorporated into the CPI somehow, albeit with a considerable time lag. To account for commodity subsitution at a low level of aggregation, some agencies are using geometric means of price observations.

Since the CPI is a sample statistic, the statistical agency must have a view on the population statistic that the estimator represents. For instance, how should one interpret the Laspeyres price index estimator and the quality adjustments made when the number of goods in the population decreases? Looking at the population Laspeyres price index we can no longer speak of a one-to-one relation between a disappearing good and its natural successor. In this article it has been suggested that quality adjustment methods be viewed as imputation procedures. Extending this idea to the Fisher price index, a generalised Fisher price index was defined on a variable set of goods in which unobservable (fictitious) current periods prices of disappearing goods and base period prices of new goods should be imputed. Note that, even if one is willing to accept the idea of the generalised Fisher price index as the ideal aggregator, there will always be room for controversy about the true value of the index number in a certain time period because different procedures to estimate the fictitious prices will lead to different outcomes.

The generalised Fisher price index rests on a simple symmetry argument: the Laspeyres price index uses forward imputation of disappearing goods' prices and the Paasche price index uses backward imputation of new goods' prices, and consequently their geometric mean uses both. It has been shown that the bias of the Laspeyres price index estimator–that is, the difference between the expected value of the estimator and the true population value of the generalised Fisher price index–can be decomposed into three additive terms: statistical bias attributed to inadequate quality adjustment (the use of incorrect fictitious current period prices of disappearing goods), substitution bias and ''dynamic universe bias.'' The latter term reduces to new goods bias when there are no disappearing goods. Although the generalised Fisher price index was derived without explicit reference to the theory of the cost of living index, these types of bias coincide with those which are usually distinguished when applying that theory.

A substantial part of this article has been devoted to the question under what conditions the generalised Fisher price index can be accurately approximated by the matched-item Fisher price index. The answer depends to a large extent on the expenditure size of the matched part, the prevailing market circumstances and possibly also the nature of the goods in question. Computing monthly-chained indexes instead of direct (bilateral) indexes has the advantage that the size of the matched part does not diminish during the course of time. In addition, monthly-chained price indexes can be viewed as approximations to Divisia price indexes, which are in some sense the ultimate economic price indexes. Empirical work by Silver and Heravi (2001) on scanner data for washing machines seems to indicate that exact matching coupled with monthly-chained superlative price indexes leads to acceptable results. Still, a danger of obtaining biased results, i.e., of overstating or understating the generalised Fisher index, remains. This has been confirmed by our empirical analysis using an approximation of Balk's (2000a) CES approach. A practical problem when it comes to using scanner data is that exact matching by the identifying EAN code ignores a hidden price change of a homogeneous item that has different codes attached to it in different periods. The appendix tries to tackle this problem by introducing a quality-adjusted unit value.

A well-known problem that has not been addressed in the article, and for which a satisfactory solution does not seem to exist, is the presence of seasonal goods. The CPI scanner database contains EAN codes, especially for fresh fruit, which disappear in certain periods and reappear in later periods. Chaining period-to-period indexes is probably not the right choice here, because reappearing EAN codes are treated as new goods. See Schut (2001) for some interesting empirical material. Therefore we selected commodity groups where seasonal patterns do not seem to emerge. Diewert (1999) recommends statistical agencies to construct (at least) three families of consumer price indexes to deal with seasonality. The first index is defined over nonseasonal goods, the second index compares the prices of a certain calendar month with the prices of the same month of the previous year, and the third index is an annual one (built up from the second index), which compares a moving total of twelve months with twelve base year months. Dalén (1999) also notes that seasonal goods cannot be handled well when constructing monthly price indexes. His idea is that one should deal only with years. The ideal index, in his view, should ultimately be based on year-to-year links. The monthly indexes ought to be taken as temporary indicators, later to be replaced by the yearly links.

The price index numbers presented in Section 7 showed an erratic pattern. Feenstra and Shapiro (2001), using scanner data on tuna, also found substantial high-frequency variation in price and substantial response of consumer demand to this variation, suggesting inventory behaviour. In their opinion, a cost of living index must compare all prices over one planning horizon to all prices over another, e.g., compare one year to the next. They argue that the chained Törnqvist index is to be avoided when using high-frequency scanner data, since it gives too much weight to price increases that follow the end of sales and is thus upward biased. A fixed-base Törnqvist index (preferably computed weekly) should be used instead. Looking at our results, however, we cannot find strong evidence of upward bias. Also, the attrition rate of EAN codes for some commodity groups is so high that the calculation of fixed-base indexes–necessarily based entirely on the ongoing goods–soon loses its meaning since the matched part shrinks fast over time. The price index volatility might be viewed as a problem also by those statistical agencies that do not like to think in terms of the cost of living index methodology. They probably look at the CPI as a short-term indicator of the trend in consumer price change and perhaps favour some kind of smoothing procedure. Clearly, there is scope for more research in this area.

Further research could be useful for helping statistical agencies to decide whether or not to use scanner data, particularly on durable consumer goods, in the production of their CPIs. One might, for example, attempt to follow the approach described in this article using hedonic regression to estimate the fictitious prices–provided that the necessary data on the goods' price-determining characteristics are available–and quantify the chained version of decomposition (9). De Haan and Opperdoes (2001b) did this using a large (non-scanner) data set on prices, sales and characteristics of new cars purchased in the Netherlands. They also show how the generalised Fisher price index can be written as a quality-adjusted unit value index when a linear hedonic model is estimated by sales-weighted least squares regression. Only under rather strict assumptions can an ordinary unit value index be called a price index or a cost of living subindex (Balk 1998). But a quality-adjusted unit value index, being the ratio of quality-adjusted unit values pertaining to two time periods, can be given a meaningful interpretation.[6]

Furthermore, one might try to calculate weekly-instead of monthly-chained index numbers. If one does so, the size of the matched part will grow in relative terms, but possibly at the expense of greater variability of the indexes. Sampling aspects can be relevant too. If a statistical agency decides to use scanner data, it may wish to do so on the basis of a sample of items. The question then arises how the sample should be drawn and what the statistical properties are of the chosen estimator.

Appendix

Quality-adjusted unit values

The EAN coding system may be too detailed, causing different codes to represent items that are in fact identical from the consumer's perspective. In that case exact matching by EAN code would ignore any hidden price increase of a homogeneous item whose code has been changed. Although in Section 3 it was argued that the fictitious, imputed prices of disappearing goods can (and should) be viewed in isolation from the fictitious, imputed prices of new goods, this might be a valid argument to link a disappearing good to a new one in a synthetic way–especially when these goods are very close equivalents and one could speak of a disappearing good and its natural successor. In order to reduce possible bias stemming from hidden price increases, the set of exactly matched

[6] Dalén (2001) rightly argues that the use of a quality-adjusted unit value index is an alternative way of aggregating price changes of heterogeneous items. He is even convinced that it has to be the statistical target for those goods where hedonic regression or other well-defined quality adjustment methods are applied. See the Appendix for a different formulation of a quality-adjusted unit value.

EAN codes can be enlarged in a synthetic manner. This is the topic explored below. The text is a slightly amended version of Section 3.4 in De Haan et al. (2000).

Suppose that Good 1 has been sold in period $t-1$; it may be sold for some time during period t also, but we expect it to be sold no longer during period $t+1$. The quantities are thus $q_1^{t-1} > 0$; $q_1^t \geq 0$; and $q_1^{t+1} = 0$. Good 2, which is regarded as the obvious successor of Good 1, was not yet sold in period $t-1$, but is sold in period t (and presumably in period $t+1$ as well). So we have $q_2^{t-1} = 0$; $q_2^t > 0$; and $q_2^{t+1} > 0$.

Suppose next that we can find a quality adjustment factor $\delta_{2/1}^t$ that serves to ''change the quantity bought of Good 2 in period t into a quantity of Good 1.'' It is assumed that the representative consumer attains the same level of satisfaction from the consumption of one unit of Good 2 as from the consumption of $\delta_{2/1}^t$ units of Good 1. Hence, the consumer is indifferent between consuming q_2^t units of Good 2 and consuming $\delta_{2/1}^t q_2^t$ units of Good 1. An average price in period t of Goods 1 and 2 can now simply be computed as

$$p_{2/1}^t = \frac{q_1^t p_1^t + \delta_{2/1}^t q_2^t \tilde{p}_2^t}{q_1^t + \delta_{2/1}^t q_2^t} = \frac{q_1^t p_1^t + q_2^t p_2^t}{q_1^t + \delta_{2/1}^t q_2^t} \tag{22}$$

where $\tilde{p}_2^t = p_2^t / \delta_{2/1}^t$ might be called the quality-adjusted price of Good 2; hence, $p_{2/1}^t$ might be referred to as a quality-adjusted unit value. To calculate the Fisher price index going from period $t-1$ to period t we should use the prices p_1^{t-1} and $p_{2/1}^t$, and the quantities q_1^{t-1} and $q_1^t + \delta_{2/1}^t q_2^t$. For the index going from t to $t+1$ we should use the prices $p_{2/1}^t$ and p_2^{t+1}, and the quantities $q_1^t + \delta_{2/1}^t q_2^t$ and q_2^{t+1}.

Expression (22) reduces to an ordinary unit value when $\delta_{2/1}^t = 1$. Such a situation can arise for example in the case of a failed match, i.e., when an EAN-code changes while the goods in question are identical from the consumer's perspective–particularly when they are the same in physical terms. Notice that $p_{2/1}^t = p_1^t$ if $\tilde{p}_2^t = p_1^t$, that is when $\delta_{2/1}^t = p_2^t / p_1^t$, which implies the use of overlap pricing. If instead a hidden price increase occurs, so that $\tilde{p}_2^t > p_1^t$, then we have $p_{2/1}^t > p_1^t$. In a perfect world the price difference between close substitutes would reflect the consumer's evaluation of the quality difference when both goods are available at the same time, and overlap pricing would be the proper quality adjustment method. Consequently, a hidden price increase can only occur when the market is not ''in equilibrium'' or when consumers are ill informed about the difference between the old and the new variety. Notice further that the average transaction price p_1^t is not defined in the case of $q_1^t = 0$. Taking $p_{2/1}^t = \tilde{p}_2^t$, which results from setting $q_1^t = 0$ in (22) had p_1^t been defined, is nevertheless the right solution.

Determining the quality adjustment factors δ^t might be rather difficult in practice because of the limited product descriptions generally available in scanner data sets. There will most likely be a need for collecting additional information on the goods' characteristics. For the sake of timeliness it could be useful to restrict such actions to those disappearing goods with natural successors whose fractions of the commodity group's turnover in period $t-1$ exceed some threshold value–say 5 or 10%.

9. References

Aizcorbe, A., Corrado, C., and Doms, M. (2000). Constructing Price and Quantity Indexes for High Technology Goods. Paper presented at the NBER Summer Institute Session on Price, Output and Productivity Measurement, Cambridge, 31 July.

Balk, B.M. (1998). On the Use of Unit Value Indices as Consumer Price Sub-indices. In Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W. Lane), 112–120. Washington D.C.: Bureau of Labor Statistics.

Balk, B.M. (2000a). On Curing the CPI's Substitution and New Goods Bias. Research paper no. 0005. Voorburg: Statistics Netherlands.

Balk, B.M. (2000b). Divisia Price and Quantity Indices: 75 Years After. Draft Research Paper. Voorburg: Statistics Netherlands.

Boskin, M.J., Dulberger, E.R., Gordon, R.J., Griliches, Z., and Jorgenson, D.W. (1996). Toward a More Accurate Measure of the Cost of Living. Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index.

Bradley, R., Cook, B., Leaver, S.G., and Moulton, B.R. (1998). An Overview of Potential Uses of Scanner Data in the U.S. CPI. Proceedings of the Third Meeting of the International Working Group on Price Indices (ed. B.M. Balk), 169–182. Voorburg: Statistics Netherlands.

Dalén, J. (1998a). Experiments with Swedish Scanner Data. In Proceedings of the Third Meeting of the International Working Group on Price Indices (ed. B.M. Balk), 163–168. Voorburg: Statistics Netherlands.

Dalén, J. (1998b). On the Statistical Objective of a Laspeyres' Price Index. Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W. Lane). Washington D.C.: Bureau of Labor Statistics.

Dalén, J. (1999). A Proposal for a New System of Aggregation in the Swedish Consumer Price Index. Paper presented at the Fifth Meeting of the International Working Group on Price Indices, Reykjavik, 25–27 August.

Dalén, J. (2001). Statistical Targets for Price Indexes in Dynamic Universes. Paper presented at the Sixth Meeting of the International Working Group on Price Indices, Canberra, 2–6 April.

Diewert, W.E. (1976). Exact and Superlative Index Numbers. Journal of Econometrics, 4, 115–145.

Diewert, W.E. (1992). Fisher Ideal Output, Input and Productivity Indexes Revisited. Journal of Productivity Analysis, 3, 211–248.

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

Diewert, W.E. (2001). The Consumer Price Index and Index Number Purpose. Journal of Economic and Social Measurement (Forthcoming).

Feenstra, R.C. and Shapiro, M.D. (2001). High-Frequency Substitution and the Measurement of Price Indexes. In Scanner Data and Price Indexes, Studies in Income and Wealth, Vol. 61, NBER (eds. R.C. Feenstra and M.D. Shapiro). Chicago: University of Chicago Press (Forthcoming).

Fenwick, D., Ball, A., Silver, M., and Morgan, P.H. (2001). Price Collection and Quality Assurance of Item Sampling in the Retail Price Index: How Can Scanner Data Help? In Scanner Data and Price Indexes, Studies in Income and Wealth, Vol. 61, NBER (eds. R.C. Feenstra and M.D. Shapiro). Chicago: University of Chicago Press (Forthcoming).

Haan J. de, Opperdoes, E., and Schut, C.M. (1999). Item Selection in the Consumer Price Index: Cut-off Versus Probability Sampling. Survey Methodology, 25, 31–41.

Haan, J. de, Poppel, P. van, and Spaanderman, C. (2000). Using Scanner Data in the Compilation of the CPI: A Dutch Pilot Study. Research paper, Voorburg: Statistics Netherlands.

Haan, J. de, and Opperdoes, E. (2001a). Geometric Average Prices in the CPI (Meetkundig gemiddelde prijzen in de CPI, written in Dutch). Research paper, Voorburg: Statistics Netherlands.

Haan, J. de, and Opperdoes, E. (2001b). Hedonic and Matched-Model Fisher Price Indexes for Cars. Draft Research paper, Voorburg: Statistics Netherlands.

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

Hawkes, W. and Smith, R. (1999). Improving Consumer Price Measurements Through the Use of Scanner Data and Market Segmentation. Proceedings of the Measurement of Inflation Conference (eds. M. Silver and D. Fenwick), 283–299, Eurostat/ONS/Cardiff University: Cardiff Business School.

Hill, P. (1999). Inflation, the Cost of Living and the Domain of a Consumer Price Index. Paper presented at the Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 3–5 November.

Ioannidis, C. and Silver, M. (1999). Estimating Exact Hedonic Indexes: An Application to U.K. Television Sets. Journal of Economics, Zeitschrift für Nationalökonomie, 69, 71–94.

Lowe, R. (1998). Televisions: Quality Changes and Scanner Data. Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W. Lane), 5–20. Washington D.C.: Bureau of Labor Statistics.

Lowe, R. and Ruscher, C. (2001). Estimating Price Movements for Consumer Durables Using Electronic Retail Transactions Data: An Empirical Study on Television. In Scanner Data and Price Indexes, Studies in Income and Wealth, Vol. 61, NBER (eds. R.C. Feenstra and M.D. Shapiro). Chicago: University of Chicago Press (Forthcoming).

Moulton, B., LaFleur, T.J., and Moses, K.E. (1998). Research on Improved Quality Adjustments in the CPI: The Case of Televisions. Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W. Lane), 77–99. Washington D.C.: Bureau of Labor Statistics.

Opperdoes, E. (2001). Some Empirical Experiments with CES Functions. Research paper, Voorburg: Statistics Netherlands.

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

Richardson, D. (2001). Scanner Data for the CPI. In Scanner Data and Price Indexes, Studies in Income and Wealth, Vol. 61, NBER (eds. R.C. Feenstra and M.D. Shapiro). Chicago: University of Chicago Press (Forthcoming).

Schut, C.M. (2001). Using Scanner Data to Compile Price Indices: Experiences and Practical Problems. Paper presented at the Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 1–2 November.

Silver, M.S. (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. and Heravi, S. (2001). The Measurement of Quality-Adjusted Price Changes. In Scanner Data and Price Indexes, Studies in Income and Wealth, Vol. 61, NBER (eds. R.C. Feenstra and M.D. Shapiro). Chicago: University of Chicago Press (Forthcoming).
Triplett, J.E. (1999). Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index? The Economic Journal, 111, 311–334.

Received May 2001
Revised January 2002