# Segmented markets and CPI elementary classifications

Timo Koskimäki[1] and Mari Ylä-Jarkko[2] [3]

Statistics Finland

*Special thanks to Eugen Koev and Antti Suoperä for insightsfull comments. Professor Yrjö Vartia also provided valuable ideas in the initial stages of the research project.*

**Abstract:** A traditional consumer price index (CPI) is typically constructed using - sometimes very limited - samples of transactions. Also, the standard CPI typically keeps the quality of the products to be compared constant by imposing on the products to be compared rather rigid classificatory rules. These rules determine, which kind of products should be treated as comparable and which not. The availability of more exhaustive scanner data sets opens up an opportunity not only to compile indices utilising more exhaustive data but an opportunity to check the validity of the standard CPI sampling- and replacement practices.

In the paper the performance of traditional CPI sampling, matching and replacement rules are evaluated using an extensive scanner data set as a simulation platform. The structure of the study is as follows:

First, the methodology of a traditional[4] CPI is replicated in every detail using a representative micro-level scanner data set. The results of the scanner- and traditional price indices are compared.

Then, alternative classification schemes that allow for varying degree of substitution between varieties offered by different manufacturers, product brands, outlets and different regions are constructed and applied to the scanner data. The results are then interpreted in the light of traditional CPI sampling and replacement solutions.

The type of data used does not seem to be of immense importance for the price change comparisons. When the CPI practices were replicated using scanner data and compared with the official CPI, the indices were fairly close to each other.

The choice of classification and replacement rules (whether the products belonging to the same product class are deemed directly comparable or not) do, however, affect the results of the index calculus quite significantly also in the case of rather simple consumer goods.

The problem of quality change has in the research literature mainly been discussed in the context of products where rapid technological change is apparent. On the basis of the results it is apparent that "quality change" issues should be considered also in the case of simple commodities. However, the nature of quality change for simple products is somewhat

[1] Timo Koskimäki wrote the first draft for the abstract, chapters 2 and 3.
[2] Mari Ylä-Jarkko wrote the first draft for chapter 1.
[3] In addition to the authors, Juha Mylläri and Antti Katainen have been active members of the research team.
[4] In this case the Finnish national Consumer Price index.

different. It is mainly a question of differentiated marketing strategies, i.e. manufacturers attempt to present their own products as distinct as possible. The idea is to try to maintain brand-based segmented micro-markets where substitution between the products should happen within the brand, not between brands. If the segmentation strategy is successful, the manufacturer can adopt quite profitable pricing schemes and plan the obsolescence of their products. The question of how price statisticians should react to this kind of situation in principle is somewhat difficult. However, in practical data collection situation this decision is taken when determining the "sameness" of the products to be priced.

As there do not exist any self-evident statistical rules on how to deal with different types of classification and comparability issues at the most detailed level of the CPI's , measures should be taken to agree on some generally accepted best practices for types of commodities and market situations highlighted in this study

## 1. The traditional CPI compiled using scanner data

### *1.1 The data*

The data used in this study was provided by ACNielsen Ltd. The "elementary transaction" in the data is sales of a specific product during one week in a single outlet. "Specific product" is identified as a distinct EAN code. Data relating to one elementary transaction are as follows:

- number of packages sold
- package size
- unit of the package (Kilogram, Litre, Ml et.c.)
- total sales during the week
- Product identification code ("EAN")
- brief description of the product
- manufacturer
- ACN's own product class
- "Brand"
- outlet code
- region

For the purpose of index calculations a simple package size adjustment was performed. The elementary price used in all the examples is the "weekly mean price" of one unit sold.

The products covered in this study are typical "fast moving consumer goods" where no rapid technical quality changes is expected to occur. The data covers the following product groups: Butter, Margarine and other vegetable fats, Vegetable oils, Soft drinks, Fruit juices and Detergents.

### *1.2 The CPI replication approach*

In this part of the study, we identify, in the scanner data set, the same products that are being priced in the Finnish national CPI. Indices applying same type of replacement rules as in the current CPI but using the scanner data will be calculated and the results compared with the official CPI results.

The purpose of this part is to examine whether there are differences between indices based on scanner data and the official indices based on standard CPI data collection which stem from

the differences in basic data. To assess this issue is important from the point of view of generalisation of the results of the other analyses.

From the data only the varieties that corresponded to the standard CPI quality descriptions were taken into the compilation of the ACN -index. From every shop the most sold variety in year 1998 that was available in January 1998 was chosen for monitoring. If there were no observations in the outlet in January 1998 the outlet was excluded from the compilation. Only the outlets that stayed in the sample over the three years of monitoring were selected. This resulted to good 250 outlets, which is slightly more than there are in the standard CPI price collection.

If a variety was not available during the monitored month, the price change has been imputed using the mean price change for all the varieties belonging to the same class as proxy information. In the same manner, when a new variety was selected for pricing, the same procedure was used to construct a new base price. In essence, new varieties were chained into the index assuming that the price development has been the same that was observed for varieties where price information for the two periods was available.

The product groups that matched with standard CPI product descriptions are presented in table 1. When aggregating the sub-indices the standard CPI weights were used also for the scanner data.

**Table 1: Sub-indices and their weights**

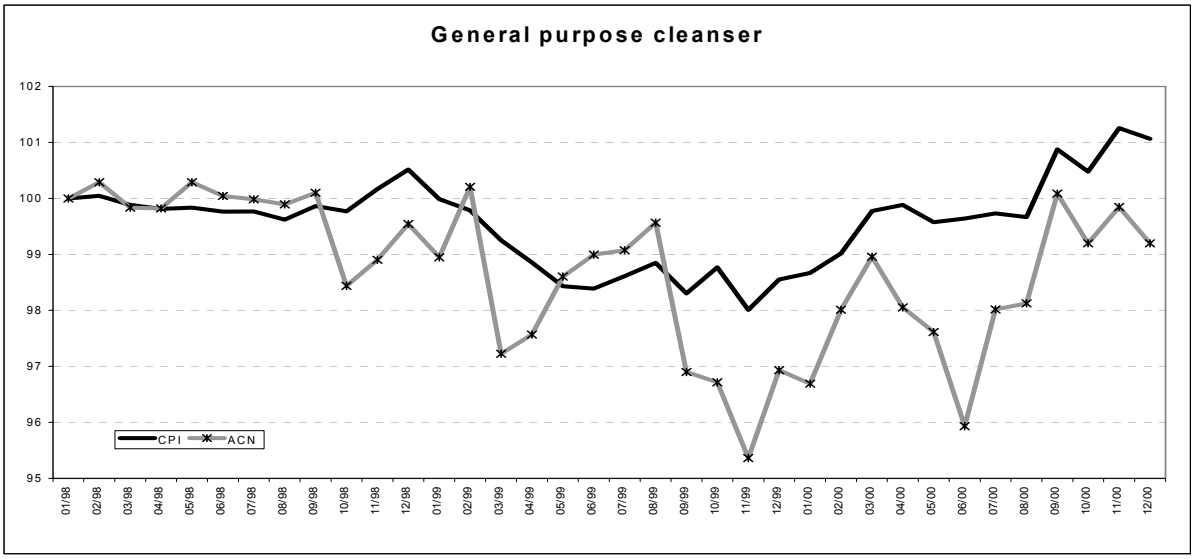
| **Total Index** | | **100,00** |
|---|---|---|
| **COICOP** | | |
| 01.1.5.1.01 | Dairy butter | 14,47 |
| 01.1.5.2.01 | Butter and vegetable oil mixture | 10,72 |
| 01.1.5.2.02 | Cooking margarine | 3,25 |
| 01.1.5.2.03 | Soft margarine | 14,81 |
| 01.1.5.2.04 | Low fat margarine | 1,17 |
| 01.1.5.4.01 | Vegetable oil | 3,22 |
| 01.2.2.2.01 | Vegetable extract drinks cola | 16,18 |
| 01.2.2.3.01 | Mixed fruit cordial | 1,08 |
| 05.6.1.1.02 | Synthetic detergent | 18,34 |
| 05.6.1.1.03 | Dish washing liquid | 5,24 |
| 05.6.1.1.04 | General purpose cleanser | 11,53 |

Despite the attempt to replicate the standard CPI in every detail, some important differences do remain. Probably the most important difference is the fact that ACN prices are weekly average prices rather than daily on the spot -prices collected for the CPI. This means, among other things, that (nation-wide) sales campaigns, when the variety on sale is purchased in large quantities, clearly show up in the ACN price development. The feature is amplified because of the fact that CPI data is typically not collected during weekends which is the typical time for campaigns. The other factor causing differences is the slightly different treatment of loyal customer sales prices and other targeted campaigns. These types of sales pricing schemes are not covered by the traditional CPI but are covered in ACN data.

### *1.3 Results (I)*

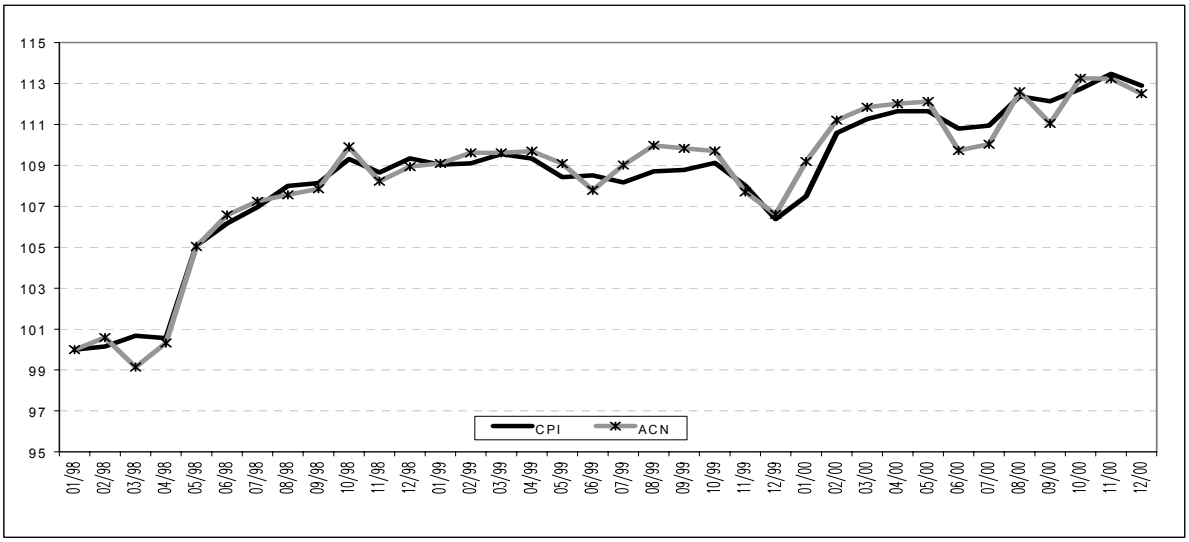
In picture 1 the indices for group "General purpose cleanser" are presented. The ACN index is clearly more volatile. This volatility is most probably due to the differences in the price concept in the two data sources.

**Picture 1: Indices of general purpose cleanser. The current CPI methodology, CPI -data and ACN -data.**



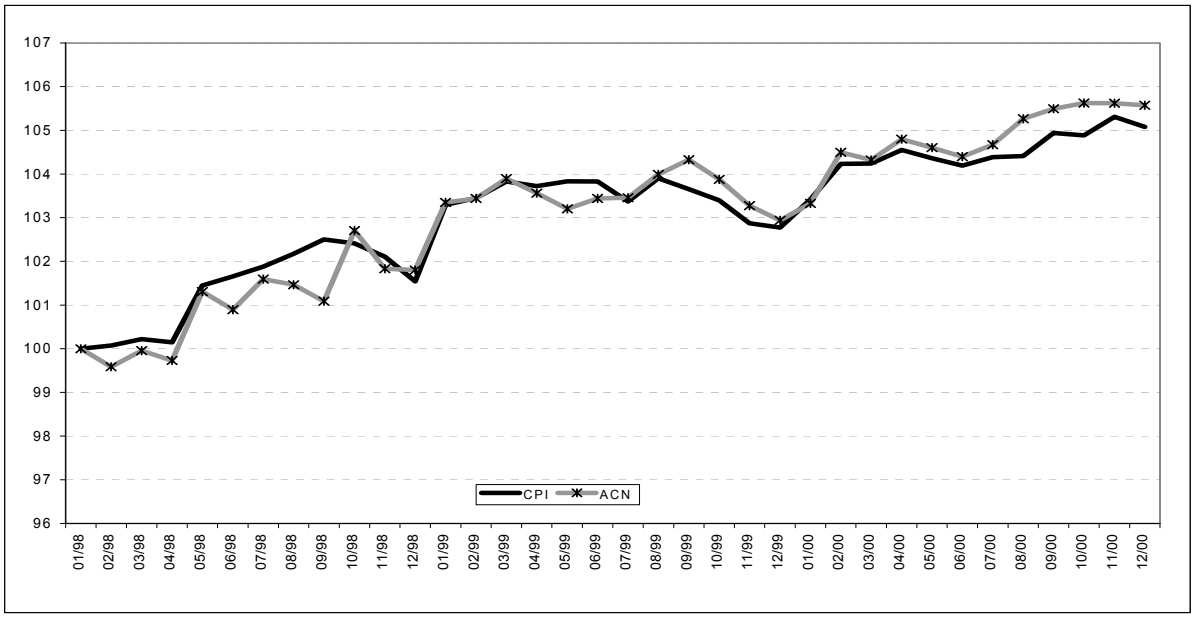

However, for other product groups the differences are relatively small. Picture 2 shows price indices for "Dish Washing liquid". Here the differences are marginal.

**Picture 2: Indices of dish washing liquids. The current CPI methodology, CPI -data and ACN -data.**



When aggregating all compiled sub-indices into a total indices, the results are reasonably close to each other. Total indices are presented in picture 3. After three years of surveillance the indices differ only by 0,5 percentage points. The largest difference is 1,41 percentage points. Generally, we conclude that both of the data-sets give the same result when same type of sampling and replacement practices are applied.

**Picture 3: Total indices 1998:01 = 100 using the current CPI methodology.**




As is apparent from above, the raw mean prices calculated from ACN data do differ from mean prices calculated from CPI data. Annex 3 gives some examples on differences. The general trend is somewhat unclear, in some of the cases studied, the mean prices are reasonably close to each other, whereas for certain product groups the difference is huge. Some of the reasons – campaigns (impacts of sales volumes) and also loyal customer rebates etc. have already been pointed out. In addition, in this study it was not possible to match CPI data with ACN-data at the outlet level. Hence, part of the observed difference may stem from differences in the outlet sample.

It is however quite clear that (raw) mean prices collected using the standard CPI methodology do, in general, differ from each other although the indices compiled from the same data are not biased.

As a consequence, mean prices from the traditional CPI data collection should not be compared directly with the mean prices from cash register systems. Such a comparison will most probably be seriously biased.

## 2. CPI elementary classifications

In chapter 1 we compiled the CPI using the scanner data but applying the traditional CPI sampling and matching rules. As we have access to total data, there is in principle no need to

use sampling. Instead, it is natural to construct a form of index compilation strategy that utilises a maximum amount of the data available.

This part of the study presents two possible strategies that may be applied when using scanner-type data in CPI compilation. We denote the two strategies as the **matching approach** and as the **classification approach**. Both of the strategies will be illuminated by empirical analysis of a geographically representative scanner data set covering some 350 outlets and 8 product categories described in chapter 1. Advantages and disadvantages of the approaches will be discussed as well as the conclusions from the point of view of the traditional CPI practices.

### *2.1 Matching, classification and unit value indices*

The most obvious strategy to deal with the excessively detailed scanner data is to treat the data in a similar manner as the data is treated in current CPI compilation, i.e. one-to-one matching of the observations at the most detailed level, within outlets, across time. Indeed, most of the current studies utilising scanner data seem to take one-to one matching as a self-evident starting point without reflecting on other possibilities.

The number of products to be included in the "scanner-CPI" would be essentially larger than in the traditional CPI. However, the products would still be matched in a very detailed level and the non-matches would be omitted from the compilation. In practice, this means matching of the products on the basis of the most detailed product codes available (in Europe the so called EAN-codes) on the level of single outlets.

From a more general perspective the matched models approach can be characterised as an extremely detailed classification system that has been formed by combining product- and outlet typologies. On the most disaggregated level of index compilation the product classification to be used is often very narrow. Products at the elementary aggregate level are typically classified by package size, manufacturer of the product, flavour or colour, specific brand name etc.

Once we have realised that the traditional one-to one matching is a form of classification we may also consider alternative ways of classifying the data. In essence, we may choose - at our will - the level the detail to be used in the formation of elementary aggregates for the index compilation. The choice of the level of classification should, of course, be justified by some form of rational reasoning (see, e.g. Durkheim and Mauss 1963/1903).

Generally, the term used for this type of approach is "unit value index". The use of unit value indices as elementary aggregates may be perceived as somewhat peculiar (Balk 1998). Diewert (1995, ref Balk 1998), however, seem to give some support for the general idea. Unit values within homogeneous sub-groups are also the standard procedure proposed by some less recent textbooks (see e.g. Allen, 1975). As will be demonstrated, there always exists a multitude of alternatives to form "unit value indices". To define the borderline between "unit value index" and "price index" is always to certain extent an arbitrary choice.

"Price indexes, almost universally, have followed one fundamental methodological principle: The quality of the product sets to be compared is held constant by following the matched pairs strategy. The price index compiling agency chooses a sample of retail outlets or sellers and a sample of products. It collects an initial period, or base period, price for each of the products selected. It then collects at some later date the price for exactly the same product, from the

same seller, that was selected in the initial period. The price index is computed by matching, observation by observation, the price at the later period with the initial price (Tripplet, 2000)".

This kind of classification strategy, admittedly, creates very homogenous sets of observations. Triplett, again, summarises the advantages of one-to-one matching quite neatly:

"The great advantages of this matching methodology are sometimes not explicitly stated, and other times not fully appreciated. The "matched model" methodology holds constant many price determining factors that are usually not directly observable. Examples are characteristics of the retailer, such as customer service, reputation of the manufacturer etc. Matching the price quotes model by model (and outlet by outlet) is not just a methodology for holding quality change constant in the items selected for pricing. It is also a methodology for holding constant non-observable aspects of the transaction that might bias the measure of price change.

The disadvantages of the matching approach are quite obvious although seldom explicitly stated. The most general formulation of the drawbacks is that the matched models approach always causes quite a considerable loss of information. This loss of information stems from the simple fact that new and disappearing varieties can not be taken into account in the index calculation.

Quite a number of contemporary problems in the field of price indices are derivatives of this general feature of the matched models approach. The need for specific "quality adjustment" procedures in cases where a variety that has been followed in the CPI disappears from the markets and has to be replaced by another variety is one class of problems. The other class of problems stems from appearance of "new products". As there is no natural pair for comparison, the new product is generally omitted from the calculation.

A third class of problems stem from the practicalities of the matched pairs -approach. It is often the case that price collectors are advised to follow the initially selected variety until it entirely disappears from the markets. This, in turn, leads rapidly to a situation where the CPI sample is no longer representative, i.e. gives a distorted picture of the markets (see Koskimäki and Vartia 2001, Silver 2001).

The main drawback of the classification strategy is often denoted as "the unit value bias". In essence this term refers to a situation where the elementary cells - or elementary classes created using pre-specified classification rules - are not homogenous enough. The idea of comparing "like with the like" may thus be violated and differences in quality might appear as differences in price. The important question in this study is, however, how should we interpret the notion "like with like".

### *2.2 The classification approach*

For the purpose of this study, two weekly data sets - last week of September 1998 and last week of September 2000 were extracted from the material described in chapter 1.

In addition, products were classified in the spirit of the standard classification of individual consumption by purpose (COICOP). The version of COICOP classification used is presented in annex 1. Also brand- and EAN- classifications provided by ACNielsen were used. The number of classes in each level of the product- and geographic classifications used in the index calculations are presented in table 2 below:

**Table 2: Classifications used in the study**

(Number of distinct product codes (EAN) refers to 1998 data)

| Regional dimension | Number of levels | Product dimension | Number of levels |
|---|---|---|---|
| Whole country | 1 | COICOP 5-digit | 6 |
| Province | 4 | COICOP 7-digit | 26 |
| ACN region | 15 | ACNielsen brand | 266 |
| Outlet | 338 | ACNielsen EAN | 1 028 |

The regional and product dimensions were then combined to form a typology - or a matrix - to be used in the compilation of indices. For each cell in the matrix an index with differing weighting structure were calculated. The typology is shown in table 3.

**Table 3: Typology of different classification strategies**

| Regional dimension | Product dimension: COICOP 5-digit | COICOP 7-digit | ACNielsen Brand | ACNielsen EAN |
|---|---|---|---|---|
| Whole country | 1. | 2. | 3. | 4. |
| Province | 5. | 6. | 7. | 8. |
| ACN region | 9. | 10. | 11. | 12. |
| Outlet | 13. | 14. | 15. | 16. |

The idea of the approach is as follows: The intersection of row and column attributes describes the elementary aggregate level of the index to be calculated. Hence, in the index to be calculated for cell number one we only fix weights for the six COICOP 5-digit groups. In cell 5 we fix the weights also according to the province (4 classes in regional dimension and six classes product dimension, i.e. 24 classes.).

In the lower right corner (cell 16) all available data has been fixed to form an entirely fixed weighting structure (338 outlets and 1028 distinct products). Theoretically, if all outlets would sell all products - and stayed in the markets over the entire period of the study - this would mean 347 464 fixed cells.

Turning back to the two basic approaches, the rightmost cells in the matrix (4, 8, 12, 16) are variations of the matching approach. The rest of the cells are examples of different classification strategies.

From the point of view of traditional CPI practices, the design of the cell structure can be interpreted as a set of distinct replacement rules. In the cell 1 we would consider - in a practical price collection situation – any product within the Coicop 5-digit class as directly comparable (say, vegetable oil is vegetable oil) whereas in cell 16 we would require also the

type of oil, exact package size, manufacturer and product line (brand) to be the same before we consider the products to be the same product.

The elementary aggregates below the given stratification levels have been calculated as weighted arithmetic means. The aim is to produce a meaningful mean price taking into account the fact that the market shares of the products under study vary from 0,5 per mill to some 10 ten per cent within a given product category.

From the point of view of traditional CPI practices, the way to construct the elementary prices within aggregates can be interpreted as a re-sampling rule that would advise the price collector performing outlet-level sampling always to select the most sold variety of the product group to be priced.

In essence, elementary aggregates used here are analogous to the price concept inherent in the scanner data on the outlet level. The weights used in the construction of elementary aggregates are allowed to vary between the two periods concerned. This is consistent with the general idea that the elementary aggregates under study are considered to be homogenous, at least from a consumer's point of view.

### *2.3 Results (II)*

The results of the exercise are shown in tables 4 to 7. To summarise:

- Increasing the level of detail (of the elementary classification) in the product dimension tends to lower the observed price increase. Using 6-class structure in the product dimension yields a price increase of 7,9 percent whereas fixing the weights on EAN-level indicates only 2,3 per cent price increase.

- Increasing the level of detail in regional dimension, especially if the calculus is fixed at the outlet level, tends to increase the observed price increase. The phenomenon, however, disappears if product dimension is tightly fixed.

- Classification which keeps the producer - but not the EAN - fixed (ACNielsen Brand - classification, 266 classes) tends to give higher price increase when compared to a reasonable consumer-oriented classification (COICOP 7-digit, 26 classes).

- Tight classification - both in product- and outlet- dimensions - tends to increase upper level substitution bias (measured as difference between Laspeyres and Fisher indices).

- The loss of information increases rapidly when classifications get more detailed. Keeping the outlet sample fixed between the two periods decreases the coverage by 10 per cent. The most detailed product classification enables only 80 per cent of the transactions to be included in the comparison. The joint effect - keeping both the products and outlets extremely fixed - excludes almost 40 per cent of the data from the index calculation.

**Table 4: Laspeyres price indices September 1998 - September 2000 (sales during one week)**

| Regional dimension | Product dimension | | | |
|---|---|---|---|---|
| | COICOP 5-digit | COICOP 7-digit | ACNielsen Brand | ACNielsen EAN |
| Whole country | 107,9 | 103,1 | 104,6 | 102,3 |
| Province | 107,8 | 103,1 | 104,8 | 102,3 |
| ACN region | 107,8 | 103,1 | 105,1 | 102,5 |
| Outlet | 108,6 | 104,0 | 106,0 | 102,8 |

**Table 5: Fisher price indices September 1998 - September 2000 (sales during one week)**

| Regional dimension | Product dimension | | | |
|---|---|---|---|---|
| | COICOP 5-digit | COICOP 7-digit | ACNielsen Brand | ACNielsen EAN |
| Whole country | 108,0 | 103,2 | 104,8 | 101,5 |
| Province | 107,9 | 103,1 | 104,8 | 101,4 |
| ACN region | 107,9 | 103,0 | 104,7 | 101,4 |
| Outlet | 108,9 | 103,4 | 104,9 | 101,1 |

**Table 6: Difference Laspeyres – Fisher**

| Regional dimension | Product dimension | | | |
|---|---|---|---|---|
| | COICOP 5-digit | COICOP 7-digit | ACNielsen Brand | ACNielsen EAN |
| Whole country | -0,1 | -0,1 | -0,2 | 0,8 |
| Province | -0,1 | 0,0 | 0,0 | 0,9 |
| ACN region | -0,1 | 0,1 | 0,4 | 1,1 |
| Outlet | -0,3 | 0,6 | 1,1 | 1,7 |

**Table 7:Turnover covered by each calculation (Whole country, COICOP 5-digit = 100)**

| Regional dimension | Product dimension | | | |
|---|---|---|---|---|
| | COICOP 5-digit | COICOP 7-digit | ACNielsen Brand | ACNielsen EAN |
| Whole country | 100,0 | 100,0 | 98,2 | 80,1 |
| Province | 100,0 | 100,0 | 97,5 | 77,4 |
| ACN region | 100,0 | 100,0 | 96,9 | 75,5 |
| Outlet | 90,4 | 90,4 | 84,6 | 61,7 |

(period-specific weights)

## 3. Conclusions

On the basis of the results of the CPI replication presented in chapter 1, it is evident, that the type of data used for CPI's is not of immense importance. However, the way data is compiled during the index calculus is clearly of importance.

It is relatively easy to figure out behavioural explanations for the results presented in chapter 2. If we keep "the producer" (ACN brand-classification above) fixed, we seem to exclude from the index price effects caused by new producers or brands entering the markets.

More important, when keeping the brand classification constant, we technically do not allow consumers' substitution between different brands in connection with brand-specific price changes or changes in consumption pattern. I.e. when, within a brand, a new variety of a product is introduced at a higher price we in this design assume that consumers only substitute between the older versions of the same brand (if these still exist in the markets) but do not go over to less expensive brands.

According to the data, however, people tend to substitute between brands as well. This is evident from the comparison of substitution effects (biases) in table 6: When using ACN brand-classification (substitution between brands not allowed), the substitution bias, measured as a difference between Laspeyres and Fisher price indices is 1,1 per cent whereas when applying Coicop 7-digit classification (where substitution between brands is allowed), the substitution bias (or effect) is only 0,6 per cent.

If we keep our sample of outlets constant, we loose the price decrease brought up by new outlets competing on the markets. Also, we loose the effect of consumers changing to less expensive outlets already in the markets.

Similarly, if we keep our product set extremely tightly defined, we loose a considerable part of the transactions in the markets and hence may create a biased index.

The above mentioned substitution effects have traditionally been discussed in the context of different index formulas (see, for example, de Haan 2001). The exercise exposed here should show that the issue can - and should - be treated also as a problem of product classification, or, in traditional CPI context, as a problem of matching, replacement and quality adjustment rules.

The problem of quality change has in the CPI research literature mainly been discussed in the context of products where rapid technological change is apparent. On the basis of the results of this study, it is apparent that "quality change" issues should be considered also in the case of simple commodities.

The nature of quality change for simple products is, however, somewhat different. It is mainly a question of differentiated marketing strategies, i.e. manufacturers attempt to present their own products as distinct as possible. The idea is to try to maintain brand-based segmented micro-markets where substitution between the products should happen within the brand, not between brands. If the segmentation strategy is successful, the manufacturer can adopt quite profitable pricing schemes and plan the obsolescence of their products. The question of how price statisticians should react to this kind of situation in principle is somewhat difficult.

However, in practical data collection situation this decision is taken when determining the "sameness" of the products to be priced.

As there do not exist any self-evident statistical rules on how to deal with different types of classification and comparability issues at the most detailed level of the CPI's , measures should be taken to agree on some generally accepted best practices for types of commodities and market situations highlighted in this study.

**References**


Allen, R.G.D: Index Numbers in Theory and Practice. The Macmillan Press Ltd. London and Basingstoke, 1975.

Balk, B. M: On the Use of Unit Value Indices as Consumer Price Sub-indices. Paper presented at the 4th meeting of the Working Group on Price Indices, Washington D.C, April 22 - 24, 1998.

Diewert, W. E: Axiomatic and Economic Approaches to Elementary price indexes. Discussion paper no 95-01, Department of Economics, University of British Columbia, Vancouver 1995. Referenced in Balk, 1998.

Durkheim, E, and Mauss, M: Primitive Classification. The University of Chicago Press, Chicago 1963 (Original: De Quelques Formes Primitives de Classification, Annee Sociologique, 1901-02 (1903)).

De Haan, Jan: Generalized Fisher Price Indexes and the Use of Scanner Data in the CPI. Paper presented at the 6th meeting of the Working Group on Price Indices, Canberra, April 2001.

Koskimäki, T and Vartia, Y: Beyond matched pairs and Griliches-type hedonic methods for controlling quality changes in CPI sub-indices. Mathematical considerations and empirical examples on the use of linear and non-linear hedonic models with changing quality parameters. Paper presented at the 6th meeting of the Working Group on Price Indices, Canberra, April 2001.

Silver, Mick: Quality adjustment and price indices. Draft chapter for the forthcoming ILO manual on CPI:s ILO, 2001.

Tripplet, Jack E: Handbook on quality adjustment of price indexes for information and communication technology products. Draft report prepared for Industry Committee, OECD Directorate for Science, Technology and Industry. October 1999.

## Annex 1: COICOP classification used in the study

| | |
|---|---|
| ***01.1.5.1*** | **Butter** |
| 01.1.5.1.01 | Dairy butter |
| 01.1.5.1.02 | Other butter |
| ***01.1.5.2*** | **Margarine and other vegetable fats** |
| 01.1.5.2.01 | Butter and vegetable oil mixture |
| 01.1.5.2.02 | Cooking margarine |
| 01.1.5.2.03 | Soft margarine |
| 01.1.5.2.04 | low fat margarine |
| 01.1.5.2.05 | other veg. fats |
| ***01.1.5.4*** | **Vegetable oils** |
| 01.1.5.4.01 | Rapeseed oil |
| 01.1.5.4.02 | Sunflower oil |
| 01.1.5.4.03 | Olive oil |
| 01.1.5.4.04 | Other Oil |
| ***01.2.2.2*** | **Soft drinks** |
| 01.2.2.2.01 | Veg. extract drinks (Coke) |
| 01.2.2.2.02 | Soda orange |
| 01.2.2.2.03 | Energy drinks |
| 01.2.2.2.04 | Soda soft drinks, other than orange |
| 01.2.2.2.05 | Other soft drinks |
| ***01.2.2.3*** | **Fruit juices** |
| 01.2.2.3.01 | Mixed fruit cordial |
| 01.2.2.3.02 | Orange juice, 100 per cent fruit |
| 01.2.2.3.03 | Other cordials |
| 01.2.2.3.04 | Other juice, 100 per cent fruit |
| 01.2.2.3.05 | Juice, less than 100 per cent fruit |
| 01.2.2.3.06 | Other juices and cordials |
| **05.6.1** | **Non-durable household goods** |
| ***05.6.1.1*** | **Detergents** |
| 05.6.1.1.01 | Dishwasher detergent |
| 05.6.1.1.02 | Synthetic detergent (Clothing) |
| 05.6.1.1.03 | Dish washing liquid |
| 05.6.1.1.04 | General purpose cleanser |
| 05.6.1.1.05 | Other detergents |

**Annex 2: Example of the brand -classification – detergents**

AIRWICK
AJAX
ANDY
ASPI
BECKMANN
BEMINA
CHEVY
CLEANI
DEONET
DETEX
DOMESTOS
EKO
EPEX
EUROSHOPPER
FINN MAID
GREEN 2000
HARDOL
HAVI
HIT KLARO
JOHNSON
KIILTO
KLORIN
KLORITE
LASER
METSANRAIKAS
MINIRISK
MONSUN
MR MUSCLE
MR. PROPER
MT-EXIMA
PINJA
PIRKKA
REMOP
SANILAV
SIISTO
SINETTI
SMART
SPAR
SUPER
SUPI
TOILET DUCK
TOLU
WC DUCK
WC-FLOWER
VIM
MUU TUOTEMERKKI

## Annex 3: Price levels in Scanner-(ACN) and CPI – data sets

Samples and varieties in different data sets have not been matched, hence the differences may reflect, in addition to differences in measurement, also differences in samples

| Orange Juice | | | | Coke | | | |
|---|---|---|---|---|---|---|---|
| Month/year | ACN | CPI | CPI/ ACN*100 | Month/year | ACN | CPI | CPI/ ACN*100 |
| 01/95 | 4,88 | 5,09 | 104,2 | 01/95 | 4,70 | 5,43 | 115,6 |
| 01/96 | 4,99 | 5,20 | 104,2 | 01/96 | 4,85 | 5,16 | 106,3 |
| 01/97 | 5,63 | 5,41 | 96,1 | 01/97 | 5,40 | 5,06 | 93,7 |
| 01/98 | 5,08 | 5,48 | 107,8 | 01/98 | 5,83 | 5,67 | 97,2 |
| 02/98 | 5,40 | 5,46 | 101,2 | 02/98 | 5,69 | 5,75 | 101,1 |
| 03/98 | 5,24 | 5,45 | 104,1 | 03/98 | 5,73 | 5,76 | 100,6 |
| 04/98 | 5,18 | 5,44 | 105,1 | 04/98 | 5,79 | 5,80 | 100,3 |
| 05/98 | 5,31 | 5,46 | 102,9 | 05/98 | 5,76 | 5,82 | 100,9 |
| 06/98 | 5,28 | 5,46 | 103,4 | 06/98 | 5,44 | 5,77 | 106,1 |
| 07/98 | 5,41 | 5,44 | 100,5 | 07/98 | 5,69 | 5,81 | 102,2 |
| 08/98 | 5,38 | 5,39 | 100,1 | 08/98 | 5,75 | 5,82 | 101,3 |
| 09/98 | 5,65 | 5,47 | 96,9 | 09/98 | 5,81 | 5,76 | 99,2 |
| 10/98 | 5,74 | 5,48 | 95,4 | 10/98 | 5,82 | 5,80 | 99,7 |
| 11/98 | 5,83 | 5,57 | 95,5 | 11/98 | 5,72 | 5,79 | 101,1 |
| 12/98 | 5,87 | 5,56 | 94,7 | 12/98 | 5,73 | 5,81 | 101,4 |

| Fruit cordial | | | | Mineral water | | | |
|---|---|---|---|---|---|---|---|
| Month/year | ACN | CPI | CPI/ ACN*100 | Month/year | ACN | CPI | CPI/ ACN*100 |
| 01/95 | 10,35 | 7,02 | 67,8 | 01/95 | 3,95 | 4,56 | 115,5 |
| 01/96 | 12,94 | 8,32 | 64,3 | 01/96 | 3,92 | 4,31 | 110,0 |
| 01/97 | 16,29 | 8,36 | 51,3 | 01/97 | 5,06 | 4,29 | 84,8 |
| 01/98 | 15,45 | 8,38 | 54,2 | 01/98 | 5,02 | 4,93 | 98,4 |
| 02/98 | 15,65 | 8,48 | 54,2 | 02/98 | 5,26 | 4,92 | 93,6 |
| 03/98 | 15,41 | 8,32 | 54,0 | 03/98 | 5,27 | 4,92 | 93,4 |
| 04/98 | 15,99 | 8,34 | 52,2 | 04/98 | 5,25 | 4,96 | 94,6 |
| 05/98 | 15,04 | 8,32 | 55,3 | 05/98 | 5,24 | 4,93 | 94,1 |
| 06/98 | 14,36 | 8,26 | 57,5 | 06/98 | 5,07 | 4,94 | 97,3 |
| 07/98 | 14,11 | 8,15 | 57,7 | 07/98 | 5,17 | 4,94 | 95,4 |
| 08/98 | 16,24 | 8,04 | 49,5 | 08/98 | 5,21 | 4,93 | 94,6 |
| 09/98 | 15,53 | 8,40 | 54,1 | 09/98 | 5,35 | 4,97 | 92,8 |
| 10/98 | 14,81 | 8,36 | 56,5 | 10/98 | 5,19 | 4,98 | 96,0 |
| 11/98 | 15,67 | 8,28 | 52,9 | 11/98 | 4,93 | 5,02 | 101,9 |
| 12/98 | 15,96 | 8,24 | 51,6 | 12/98 | 5,49 | 5,06 | 92,1 |