

Distr.  
GENERAL

CES/AC.49/1999/12  
14 September 1999

ENGLISH ONLY

STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE

INTERNATIONAL LABOUR  
ORGANIZATION (ILO)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/ILO Meeting on Consumer Price Indices  
(Geneva, 3-5 November 1999)

#### IMPLEMENTATION OF HEDONIC METHODS IN STATISTICS FINLAND

Invited paper submitted by Statistics Finland\*

##### Summary

1. Statistics Finland puts high priority to the quality of price and cost indices. In this paper we present some outcomes and future perspectives of the co-operation between Statistics Finland and the University of Helsinki in a project concentrating mainly on bias due to quality change. Different other sources of potential biases in index numbers have also been identified, such as biases due to the index formula and substitution, or outlet and item selection biases, but it appears to be an accepted fact, that the bias due to changes in the quality of goods and services is the biggest problem in price indices.

2. In Statistics Finland hedonic method has been studied and applied at least tentatively in the following areas: used cars, new cars, house prices, rent statistics, wages and employment costs and some consumer durables. A group of economists and statisticians are further experimenting with existing databases and developing new methods to collect quality related data in CPI.

---

\* Prepared by Jarmo Hyrkkö, Arja Kinnunen and Yrjö Vartia. Mr. Jarmo Hyrkkö is the director and Mrs. Arja Kinnunen a former senior statistician in Prices and Wages Department in Statistics Finland. Mr. Yrjö Vartia is a professor in the Department of Economics, P.O.BOX 54 (Unioninkatu 37), FIN-00014, University of Helsinki, Finland, E-mail: Yrjo.Vartia@Helsinki.fi

## **Introduction<sup>1</sup>**

3. Statistics Finland has put very high priority to the quality of price and cost indices and has participated actively in the development projects financed by the European Union. Finland has worked actively in four Study Groups and Pilot Surveys acting as a leader in two. Finland has a long tradition both in the theory and practice of index numbers. It is worthwhile to mention, that more than 60 years ago the much cited article of Törnqvist<sup>2</sup> was published in the Bank of Finland Monthly Bulletin. In the paper Törnqvist described his ideas, methods and practical solutions to produce CPI figures *twice a month*.

4. In this paper we present the outcomes and future perspectives of the co-operation between Statistics Finland and the University of Helsinki in a project concentrating mainly on bias due quality change. Besides the authors, Antti Suoperä (Statistics Finland), Eugen Koev (University of Helsinki), Anja Roikonen (University of Jyväskylä) and Markus Halonen (Autodata) have contributed to the development of methods. Also Jarko Pasanen, Jari Haapasalmi and Mari Ylä-Jarkko (Statistics Finland) have been active members of this group.

5. The development of (Consumer) Price Indices has received very much attention during the past years<sup>3</sup>. Different sources for potential bias have been identified. It appears to be an accepted fact, that the bias due to changes in the quality of goods and services is the biggest problem in price indices. Although many other sources for bias exist, such as the index formula and substitution bias, outlet and item selection bias, definition of more or less homogenous index commodities and quality adjustment need a lot of attention.

## **Price indices**

6. The European Commission has financed index development work for creating the Harmonised Index of Consumer Prices (HICP). This work has improved the national Consumer Price Indices (CPI) which will remain in most countries as main inflation measures for domestic purposes. As the national CPIs measure temporal price changes *within* a country, the HICP is constructed to measure differences in inflation *between* member states on comparable basis as required by the inflation convergence criteria according to the Maastricht treaty. Therefore, the potential bias due to index formula mostly cancels away in the HICP as long as the same formula, actually Laspeyres, is used in all countries.

7. Statistics Finland calculates several price and cost indices which all require control of quality change:

\* Consumer Price Index (1990=100) and the there of derived HICP (Harmonised Index for Consumer Prices) and the Indicator of Underlying Inflation (IUI)

---

<sup>1</sup> Based on a Discussion Paper No. 450:1998, Department of Economics, University of Helsinki.

<sup>2</sup> Törnqvist: The Bank of Finland's Consumption Price Index, Bank of Finland Monthly Bulletin, Vol. 10, pp. 1-8. 1936.

<sup>3</sup> The Boskin report, The HICP work co-ordinated by Eurostat and discussions conducted in many countries.

\* Producer Price Indices (Wholesale Price Index, Producer Price Index for Manufactured Goods, Basic Price Index for Domestic Supply, Export Price Index and Import Price Index, all having 1990=100)

\* House Price Index (1983=100)

\* Rent Statistics

\* Price Index for Real Estates and Unbuilt Land (ongoing study in Statistics Finland).

\* Employment Cost Index (pilot study relating to the European Employment Cost Index, EECI) and Index of Wage and Salary Earnings (1990=100)

8. Up till now the quality bias in the CPI has been tackled with by price collector's assessments. Specific quality change codes are used<sup>4</sup> which in principle should work fine but which in practice has been shown to lead to inconsistencies. A major improvement has been the EU regulation that bans *automatic linking* as a default in forced replacements of index commodities. In the case of introducing a new variety for a disappeared one, prices are calibrated in this method so that the price level is kept intact. Therefore, automatic linking causes downward bias in inflationary periods.

9. Instead of subjective assessments, more reliable objective and quantitative methods are called for. Hedonic methods provide a general technique for estimating the price effect of quality change. There has been some scepticism for the use of hedonic regression but with sufficient and detailed data there is no doubt that hedonic methods yield good estimates of the impact of quality change<sup>5</sup>. The House Price Index, Used Car Index in the CPI and Employment Cost Index are good examples of this. With modern software hedonic models with even 100 000 observations and 50 variables can be calculated in few seconds. Expert knowledge is required when implementing sophisticated statistical methods, but the basic ideas and results are intuitively understandable.

#### **Hedonic method for used cars**

10. Hedonic approach was first adopted in the measurement of used cars. The method has been described in the work of HICP Study Group: Vehicles<sup>6</sup>. The actual price collection and modelling are done in a company Autodata and the CPI gets every month the aggregate car index as an input to the CPI system.

11. Autodata's index is based on asking prices obtained directly from car dealers and the material covers approximately 90-100 per cent of all second hand cars traded through car dealers in Finland. The index is based on about 70 regression equations where the log-asking price is explained with:  
-age of the car

---

<sup>4</sup> Kinnunen: CPI 1990=100, Methodology and Practice, Statistics Finland Research Paper 200

<sup>5</sup> Vartia, Koev and Suoperä: Combining classification and hedonic technique, Study Group Paper presented in Helsinki, February 1997.

<sup>6</sup> Vartia (University of Helsinki) and Halonen (Autodata): Price Measurement of Second Hand Cars currently applied in the Finnish CPI, Study Group Paper presented in Helsinki, February 1997.

-mileage

-its price as new

-time dummies for 5 preceding months

-minor corrections for area price differences

12. The index is based on quality adjusted prices of 70 most popular car models. From these a weighted Laspeyres index is calculated using monthly chaining and quality adjusted prices at standard quality points.

13. Co-operation with Autodata enables us to have very detailed approach to hedonic modelling. But even with lesser data the same approach can be used as in fact has been recommended by Eurostat's Study Group<sup>7</sup> as one possible option for used car measurement. The calculations are based on standard quality points defined by model specific ages and mileages. Results prove to be nonsensitive with respect to a reasonable range of standard quality points.<sup>8</sup>

### **Hedonic modelling for new cars**

14. For new cars hedonic approach was also planned but finally a much easier and simpler method was developed<sup>9</sup>. This method based on relatively simple hedonic models resulted to expressing the price of a car in relation to its size and power. This is argued to capture the effect of "real technical" quality change (improvement) in time together with proper sampling strategy<sup>10</sup>. This is shown in Figure 1.

15. What in practice is experienced as a big problem (options becoming standard) can as an intermediary solution be solved with option cost method favoured by the UK. This means, that when for instance air conditioning becomes part of a standard model, the resulting price increase can be adjusted by using say half the price of the air conditioning when it was optional. A more objective way would use hedonic technique for adjusting for quality, but care should be taken not to overadjust. Many new features of a car may namely correlate with size and power<sup>11</sup>. This method is implemented in the current Finnish CPI and how it works in practice will be reported in the future.

---

<sup>7</sup> Progress Report of the Vehicles Study Group HICP 97/145, Luxembourg June 1997.

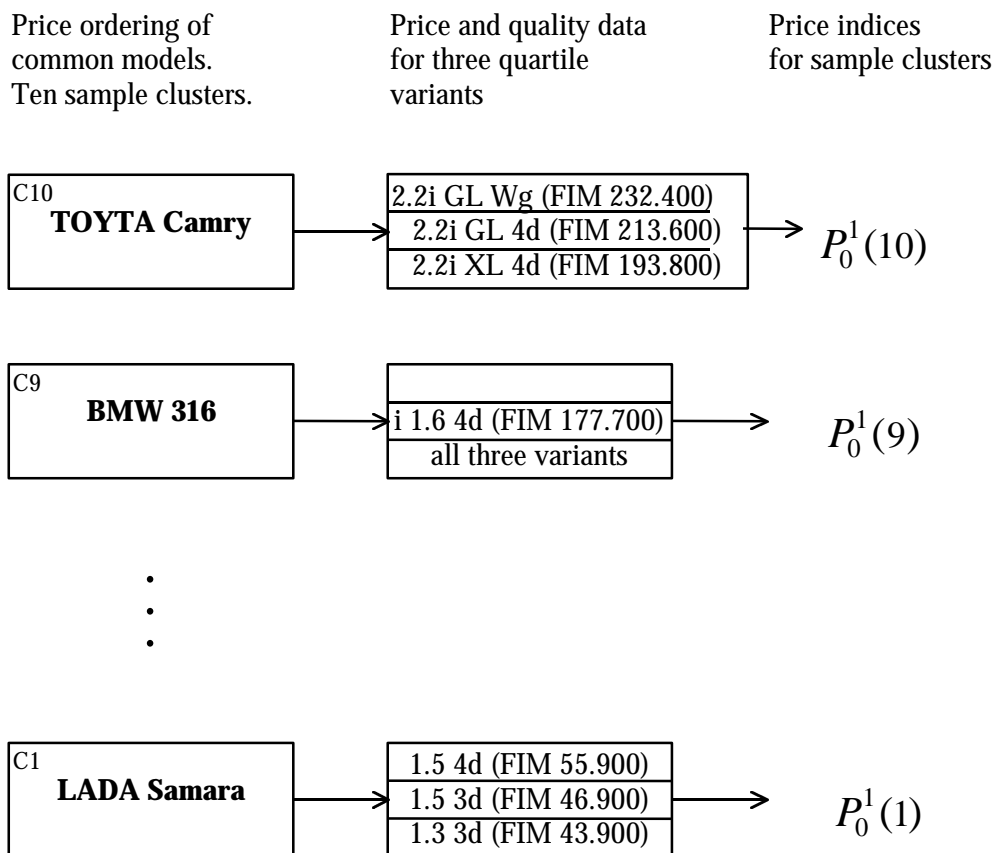
<sup>8</sup> Vartia (University of Helsinki) and Halonen (Autodata): Sensitivity Analysis Concerning Standard Quality Points, Study Group Paper presented in Luxembourg, April 1997.

<sup>9</sup> Vartia (University of Helsinki): Quality Differences and Prices of New Cars: Price per Quality Unit Approach, Study Group Paper presented in Helsinki, February 1997.

<sup>10</sup> Vartia (University of Helsinki): Item Selection, Quality Change and Grouping in the CPI: the Case of Cars, Study Group Paper presented in Luxembourg, April 1997.

<sup>11</sup> Miss Tanja Wahlberg is preparing her master's thesis in Economics on the subject.

Figure 1: Design for a New Car Price Index



$$P_0^1(j) = \prod \left( \frac{p_{ij}^1}{p_{ij}^0} \right)^{1/3}$$

= a geometric average of the three representative prices in the model cluster j. Prices  $p_{ij}^1, p_{ij}^0$  are **quality corrected** ( or expressed in quality units), say divided by “size and power” of the car.

$$P_0^1 = \sum w_j^0 P_0^1(j), \text{ where}$$

$$w_j^0 = \frac{\bar{p}_j^0}{\sum \bar{p}_k^0}$$

**Quality adjustment in the House Price Index**

16. Housing is the biggest main category of the Finnish CPI (20 per cent in the CPI 1990=100). Owner occupied housing has been and still is the principal type of shelter and it plays also important role in the CPI. Both the House Price Index (as a proxy for depreciation costs) and the rent index use hedonic technique.

17. Statistics Finland calculates a quarterly index for house prices, that is apartments, flats and detached houses. The prices are obtained from major real estate agents comprising some 30 per cent of all housing deals concluded.

18. The old method for calculating House Price Index classified the prices per m<sup>2</sup> according to regions, flat sizes (studio, 2 -bedrooms, etc.) and the year of construction. This method gives satisfactory results only when the datasets in each period do not differ as to their characteristics.

19. But the population of dwellings sold in each period (quarter) varies with respect to certain characteristics. For this reason hedonic method has been implemented in the house price index<sup>12</sup>.

20. Equation (1) shows how the market prices of dwellings vary according to their characteristics:

$$(1) \text{Log}P_{iat} = a + \sum_{k=1}^p b_k x_{kiat} + \sum_{a=2}^{Po} m_a A_{ait} + \sum_{t=2}^J l_t T_{ti} + e_{iat}$$

$P_{iat}$  = i.th price observation in area a in quarter t

$x_{kiat}$  = value of characteristic k in observation i, area a in quarter t

$A_{ait}$  = dummy variable for areas (zip code areas are used)

$T_{ti}$  = time dummy for quarter t.

21. The equation (1) is modified in to more operational form:

$$(2) \text{Log}P_{iat} = \text{Log}P_{\bar{a}} + \sum_{k=1}^p b_k (x_{kiat} - x_{k\bar{a}}) + \sum_{t=2}^J l_t (T_{it} - T_{\bar{a}t}) + (e_{iat} - e_{\bar{a}}).$$

22. In this equation the area dummies have been removed by calculating area means of both sides of (1) and subtracting it from (1).

23. In the practical estimation, the following 10 variables are used: price per m<sup>2</sup>, floor area and its square root, the age of the flat and its square root, dummy for a block of flat, number of rooms and the condition of the flat (good, satisfactory, unknown). In addition 7 time dummies relate to the quarters of the two years.

24. The estimate of the quality adjusted price change is based on the coefficients of the time dummies. To increase stability, data for the latest two years is always included; as a new quarter enters, the oldest quarter is

<sup>12</sup> The method derives from the master's thesis in Economics by Markus Halonen in 1996. See also House prices, Statistics Finland, Housing.

dropped. The first quarter of 1994 is the base period and is always included; the difference of the coefficients of base period and of the last quarter defines the log-change of the index.

25. The old and new methods have given fairly similar estimates for the price development for the whole country and major regions. Differences are bigger in more detailed aggregates because in some periods in some areas a bigger share of better flats might have been sold. This is accounted for in the index based on hedonic regression but not in the old traditional method.

26. We are currently reviewing the method used in House Price Index. A future option is that this Griliches type hedonic approach and time dummies will be replaced by hedonic imputation (or rather hedonic interpolation).

#### **Quality adjustment in the Rent Index**

27. In May 1997 a new method was adopted in the rent statistics. Hedonic technique applied here is very similar to the model of House Price Index. Since the model is brand new and it gave very different estimates for price movements compared to the old methods, no changes in the CPI are yet made

28. The purpose of the new rent index<sup>13</sup> is to describe how much more or less people will have to pay during the period under review for a rented dwelling of roughly the same standard compared to the base period. If dwellings of exactly the same standard were rented during each period, the measurement of changes in housing rents would present no difficulty: the average rent per square metre, for example, would directly provide an accurate and reliable measure. However, the stock of dwellings rented at any given time has changed in recent years, and therefore changes in mean rents per square metre do not provide a proper measure of price changes.

29. The hedonic price index is based on the idea that the rents of dwellings are determined in the market on the basis of their characteristics. The rents and the quality characteristics provide a basis for estimating the hedonic price model, which determines expected rents of dwellings in terms of their characteristics just like in the case of house prices above.

#### **Finnish Study relating to European Employment Cost Index (EECI)**

30. According to the Finnish Study and empirical tests it is possible to utilise existing large individual wage databases supplemented by wage tariffs in constructing EECI, even though new practices and procedures have to be implemented.

31. Hedonic methods may be effectively used to control quality differences (or structural changes) and to produce sufficiently homogenous labour classifications. Tests with large databases showed that *hedonic imputation*

---

<sup>13</sup> Rent Statistics, Statistics Finland: Housing, 1997:3. Also a master's thesis in Economics by Mr Kari Takala is under preparation on this topic.

produced as or more reliable results as traditional index calculations based on very detailed classifications<sup>14</sup>, see "Traditional and Hedonic Methods Compared" below. The current Index of Wages and Salaries uses a much rougher classification and is prone to biases caused by structural changes.

---

<sup>14</sup> Vartia (University of Helsinki): Effects of classification and index number formula on labour price and quantity development, A paper produced for the Finnish Pilot Study.

## ***Traditional and Hedonic Methods Compared\****

### **Key stages in constructing a nominal wage/labour cost index:**

1. Determine as well as possible the different kinds/qualities of labour.
2. Obtain an estimate of the expected wage for each labour type in the base and the current period.
3. Use the estimates from 2. together with appropriate weights as input in the index number formula chosen.

### **Implementation:**

#### **Traditional**

1. Cross-classify the employees according to the most important characteristics determining pay. (education, experience, skill requirements of the job etc.)
2. Compute average wage for each class in the current and the base period.
3. Use estimates from 2. as inputs in the chosen index number formula.

**Pros:** No need to assume anything as to how different labour characteristics affect wage.

**Cons:** The number of classes goes to thousands even with a small number of cross-classifiers. Many classes will be empty or containing very few observations. In such cases no average wage estimates can be obtained for each period or estimates may be unreliable. In practice this means that one has to use classifications which are too rough to account properly for the labour quality differences (the labour heterogeneity).

#### **Hedonic imputation**

1. Decide on the most important characteristics of the work and the worker determining pay.
2. Specify a regression model explaining wages in terms of the underlying characteristics (specification may differ e.g. by branch of industry). Estimate the regression model separately for the base and the current period. Use the results to obtain estimates of the expected wage in each period for each class implicitly defined by the regressors in the model.
3. Use the estimates from 2. as inputs in the desired index number formula.

**Pros:** Compact representation of the wage formation, ability to "impute" wages for labour types which are not observed in every period. Wage regressions are common practice in empirical labour economics.

**Cons:** Regression analysis requires a number of assumptions concerning the form of the regression function and the statistical error term. Econometric modelling always involves some subjective judgement.

\*Based on seminar presentation by E. Koev at the Finnish Statistical Society, Helsinki, 22.10.1997

32. In hedonic imputation quality changes are controlled independently of the actual index number calculations meaning that any traditional index number formula may be applied to hedonically imputed, quality controlled prices<sup>15</sup>. For example, Laspeyres and Paasche formulas are shown to give practically equal results for the price index  $P_0^1(Y)$  of the gross wages  $Y$ .

33. EECI is based on a wide cost concept: in addition to gross wages  $Y$  for example employer's social costs are included in the total costs  $Z$ . The study suggests<sup>16</sup> that the social costs should be treated in a multiplicative way by calculating a separate index  $P_0^1(T)$  measuring employer's social costs. The index  $P_0^1(Z)$  of total costs  $Z$  is defined as the product  $P_0^1(Y)P_0^1(T)$  of the index of gross wages  $P_0^1(Y)$  and the index of employer's social cost factor  $P_0^1(T)$ . This seems to be a better method<sup>17</sup> than trying to apply traditional classification or hedonic method directly to total costs  $Z$ .

#### **Testing hedonic method in practice in the CPI; the case of vacuum cleaners and fridge-freezers**

34. National Consumer Research Centre filled the missing parts in our CPI database for vacuum cleaners and fridge-freezers. The aim was to get all relevant quality information to enable hedonic regression, which was also done<sup>18</sup>.

35. The next step was to widen the target sample of both the collected varieties and the number of prices collected. A "check list" was designed with the intention that the price collectors are forced to collect the same quality identifying information that is vital input in the process of hedonic modelling.

36. Price collectors have now used the check lists for these two items since August 1997. There is not yet enough information for actual modelling, since the learning process is still going on. Nevertheless, we have already learned among other things that:

\* Although we tried to concentrate on very essential quality information (13 variables for the vacuum cleaners and 10 for the fridge-freezers) it appears still very difficult to get the information from the brochures or from the shop keepers.

\* Price collectors claim that the quality information should be collected centrally since the models sold are basically the same in the whole country.

---

<sup>15</sup> Koev (University of Helsinki): Constructing a hedonic wage index: Pilot study for the Finnish metal industry, A paper produced for the Finnish Pilot Study.

<sup>16</sup> Study Report p. 9. See also Vartia: List prices and transaction prices: reduction correction factor, Study Group Paper presented in Helsinki, February 1997.

<sup>17</sup> Mr Heikki Pursiainen is preparing his master's thesis in Economics on the subject.

<sup>18</sup> Halonen: Price formation and quality changes for vacuum cleaners and fridge-freezers (in Finnish), A paper produced for the Finnish Pilot Study, May 1997.

But, the number of different models for each manufacturer even in Finland is huge<sup>19</sup>.

\* The cost of extra price collection is important. When the check lists were completed for the first time, the price collectors needed from 30 min to 2 hours extra time per outlet (for these two items).

37. The practical tests are still going on. We are not yet convinced, that the price collectors should use very detailed check lists, but the quality identification information is essential in the quality management of CPI production.

38. Finally, we are working with different strategies concerning the actual hedonic modelling, how and when to do it. Our strategy is to make stepwise progress, first learn the process, manage the price collection and control the price collection costs. After vacuum cleaners and fridge-freezers we will most probably attack clothing.

**Efficient data capture is a necessary condition for  
controlling quality change in the CPI**

39. Our aim is to produce a reliable CPI as efficiently as possible. For this the basic data has to be of good quality. The biggest problem in the CPI price collection and biggest source of potential bias is the uncontrolled quality change occurring with each replacement of commodity variants. The potential bias due to quality changes can be reduced by applying different quality adjustment methods for different items and item groups. Hedonic method is a good alternative if we have relevant quality information. Therefore, we have to start collecting also quality data together with prices.

40. Computer assisted data collection is very important if we want to implement hedonic methods. Paper and pen method will work too slowly if we want to keep to current timetables and cost constraints. In the HICP framework computer assisted data collection has been discussed.

41. EU has had a significant impact on raising the index number construction on the forefront of research after some stationary decades. More international co-operation is needed in this area to avoid the reinventing of the wheel in many countries. The Finnish approach is to emphasis the control of quality change in the data capture and to develop operational systems for the implementation of hedonic methods where appropriate. Co-operation between the statistical agencies and university researchers should be an essential part in this development.

---

<sup>19</sup> Currently there are 13 different brands and 93 different model notations from 237 sheets for the vacuum cleaners and 17 brands and 88 model notations from 238 sheets for the fridge-freezers (in the CPI test).