Statistics Finland

Paper for the 4th International Conference of the Ottawa Group

Arja Kinnunen

Hedonic method is practicable in CPI compilation

This paper aims at describing the rather extensive and serious work Statistics Finland has invested on developing new methods of price index compilation. I try to discuss the quality control of the CPI from rather practical point of view. I also present briefly the work done using hedonic methods at Statistics Finland; what we have learned and what makes us still to believe that quality change can be managed in a professional way.

I thank professor Vartia[1] for his very inspiring thoughts and my many colleagues[2] at Statistics Finland for their courageous work. I also apologise for being verbal and not providing numerical evidence for what I argue. These have been done in those many papers my colleagues have produced during the last years and which will be published in the near future. Also Jörgen Dahlen's paper[3] has to some extent discussed more theoretically what we have discussed in Finland.

1. Some elementary thoughts about the CPI

The CPI is in practise used for analysing whether there is inflation (or deflation) faced by consumers acquiring consumer goods and services. Though the theoretical background is in consumer choice theory and CPI might be used as a proxy for "cost of living index" the basic question the CPI should answer is: "have prices of consumer goods and services changed during a specified time period taking also into account changes in quality". The Finnish CPI has traditionally been a compensation index with broad definition of consumption. This does not however mean, that the CPI intends to be a cost of living index.

The CPI is based on a detailed basket which should represent all relevant consumer goods and services bought by the index population. In Europe COICOP classification is used , which divides the consumer good universe into carefully defined classes (by purpose). In practical CPI work these classes are often divided into sub-classes which then form the finest level where explicit weighting can (or is) used. These we call the elementary aggregates.

For each elementary aggregate we tend to choose one representative item for which we then month after month collect prices. This makes economic sense but is problematic. Modern world with rapidly changing markets has made consumer goods rather difficult to follow. Qualities are changing, new models appear, older ones disappear, existing varieties are changed, for better or for worse causing a lot of trouble for CPI statisticians.

---

[1] Professor Vartia (University of Helsinki, department of economics) has been our scientific consultant since February 1997. With a recent contract with University of Helsinki we will continue this very fruitful co-operation with the academic world. There are several masters theses under preparation and two licentiate theses on price indices.

[2] Jarmo Hyrkkö, director of Price and Wages Statistics, Eugen Koev, Antti Suoperä, Mari Ylä-Jarkko, Jarko Pasanen and Mari Suviranta.

[3] On the statistical objective of a Laspeyres' price index, Eurostat's Task Force paper PTF2/98/97, Lisbon, 9-10.3.1998.

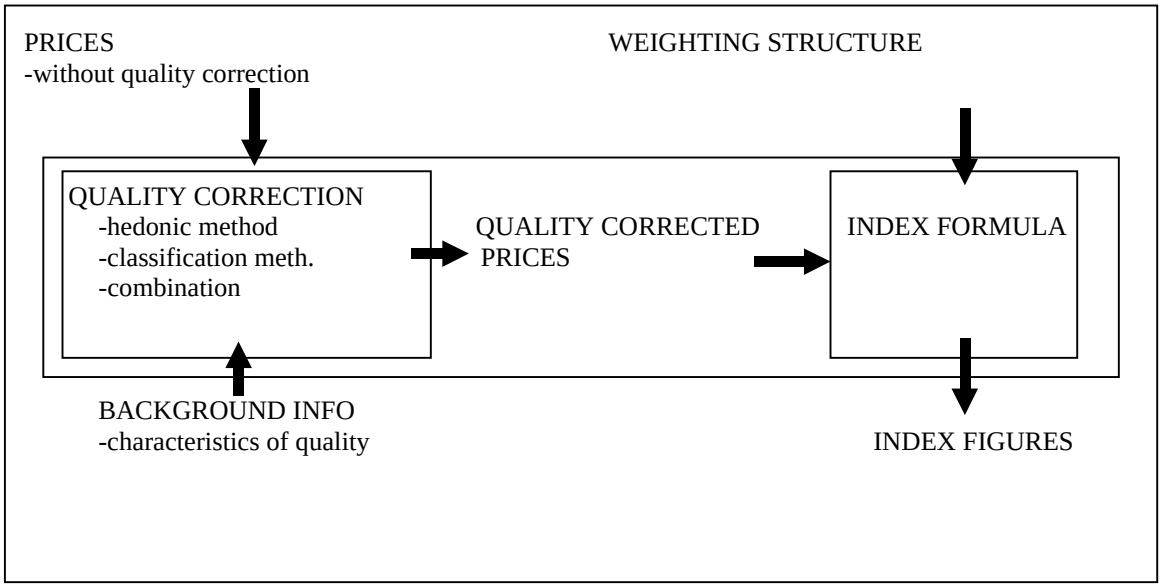
Very common way of treating heterogeneous items and item groups is to use finer and finer classifications, isolating every differing factor and creating new classes in the hope of arriving at controlled quality level. When something happens to the collected variety representing a given class we have difficulties in interpreting and adjusting the figures because we lack information.

The process of sub-grouping is too often forced up to the ignorance level where no further information within the group is available. Our point is, that classification should not be forced to the ignorance level. Instead of taking only one variety from a fine-group we propose to take, say three representatives (low, medium, high) on a less detailed level.

In Europe we have had numerous meetings discussing what to do when quality change occurs; when to compare directly assuming no quality difference, when to apply explicit estimates, when to use overlapping prices, when to link. So far we have not found consensus let alone developed widely accepted practices which would solve our problems.

We should not seek for a general method solving quality adjustment methods. Instead we should concentrate on identifying our elementary aggregates, designing appropriate sampling schemes, using classification and/or hedonic methods for managing quality differences. Careful and well managed quality correction does not restrict the choice of index formula.

The basic structure of CPI compilation can be characterised with following picture[4]:



[4] Report of Finnish Study Relating to European Employment Cost Index (EECI). DG34/E-1/0004211/16 July 1996.

2. The quality of data and the nature of the representative item determine the practical method for controlling the quality change

CPI statisticians so often work with scarce and not optimal (quality) data. Also representative items differ; some are rather easy to follow some are very difficult. Two cases could be distinguished:

1. The quality is easily defined and can be observed. These are typically items where quality can be associated with quantity and classification is a useful tool.

    a) Butter in a packet of 500 g. There are unsalted, salted and "traditional" types. If we know the weights of all different butters we use them, otherwise we collect all different butters (if they all are relatively important) and allow possibly slightly changing package sizes. Observed prices converted to kilo prices.

    b) The bananas and ECO bananas. The kilo price for "normal" bananas and for "ECO bananas" are followed separately and weighted together to get a banana index.

    c) If a clothing store uses price brackets, a summer dress costing 19,99 can be compared with all dresses in the same rack costing 19,99. Here the price bracket is assumed to indicate "comparable quality".

2. Quality of a product consists of several quality components.

    a) We have enough observations and 2-5 quality relevant variables and can possibly run hedonic models every now and then (annually, biannually,..)

    b) We have access to quality relevant data and can actively observe the market dynamics. Hedonic interpolation solves the quality adjustment problems

The traditional way of looking at prices has been to observe the prices per piece, or prices per kilogram, litre, etc. Should not the prices be, where possible, expressed in terms of quality units?

Examples:

- Compact washing powder in packets of 1 kg, price per number of washes possible with one kg (as advertised on the package) could be the quality unit.
  We need to collect besides the price, the number of washes. Although price per wash should be better than price per package, the quality could also be seen as a function of chemical details (phosphates, etc.) requiring hedonic technique and making the work more complicated.

- New cars: prices expressed in quality units with relation to their weight and torque[5]. As a minimum solution, only data for two variables have to be collected. If this sounds too simplistic, several other variables may be introduced. Weight and torque seem to perform well at least in Finland (80 per cent of the price variation is explained by these very two variables).

Prices expressed in terms of quality units should be much more convincing than prices per piece. For the simplest items quantity could be used, for more complicated either quality units or hedonic modelling should be used.

3. Hedonic method

Hedonic technique is an objective method for assessing quality difference (or change) where the effect of several quality components can be analysed. Only observable and measurable characteristics determine the quality. Statistical models and methods (provided that we have sufficient and reliable data) have much firmer basis than any intuitive method never free of subjectivity.

Very often we do not have sufficient data. More data could be collected, but it costs. Consumer goods markets are so dynamic that coefficient estimated more than a year ago might not have relevance next year.

The experience from some countries has shown, that hedonic technique becomes extremely expensive in practice, data collection and the actual modelling are expensive and results are not always feasible. Coefficients are not stable over time, they might be of "wrong" sign, users do not understand this complicated method.

3.1 Griliches type of modelling vs. hedonic interpolation

The work done by Eugen Koev, Antti Suoperä and professor Vartia[6] has convinced us that hedonic method should be seen through hedonic interpolation (earlier we used the term hedonic imputation) and the mostly used "Griliches type" model should be seen as a special case of hedonic interpolation.

By Griliches type we mean the following modelling:

$$(1) \quad P_{it} = \alpha + \sum_{k=1}^{p} \beta_k x_{ik} + \sum_{t=2}^{J} \lambda_t T_{ti} + \varepsilon_{it} \ ,$$

[5] See professor Vartia: Quality differences and prices of new cars: price per quality unit approach, Helsinki 27-28.2.1997 (EC Study Group working papers).

[6] Vartia, Koev, Suoperä: Combining classification and hedonic technique, Study Group: Vehicles working papers, Helsinki 1997.

where α is constant, β_k parameter estimate, x_{ik} value of characteristic k in observation i, T_{ti} time dummy for observation period t and λ_t its coefficient. λ_t describes the pure, or quality adjusted price change compared to the base period.

Griliches type modelling requires stable coefficients in time. This means, that the elementary aggregate the model should describe has same relative price changes. Same relative price changes can be assumed only for very homogenous elementary aggregates. But, if we can achieve very homogeneous classes (in terms of quality and perfect substitution) simple classification approach would be enough. This in itself would be a very relieving solution for forced replacements but would not solve the problem of how to measure quality change in time.

Hedonic technique requires enough data and enough variation in quality. By applying Griliches approach the assumptions we have to make are rather strong.

What if we gave more freedom to the modelling and could run hedonic regression in current period? The latter is wishful thinking for most of my colleagues, but it will be soon the reality with modern software and good understanding of statistical methods.

Hedonic interpolation is basically the following[7]:

1. Decide the most important quality characteristics of the item in question by using statistical methods, usually hedonic regression. These characteristics implicitly stratify the item into homogenous "sub-items".

2. Collect data for these variables either by using the price collectors or centrally at the office.

3. Decide the functional form of the model (semi-log or logarithmic are the usually recommended ones).

4. Estimate regression models separately for the base and the current period.

5. Use the estimator results to obtain the price estimates of each variety for the base and the comparison period.

6. Use the price estimates from (5) as inputs in the decided index number formula chosen.

Hedonic interpolation combines classification and hedonic regression making maximum use of the available data. Hedonic interpolation solves automatically the problem of missing prices, missing observations are imputed with the model. It also allows the coefficients to change in time if the markets start to value more certain features than before. If coefficients are stable in time (according to hedonic interpolation) Griliches type approach is valid.

[7] I have modified Eugen Koev's seminar presentation at the Finnish Statistical Society, Helsinki 23.10.1997.

Stable coefficients should not be assumed as a default option. The aim should be to observe the market and consumer's choices and try to go towards hedonic interpolation or otherwise efficiently use the available information.

4. Some experiences about hedonic methods in Finland

The control of quality change in our CPI has long traditions. Price collectors have been trained to understand that the aim is to measure "pure price" change and an estimate of the value of quality difference[8] has been obligatory in every replacement situation. Price collectors have complained, that the evaluation and decision making is difficult. Also studies have shown, that price collectors are not consistent in their judgements.

First serious attempt into hedonic method was made in 1994 when the basic methodology was reviewed and test were made for cars, compact cameras and vacuum cleaners[9]. The work has continued and currently hedonic technique is used to measure prices of second hand cars and housing. Several studies have been done and many more have been planned.

Though theoretical work in Finland has most recently concentrated on advocating hedonic interpolation Griliches type approach has mostly been used. Practical tests have only begun and there is not enough data to simulate hedonic interpolation vs. stable coefficient & time dummy - method. In the construction of the pilot study for European Employment Index EECI[10] hedonic interpolation was tested together with most index formulas.

4.1. Fridge-freezers and vacuum cleaners

Fridge-freezers and vacuum cleaners both are still being tested and hedonic technique is not yet implemented in the practical CPI computation. Price collectors collect quality identifying data on separate forms. This data collection started in August 1997 and for both items there are about 700 prices.

It has been difficult to get correct data for pre-specified[11] quality variables, information has been insufficient and erroneous. Price collectors will soon get hand held computers which should make the data management easier and more efficient.

The estimated model is[12]:

$$Log P_{it} = \alpha + \sum_{k=1}^{p} \beta_k x_{ik} + \sum_{t=2}^{J} \lambda_t T_{ti} + \varepsilon_{it} ,$$

[8] See Kinnunen: CPI 1990=100, methodology and practice, pp. 31-33 for how it works. Statistics Finland´s Research Papers No 200, 1993.

[9] See Wangqui Song: Hedonic Regression Analysis in Consumer Price Index, Theory and Empirical results, Statistics Finland´s research papers No 207, 1994.

[10] Report of Finnish Study Relating to European Employment Cost Index (EECI). DG34/E-1/0004211/16 July 1996.

[11] A pilot study was made by Finnish Consumer Research Institute by product experts to identify all relevant information affecting the quality of these investigated two items. Also a consumer panel was interviewed.

[12] Mari Ylä-Jarkko: Study on hedonic modelling of consumer durables, Study Group: Consumer Durables, Helsinki, 23.2.1998.

where α is a constant and x_k is dependent variable and may be transformed and β_k is its parameter estimate and constant over period $t = 1, \ldots, J$ and T is time dummy for observation periods and λ_t is its coefficient. The index can be obtained straight from λ_t and it can be considered as a change of prices from the base period to the period t.

Following tables present parameter estimates when ordinary prices have been used:

Table 1. Hedonic model for fridge-freezers. R^2 is 0.6583 and adjusted R^2 is 0.6520 and root MSE is 0.1141.

	Parameter estimate A	**Prob > |T|**
Intercept	12.4159	0.0001
Manufacturer 1	-0.03730	0.0655
Manufacturer 2	-0.1057	0.0001
Energy consumption (log)	-0.4622	0.0001
Volume of fridge (log)	0.2291	0.0001
Volume of freezer (log)	0.3781	0.0001
Freezing capacity (log)	0.1757	0.0001
Automatic freezing system	0.0332	0.0026
Exterior thermometer	0.0428	0.0001

Table 2. Hedonic model for vacuum cleaners. R^2 is 0.7086 and adjusted R^2 is 0.7014 and root MSE is 0.2014.

	Parameter estimate A	**Prob > |T|**
Intercept	6.8329	0.0001
Noise (log)	-0.2750	0.0058
Suction power (log)	-0.1445	0.0131
Weight (log)	0.5834	0.0001
Power (log)	0.3246	0.0302
Range (log)	1.4656	0.0001
Special nozzle	0.2332	0.0001
Automatic control of suction	0.0520	0.5802
Container for nozzles	-0.1173	0.0187

4.2. Mobile phones

Most recent experiment has been done with mobile phones[13]. With experts' assistance 11 important quality characteristics were identified. These characteristics were: talk time according to manufacturer (in minutes), talk time according to measurement, standby time according to manufacturer (*h*), standby time according to measurement, number of ringing tones, size (cm^3), weight (*g*), size of display (cm^2), clock, memory of missed calls and height of font (*mm*). Also all brands were included in the model. Measured standby and talk time were taken from a Finnish magazine 'Tekniikan Maailma', which runs tests for mobile phones regularly.
The model was estimated using a Griliches-type model, with time dummies;

$$(1) \quad P_{it} = \alpha + \sum_{k=1}^{p} \beta_k x_{ik} + \sum_{t=2}^{J} \lambda_t T_{ti} + \varepsilon_{it} \quad ,$$

where α is constant, β_k parameter estimate, x_{ik} value of characteristic k in observation i, T_{ti} time dummy for observation period t and λ_t its coefficient. λ_t describes the pure, or quality adjusted price change compared to the base period. Because the data was for 11 months, 10 time dummies were included in the model. Both semi-log and double-log models were estimated. In double-log model R^2 was better and thus the double-log model was chosen to describe price variation of mobile phones. Modelling was done with different sets of explanatory variables as well as with stepwise regression. Results from stepwise regression were not very reasonable due to multicollinearity, so the final model was not run with this method.

Because of multicollinearity the number of quality variables in the final model became rather small. This is only a good thing, if the model is going to be used in practise. The variables in the model are size (cm^3), talk time, number of ringing tones and memory of missed calls. Also two brands were included in the model. Estimation results are presented in table 3.

Size of a mobile phone has a strongest effect on prices. Being the only explanatory variable in the model, variable size explains 46 % of price variation. Variable weight has almost equal effect on prices, but it was not included in the model because variables size and weight were highly correlated (r^2 =0.73). Number of ringing tones has a positive coefficient, which is very reasonable because number of ringing tones seems to reflect mobile phone's quality. The cheapest and simplest models have only few different ringing tones but for example newest Nokia has 35 of them. Memory of missed calls behaves the same way as number of ringing tones - price rises when memory becomes bigger. Talk time has a statistically significant coefficient. That seems right, because consumers value long talk time. Standby time was not statistically significant, which is most likely due to the high correlation between talk and standby time.

All seven brands were included in modelling, but only two of them were statistically significant. Both these brands, Philips and Motorola, are inferior. What was surprising is that brand Ericsson did not get a significant coefficient, although prices of Ericssons seem to be higher than prices of other brands.

[13] Mari Suviranta: Hedonic model for mobile phones, unpublished paper, for her master's thesis, Statistics Finland, March 1998. Practically the whole chapter 4.2 is a direct quote from her paper (unpublished).

Table 3. Hedonic model for mobile phones

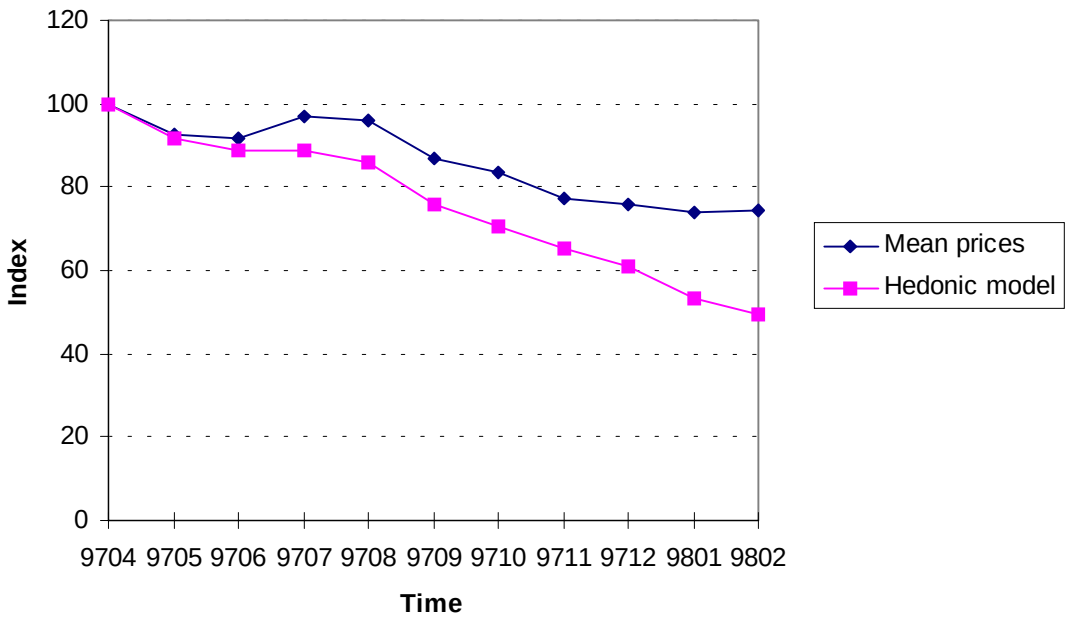
	PARAMETER ESTIMATE	STANDARD ERROR	T for H0: Parameter=0	Prob> $\|T\|$
Intercept	13.213816	0.22302726	59.248	0.0001
Size	-1.324938	0.03851667	-34.399	0.0001
Number of ringing tones	0.152695	0.01452240	10.514	0.0001
Memory of missed calls	0.154126	0.01816858	8.483	0.0001
Talk time	0.194559	0.01509920	12.885	0.0001
Philips	-0.338831	0.03339970	-10.145	0.0001
Motorola	-0.326634	0.04029932	-8.105	0.0001
0597	-0.082278	0.03426733	-2.401	0.0166
0697	-0.112604	0.03426760	-3.286	0.0011
0797	-0.109961	0.03431097	-3.205	0.0014
0897	-0.140175	0.03435508	-4.080	0.0001
0997	-0.241666	0.03421059	-7.064	0.0001
1097	-0.292444	0.03447818	-8.482	0.0001
1197	-0.347194	0.03318143	-10.463	0.00001
1297	-0.390334	0.03039030	-12.844	0.0001
0198	-0.469492	0.03038325	-15.452	0.0001
0298	-0.506351	0.03068976	-16.499	0.0001

$R^2 = 0.7969$

$Adj.R^2 = 0.7925$

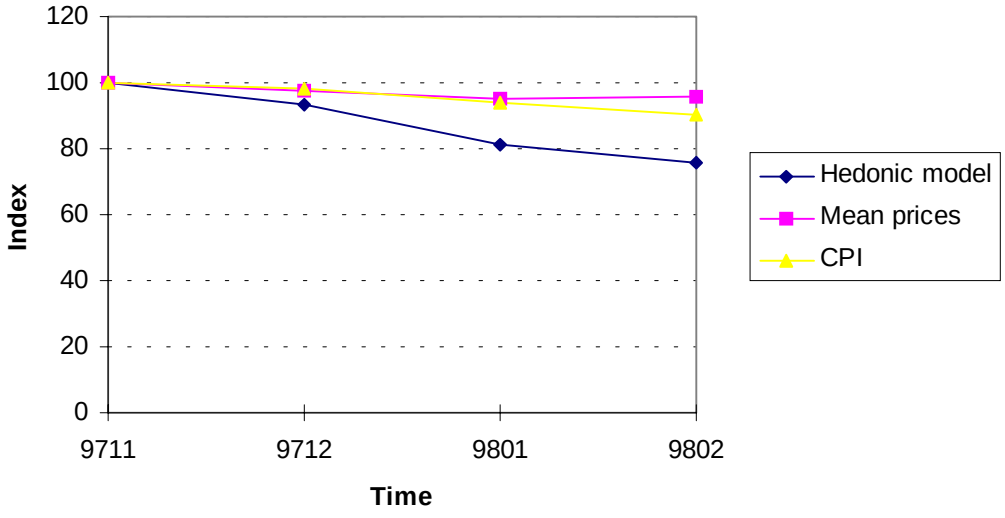
All time dummies in the model are statistically significant. The pure price change can directly be read from the time dummies. The model shows a dramatic decline in mobile phone's prices during the past 11 months. From April 1997 to February 1998 the prices of mobile phones have declined over 50 percent. This may first seem a little too much, but further research shows that it really can be possible. When comparing the results of the hedonic model and geometric mean prices in picture 1, one can see that mean prices have also declined dramatically. When we sum the effect of price decline and quality improvements, the real price decline is drastic.

PICTURE 1. Different index series for mobile phones.



Due to technical problems the price index computed with the hedonic model can only be compared with the official CPI from November 1997. Results can be seen in picture 2.

PICTURE 2. Different indices for mobile phones. Base period November 1997.



Picture shows that the official CPI is clearly higher than the hedonic index. Geometric mean prices are even higher. During this short period the CPI seems to overestimate inflation. We should be careful when interpreting the results, data is still insufficient and more research is needed.

Hedonic interpolation should be an attractive option especially for mobile phones. Quality information can be collected centrally and tests are run very frequently by specialised magazines allowing us to analyse new features and market behaviour almost on-line.