SUPPORTING PAPER No 8 (19)*
19 October 2001

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

INTERNATIONAL LABOUR ORGANISATION (ILO)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/ILO Meeting
on Consumer Price Indices
(Geneva, 1-2 November 2001)

INTRODUCING QUALITATIVE COMPONENTS INTO PRICE SURVEYS

Supporting paper submitted by Central Statistical Office, Poland

---

* Paper posted on Internet as submitted by the country.

Introduction

The paper presents the results of a survey[1], which was carried out to define the influence of quality changes of products on their prices using the regression analysis method. The survey yielded positive results, which confirmed the impact of the specified qualitative features on the price of a product.

Application of econometric models in the evaluation of impact of quality on prices

The necessity to find an appropriate methodology to determine and evaluate the impact of qualitative factors on prices and to obtain the so-called 'pure' price index seems to be obvious after the reading of the Boskin Report. It is very difficult to establish such price index from the methodological point of view however efforts have to be made to obtain the most reliable results.

In the first phase of the survey, the results of which are presented in this paper, an experimental attempt was made to capture the influence of qualitative components on the features of selected model variants of one of household appliances, i.e. an iron. Regression analysis methods were applied to construct and estimate the econometric models.

The model was developed basing on a cross-section sample. Information from the market on the pricing of numerous variants of irons[2] from diverse manufacturers (brands) and types at a given point in time (1st quarter of 2000) was used. The number of observations was n=48. The sample included data on iron models from Bosch, Braun, Eldom, Holden, Kenwood, Moulinex, MPM, Philips, Rowenta, Severin, Tefal, and UFESA. The prices of the irons were very diverse and showed the variation coefficient of 39.44% and the range was PLN 266. One of the reasons for such high price diversity was that various usability features included in the survey were also qualitative components. These qualitative features, i.e. independent variables describing the product price (dependent variable) were as follows: x_1 - electric power, x_2 - steaming, x_3 - sprinkler, x_4 - water-tank capacity, x_5 - steam blow, x_6 - self-cleaning, x_7 - dripping protection, x_8 - automatic switch-off, x_9 - cable length, x_{10} - brand (if recognised as a brand-name company).

Hence, the price of iron (Y) is the k-function of qualitative features (x_k), which can be expressed as the following formula:

[1] The work has been carried out by one of regional statistical offices in Poland (Lodz) under the supervision of Dr Teresa Smi³owska, deputy director of the office, in co-operation with the Price Statistics Division of the Central Statistical Office in Warsaw.

$$Y = f(X_1, X_2, ..., X_k, \varepsilon) \quad (1)$$

where:

k = 1, 2, 3, ..., 10 – number of independent variables (qualitative features)

ε – random component

The difficult stage of specification of usability features of the irons was followed by quantification. Out of the independent qualitative variables, three were measurable ones and included x_1 - power (W), x_4 - water-tank capacity (ml), and x_9 - cable length (cm). Other seven variables were introduced as dummy binary variables. The use of binary variables was possible owing to a large number of observations applied in the survey (n=48).

In order to construct an appropriate function type, a review of statistical information was made (see graphs 1 – 3). Internal correlation (among the independent variables) and the correlation between the qualitative features and the price (see Table 1) were examined. Since the x_2 - steaming and x_3 - sprinkler were identified as carriers of the very similar information, the latter was eliminated and the remaining nine variables were used to develop the first version of the model.

Based on a preliminary review of the pricing data on irons and their qualitative features, two forms of the function type represented by the formula (1) were identified: a linear one and an exponential one (see Kokoski M. et al. (1999), Moulton B. R. et al. (1999)).

The comprehensive model of the linear function can be expressed by the equation:

$$Y = \sum_{k=1}^{9} \alpha_k x_k + e \quad (2)$$

where:

Y – price of an iron

α_k – structural parameters of the model

x_k – independent variables

e - random component

[2] Data from Art.-Dom.Company; internet www.artdom.com.pl

Now, the equation for the exponential function after the logarithm transformation looks as follows:

$$\ln Y = \sum_{k=1}^{9} \alpha_k x_k + \varepsilon \qquad (3)$$

where:

Y, α_k, x_k, ε - as above.

The estimation of the structural parameters of both models was based on the classical least squares method[3].

The estimated structural parameters of the a_k model give rise to various interpretations depending on the type of function. In the case of the linear function, absolute changes in the mean price of irons are shown, whereas in the case of the exponential function relative (percentage) changes are displayed.

However, some modifications of the price vector **[*Y*]** and in the qualitative feature matrix **[*X_k*]** were introduced to the linear model (2). Namely, since the median value computed for a range of prices of irons included in the survey, amounted to PLN 194 and referred to the price of the Kenwood ST 675 model, this iron was therefore taken as the reference model. The initial dependent variable was transformed in relation to the reference object and intervals from the reference price were calculated. Hence, a new vector of dependent variables was established. A similar procedure was applied to the independent variables (qualitative feature matrixes). The values of measurable features describing the particular iron models were replaced by the values of intervals to the features of the reference model. Hence, measurable variables x_1, x_4, x_9 were transformed into the r_1, r_2, r_3 variables. In the case of non-measurable features, a feature carrying qualitative information received the value ‘1’, or ‘0’. The variable describing the brand (x_{10}) took the value ‘1’ in the case of companies regarded in the survey as brand name companies, i.e. Bosch, Braun, Kenwood, Moulinex, Philips, Rowenta, Severin and Tefal; otherwise ‘0’. Based on the matrix of such observation and price interval vector, the following econometric model was constructed:

$$Y = \sum_{i=1}^{3} \alpha_i r_i + \sum_j \beta_j x_j + \varepsilon \qquad (4)$$

where:

[3] see Goldberger A. S. (1972), Paw³owski Z. (1978).

Y – interval from the reference price (PLN)

$\boldsymbol{a}_i, \boldsymbol{b}_j$ – structural parameters of the model

r_i – measurable independent variables

x_j – binary independent variables

$\boldsymbol{e}$ – random component

After the regressive evaluation of the structural parameters, the model (4) can be transformed as follows:

$$\hat{Y} = \underset{(7.202)}{0.125\, r_1} - \underset{(-1.247)}{30.120\, x_2} + \underset{(7.066)}{0.256\, r_2} + \underset{(2.265)}{24.664\, x_5} + \underset{(0.197)}{2.172\, x_6} - \underset{(-0.017)}{0.172\, x_7} + \underset{(1.565)}{27.398\, x_8} + \underset{(2.008)}{0.377\, r_3} +$$

$$\underset{(2.942)}{36.896\, x_{10}} - \underset{(-1.016)}{30.171} \tag{5}$$

$$R^2 = 0.883 \qquad S_e = 28.142$$

The round brackets under the values of parameters show statistic values $t(a_j)$, that have Student's distribution. On the significance level of $\alpha = 0.05$ no substantial influence on the following price variables could be detected: x_2 - steaming, x_6 - self-cleaning, x_7 - dripping protection. However, the remaining variables turned out to be significant, i.e. they influenced the prices of irons in a substantial way. Even though the significance of the x_8 - automatic switch-off [$t(b_8) = 1.565$] can give rise to doubts, it was decided to keep it for further studies.

After elimination of the insignificant independent variables, the econometric model including all the remaining variables took the following form:

$$\hat{Y} = \underset{(8.880)}{0.125\, r_1} + \underset{(7.246)}{0.242\, r_2} + \underset{(2.073)}{10.787\, x_5} + \underset{(1.650)}{14.152\, x_8} + \underset{(2.202)}{0.380\, r_3} + \underset{(2.907)}{17.504\, x_{10}} - \underset{(-0.892)}{11.822} \tag{6}$$

$$R^2 = 0.879 \qquad S_e = 27.648$$

In both estimated models, the value of determination coefficient (R^2) was high with the value above 0.87. In the model (6), the overall variability of the prices of irons could be explained by the estimated model up to 87.9%, and in the model (5) up to 88.3%. The prices of the irons in the sample differ from the theoretical price in the model on average by PLN 27.65, which is by 14.3% of the mean price of an iron.

The estimation of structural parameters presented in the model (6) implies that in the case of brand-name irons their prices are on average over PLN 17 higher with the remaining features being the same. When a given iron is equipped with the automatic switch-off, then its price jumps on average by more than PLN 14, and when it has the ‘steam blow’ function, the price is higher on average by more than PLN 10. The prices of irons increase also along with the rise of such variables as electric power, water-tank capacity, and cable length.

A correct estimation of the structural parameters allows to develop a model of the exponential function (3), which after linear transformation and application of the least squares method, took the following form:

$$ln\,\hat{Y} = \underset{(6.688)}{0.001}\,x_1 + \underset{(0.689)}{0.094}\,x_2 + \underset{(6.880)}{0.001}\,x_4 + \underset{(2.394)}{0.147}\,x_5 + \underset{(0.310)}{0.019}\,x_6 + \underset{(0.719)}{0.042}\,x_7 + \underset{(0.950)}{0.093}\,x_8 + \underset{(2.149)}{0.002}\,x_9 + \underset{(4.503)}{0.318}\,x_{10} + \underset{(11.888)}{2.853} \qquad (7)$$

$$R^2 = 0.902 \qquad S_e = 0.158$$

One can assume on the significance level of $\alpha = 0.05$ that the following variables are crucial: x_1 - electric power, x_4 - water-tank capacity, x_5 - steam blow, x_9 - cable length, x_{10} - product brand.

Hence, the econometric model including only the significant variables can look as follows:

$$ln\,\hat{Y} = \underset{(9.138)}{0.001}\,x_1 + \underset{(7.709)}{0.001}\,x_4 + \underset{(2.730)}{0.159}\,x_5 + \underset{(2.610)}{0.003}\,x_9 + \underset{(4.830)}{0.326}\,x_{10} + \underset{(14.298)}{2.814} \qquad (8)$$

$$R^2 = 0.896 \qquad S_e = 0.155$$

The function of the exponential form of the models was the most suitable to capture links between the price of various types and brands of irons and their qualitative features. This was among others due to the fact that the exponential models have slightly higher values of the determination coefficients on the level of 0.902 for the model (7) and 0.896 for the model (8).

In both function models, the linear (6) and the exponential (8) one, the same variables proved to be significant, i.e. brand, electric power, water-tank capacity, steam blow, and cable length. The estimation of the model (8) indicates that a recognised brand of an iron affects its price by 32.6% on average (if values of the remaining variables, i.e. qualitative features stay the same). This provides an undeniable proof that the buyers of consumer goods readily accept a significantly higher price for a product if manufacturer, apart from the similar usability features and technical parameters, offer also a nice design, know-how and an extensive and efficient service. Such situation can be observed with regard to the recognised brand companies.

Summary

The problem of implementing quality adjustments in the calculations of consumer prices indices is one of the most difficult aspects of methodology.

The analysis of the construction and estimation of econometric models presented in this paper lead to conclusion that the application of the regression analysis method to capture the influence of qualitative components of products on their prices yields positive results. This being the first phase of the survey, the next one is aimed at constructing models capable of demonstrating the influence of qualitative changes on the price variation of other household appliances.

Bibliography:

1. Goldberger A. S. (1972), *„Teoria ekonometrii”*, PWE, Warszawa.
2. Paw³owski Z. (1978), *„Ekonometria”*, PWN, Warszawa.
3. Boskin et al. (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.*

4. Kokoski M., Waehrer K., Rozaklis P. (1999), „*Using Hedonic Methods for Quality Adjustment in the CPI: The Consumer Audio Products Component"*, Division of Price and Index Number Research U.S. Bureau of Labor Statistics, Washington.
5. Moulton B. R., La Fleur T. J., Moses K. E. (1999), „*Research on Improved Quality Adjustment in the CPI: The Case of Televisions"*, Bureau of Labor Statistics, U. S. Department of Labor Washington.