STATISTICS SWEDEN
Economic Statistics
Anders Norberg

QUALITY ADJUSTMENT IN THE SWEDISH PRICE INDEX FOR CLOTHING[1]

0 Summary

The method for computing the price index for clothing has been revised in two stages since 1990. Large resources has been laid on a system for data collection and computing of a hedonic price index. The method has been in use since 1994 and on average the hedonic adjustment for quality differences has given index numbers that are about one percentage unit higher than an index computed on all data without quality adjustment. The method for quality adjustment used before 1994 gave index numbers that were generally lower than an index computed on all data without quality adjustment.

The price collectors tend to choose varieties that are new, still having the regular price. This imperfect sampling would cause a significant downward bias, about two percentage units annually for clothing. A simple and unsufficient method to reduce this bias has been used since 1991 and from September 1993 the bias is eliminated with a new method.

1 Introduction

The measurement of price changes for clothing items in the Swedish Consumer Price Index, CPI, involves a number of methodological solutions that makes the survey more complex and manpower-demanding relative to the weight than most other surveys in the CPI. Considerable problems arise as the sampling of varieties (garments) in the outlets is not done by probability sampling, as in practice the whole assortment of clothes is substituted at least once a year, as there are different assortments in summer and winter, as some of the most important physical properties are difficult to measure and as the consumers' appraisals of the properties of the varieties to a large extent are emotional rather than rational.

An annual index for clothing computed without quality adjustment 1984-86 would have been on average 108,7, while the official index, with quality adjustment, was only 104,3. Data show a large quality increase during these years which is not fully trustworthy. At that time quality differences were evaluated by the price collectors.

A first stage in a revision of the Swedish CPI for clothing was taken 1990/91 when a new sampling design and new price collection forms were introduced. The new design would fit the use of a hedonic price index method. The price collectors no longer made the quality evaluations, but collected information on brands, manufacturing countries and physical properties, which were used centrally to decide whether a substitute could be treated as comparable or non-comparable to the variety in the base period. This method was in use in 1991-1993. The files of data were meanwhile used in pilot studies with hedonic indices. In 1994 hedonic indices were computed for 14 of 24 item groups and 1995 hedonic indices are computed for all clothing items.

The Swedish CPI is a chain index with annual links. December is the price reference period. For clothing a major part of the sample of varieties is renewed in December each year. One

[1] Paper prepared for the Second Meeting of the International Working Group on Price Indices, Stockholm, 15-17 November 1995. Printed 1995-10-24.

problem is that the price collectors tend to avoid varieties that have sales prices in the new sample, but they cannot avoid them in December twelve months later. This means that the index would be downwards biased. To reduce this bias, since 1991 the sample of varieties are chosen in November even though prices are collected for the first time in December. The object is to get a representative proportion of sales prices in the price reference period sample. This method has turned out to be insufficient and since September 1993 a method which estimates the bias and eliminates it is used.

In Norberg (1993) the considerations that have been taken before using the hedonic approach for adjusting for quality differences is presented. The first country to implement hedonic methods for clothing in its CPI was the U.S. as reported by Armknecht (1991). Methods and empirical findings are further discussed in Armknecht & Weyback (1989) and Liegey (1990).

2 The consumers

Clothes serve many different needs, and the consumers appraise their properties according to different kinds of yardsticks. With some consumers or in certain purchase situations the comfort, confidence and safety that a garment gives its owner by its suitable or aesthetic cut and colour is given priority. In other cases care is appreciated. There are also consumers for whom it is important to identify themselves with a certain lifestyle or accentuate membership in a group by their choice of garments.

Many consumers lack the knowledge to discriminate and estimate the properties of a garment that determine how well it will function in a certain respect. For instance, the resistance to wear, the crease-resistance of the fabric, the quality of the sewing or the shape after use for some time. The consumers' experience from former purchases is often of a secondary importance for the purchase decision, as the same articles rarely can be found in the market two or more seasons. In many situations, the purchase decision therefore is based on information from other sources, for instance from acquaintances and not the least the seller. In markets, where the consumer acts with uncertainty, the experience of a certain brand or outlet or the picture of the brand/outlet that the seller has succeeded to create in the customer's mind, can be decisive for the purchase.

3 The retail trade

Clothing accounts for 4,7% of the total private consumption according to the CPI-weights. There are two main seasons; March - August and September - February. The main delivering during the year is in March and August/September. The two main seasons have subseasons as early and late spring, early and late summer etc. The patterns of seasons varies slightly from the south to the north, but the adaptation to temperature and weather is negligible. There are also special seasons for certain item groups, one example is the Christmas trade which is a special season for ladies under-wear and night-gowns. Sales are special events which traditionally occur at special periods, for example in the days between Christmas and New Year's Eve.

At the deliverers' level there are many sellers and many buyers. The extensive offer leads to many similar substitutes for every article. The fixing of prices can be considered to work well in good competition, at least for that part of the assortment and selling, 70 percent or more, which cannot be said to be proprietary articles.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

In the retail level there are many sellers as well as many buyers, but in general the buyers lack an overview of the products and the market. In such markets the price, theoretically, gets secondary importance as a means of competition. There are, however, submarkets with hard price competition. The offer of cheap as well as expensive articles is extensive.

The retailer calculates a selling price for an article before he orders it from the producer/ wholesale house. When the clothes are delivered about six months later it will often be necessary to revise the calculation, considering the present fashion and the competition from other retailers. The retailer now fixes his first regular consumer price, as high as he judges the main part of the circle of customers to be willing to pay, considering other offers in the market. The retailers' percentile mark-up, before any selling off, varies mostly between 50% to 200%, tax excluded. On average the percentile mark-up in 1989 was fully 100%, a fraction more for fashionable articles such as blouses, skirts and dresses and a little less for more standardised articles as trousers and coats. In general, the retailers apply a higher percentile mark-up on cheap articles than on expensive articles.

As the demand for style, colour and pattern is difficult to foresee in the regular pricing, the retailer must calculate the risk of selling some part of the assortment at a lower sales price. His skill in judging the demand guides him towards an optimal regular price and towards the right moment to reduce the price for unsold garments of the season. The price reduction can be done stepwise, starting at a reduction of 10% as a "special price", and then going on to cuts of 20, 30 and 50% in the "sales price" and finally even more. For the CPI-data for women's clothes in 1991/92 then median of price reductions was about 40%, i.e. the reduced price was on average about 60% of the regular price, with a large variation. The retail companies also buy special assortments for the traditional sales.

4 Sampling and preparation of data

The Swedish CPI is a chain index with annual links and December as price reference period. In the annual link year t, price indices for different commodities are aggregated by weights, obtained mainly from the National Accounts and the Family Expenditure Survey, which are estimates of consumption values for the entire year t. The last link in the chain, which measures changes from December year t to the current month year t+1, uses weights that are estimates of consumption values in the year t too. For the price reference period new or updated samples of outlets and varieties are selected.

Statistics Sweden's Business Register, containing enterprises and local units, is used as a sampling frame of outlets. The frame is stratified into three strata for clothing items, department stores, super markets and specialised shops. A rotated sample of about 60 outlets with probabilities proportional to size is selected in the autumn of each year. A survey of the assortment in the selected outlets is made in October before the price reference period.

A non-probability sample of 24 item groups was selected 1990, based on statistics from the Family Expenditure Survey. The sample of item groups covers men's, women's and children's outer- and underwear, light and heavy ready-made clothes and sports clothes.

The samples of outlets and item groups are co-ordinated in such a way that a subsample of item groups is measured in each outlet. The number of item groups varies between the outlets, but is at most 15, while the number of surveyed outlets per item group is 20- 25.

The price collectors, interviewers from Statistics Sweden, choose 4-5 varieties of each item group in each outlet. The very first selection is made in November, one month before the base period December. In the sampling of varieties the price collectors shall choose varieties within the definition of the item group, that are different with respect to at least one important property, for example brand, country of manufacture, material or style. The price collectors shall select varieties that can be expected to be representative for the selling in the outlet.

The price collectors shall try to find the selected varieties once a month. Varieties that are sold out shall substituted. A variety in the same outlet that is much sold and reasonably comparable with the one from the price reference period (a variety selected in November can have been substituted in December) is then to be selected. Primarily the substitute should have the same brand as the original variety. If the price collector cannot find a suitable variety with the same brand a variety with a brand representing the same quality level should be selected. Among these the variety with properties that correspond best with the properties of the original variety shall be selected. The properties are listed in a priority order in the price collection forms. One form has been constructed for each item group.

Statistics Sweden utilises the service of the staff of the largest fashion magazine in Sweden to group brands into "status classes".

5 The quality concept

In this paper the quality concept is given a broad interpretation[2]:

i) quality of outlets, i.e. service, assortment, location etc.,

ii) origin of garment i.e. brands, and to some degree manufacturing country,

iii) physical properties of the garment,

iv) the usefulness and/or pleasure of a garment in different time periods of the season for which it is aimed, including the sales period. The quality in this respect is highest in the beginning of the season and declines towards the end. Lack of fit to season can be difficult the measure but it is indicated in collected data by a sales price.

The general rule is that structural changes in the population of outlets are disregarded in the index by using the chaining technique with annual links and new samples of outlets. Changes within the individual outlets during the year are however mostly accepted with no adjustment.

By experience, it is known that the variable brand is highly significant for explaining price levels of clothing. On average there is also a considerable difference in price levels between manufacturing countries, even for such garments whose brands are unknown.

---

[2]Regarding the usefulness and/or pleasure of a garment at the time being in relation to the season for which it is aimed as a quality component is the view of the author and does not necessarily reflect the policy of Statistics Sweden.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

It is a delicate task to decide which physical properties are to be treated as quality and which are to be treated as fashion details. There are some variables as presence of lining, wool and long sleeves that are stable and important characteristics of high quality. One also wants to include variables that separates working clothes from everyday clothes from festivity clothes.

In forced substitutions, when a variety is out of stock and is replaced by another much sold variety, the quality differences regarding origin and physical properties are treated with the use of the hedonic method.

The fourth quality aspect, the usefulness and/or pleasure of a garment in different time periods of the season for which it is aimed, including the sales period, have come into consideration in three situations since September 1993. Firstly the CPI-change from one month to the other and from one year to the other shall reflect different volumes and/or different markdowns in price and the method is generally to compute index on actual prices. The usefulness/pleasure of a garment in different months is not considered to be a quality aspect to be adjusted for in the index. Secondly, in December, when the present link ends and a new link starts, the proportion of varieties that have sales prices is generally lower in the new sample than in the old sample. In this case one could say that "sales" is treated as a quality characteristic and an adjustment is made in the computation, see chapter 8. Thirdly, regular and actual prices are alternatives in the choice of the dependent variable of the hedonic function.

Problems with quality differences arise from changes in the market and from imperfections in the way data are selected for the CPI-computation. The latter reason seems to require the largest adjustments in the Swedish CPI for clothing.

6 The hedonic index

For practical reasons some restrictions were initially put up for the hedonic index. The hedonic functions, i.e. the regression models, should be as similar as possible for all the 24 item groups, the models should be simple enough to be estimated using a statistical program package on a PC, the parameters of the hedonic function must be estimated at most once a year and only CPI-data can be used for estimation of the regression parameters.

6.1 Dependent variable in the model

The hedonic functions can be estimated with either regular price or actual price as the dependent variable. If actual price is used as dependent variable, "being out of season", indicated by a sales price, should in this situation be treated as a quality variable, thereby making possible better estimates of the effects of the quality variables brands and physical properties. The explanatory variables should therefore be accompanied by a variable for type of price.

In the Swedish application the hedonic functions are estimated using regular prices, thus omitting a lot of variation in the dependent variable to be explained. Empirical studies on 1991-92 data have also shown that the regression coefficients for all other explanatory variables than type of price are very similar whether estimated with the regular price or the actual price as the dependent variable., and the effect on price index is negligible.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

6.2 Explanatory variables in the model

Omitting "out of season", indicated by a sales price, as a quality variable the explanatory variables are grouped into outlet type, origin (brand and manufacturing country) and physical properties that are not essentially fashionability details. The estimated price effects of variety properties are exaggerated, if the model does not contain a set of outlet variables since there is a correlation between quality of shop and average quality of assortment.

The 1 600 brands (some of them misspelled) that appear in the CPI-data for clothes are classified by the classes 0-5, where the 0-code represents varieties without any brand at all, and 1-5 different "status" as valued from a customer's point of view. Unknown brands are classified as "missing value". Varieties with unknown brands are classified by manufacturing country.

The price collection forms describe the varieties in many details. The information can be found either on the labels of the garments, i.e. manufacturing country, fibre contents, washing instructions or by inspection, i.e. lining, length, pattern, number and type of pockets etc.. Unfortunately it is not possible to analyse different qualities of cotton and other fibres which may have a large impact on the production cost and finally on the consumer price.

Though fibre content is measured as a percentile share from 0 to 100% it is made dummy variables in the model. A continuous relation between price level and fibre content, when the latter varies from 0 to 100, is not realistic to assume. As regards wool for instance, the price of a garment is high as soon as there is any wool in the material, compared to say 100% cotton. Furthermore, the "best" wool-garments often contain some synthetic fibre to make the garment easier to handle meaning that 85% wool can be better than 100%. Hereby all explanatory variables are dummy variables

A systematic search of good explanatory variables has been done in the results of regression analysis that has been carried out by the Swedish Price and Competition Board in the 1980's, the pilot studies on CPI 1991-1993 at Statistics Sweden and findings from the US Bureau of Labour Statistics. A few variables in this list have been especially important in all or most analyses, for example the presence of a lining, while other variables are significant in some analyses and not important in others.

6.3 Data for estimating the hedonic function

Each month about 4 varieties are observed in about 20-25 outlets, which gives about 80-90 observations for each item group. In 1994, with monthly measurements, altogether 120-420 different varieties were encountered because of forced substitutions. The turnover rate is high, especially for ready-made clothes like women's blouses and dresses. Fewer data are available for four item groups of winter clothing. To these data are added the 80-90 observations from the sample of the price reference period of the next year's link

When the parameters of the regression model are to be estimated with data from different time periods the price change during the year must also be taken care of. Twelve new parameters in the model would be necessary to do this fully, but in times of low inflation fewer parameters suffice.

Assuming that the outlet effects are equal for two or more item groups, for example women's dresses, skirts and trousers, one large model for all these item groups together can be estima-

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

ted, which saves degrees of freedom, i.e. uses data more effectively. Estimating parameters this way should not be done for very different item groups as the fixing of prices within an outlet can vary from one item group to another.

6.4 The regression model

The model of pricing starts with the outlet. The outlets have different locations and give different services and they therefore have on average different percentage mark-ups. Within the outlet there are varieties of different origins with different price levels, demanded by different groups of consumers. Given an outlet and a brand, material and details of the clothing have specific production costs which the retailer and the consumers are willing to accept. The conclusion of these arguments is that the effects of outlet, origin and physical properties should be multiplied when explaining the price variations.

The regular price level for a specific variety i is described as the product of

- a mean price for the entire market, a,
- multiplied with a factor for the group of outlets where the variety is sold, which is one of the eight factors b_k, k=1-8,
- multiplied with a factor for the origin which is one of the eight factors c_k, k=1-8
- multiplied with a subset of factors for the physical properties of the variety, i.e. a subset of the factors $d_1 - d_7$,
- finally multiplied with a time factor e_k, k=1-4, where k denotes the time period when this variety was first entering the survey.

Let V_{ik} denote a dummy variable which corresponds to the coefficient b_k for the variety i, k=1-8, i.e. V_{ik} takes the value 1 if the variety i is measured in outlet group number k and 0 otherwise. Let in the same way X_{ik}, k=1-8, denote origin, Y_{ik}, k=1-7, denote physical properties, and let Z_{ik}, k=1-4, denote time of data collection of the variety i. As one and only one of the eight V_{ik}, k=1-8, takes the value 1 and the rest of them are 0, only seven of them can be in the model. Analogously seven X_{ik} and three Z_{ik} can be in the model. The model can now be written:

$$\mathrm{Price}_i = a\prod_{k=1}^{7} b_k^{V_{ik}} \prod_{k=1}^{7} c_k^{X_{ik}} \prod_{k=1}^{7} d_k^{Y_{ik}} \prod_{k=1}^{3} e_k^{Z_{ik}} \qquad (1)$$

Let $a = \exp(\alpha)$, $b_k = \exp(\beta_k)$, $c_k = \exp(\gamma_k)$, $d_k = \exp(\delta_k)$, $e_k = \exp(\varepsilon_k)$. The model (1) can now be written as

$$\mathrm{Price}_i = \exp(\alpha)\prod_{k=1}^{7} \exp(\beta_k V_{ik}) \prod_{k=1}^{7} \exp(\gamma_k X_{ik}) \prod_{k=1}^{7} \exp(\delta_k Y_{ik}) \prod_{k=1}^{3} \exp(\varepsilon_k Z_{ik}) \qquad (2)$$

Taking the logarithm of both sides of equation (2) gives the hedonic function

$$\mathrm{Log}(\mathrm{Price}_i) = \alpha + \sum_{k=1}^{7} \beta_k V_{ik} + \sum_{k=1}^{7} \gamma_k X_{ik} + \sum_{k=1}^{7} \delta_k Y_{ik} + \sum_{k=1}^{3} \varepsilon_k Z_{ik} \qquad (3)$$

The parameters of this additive linear model can be estimated in a multiple linear regression analysis. This form of the model is known as the semilogarithmic model in the literature of

hedonic price indices. The parameters of equation (1) are calculated using the relations $a = \exp(\alpha)$, $b_k = \exp(\beta_k)$, $c_k = \exp(\gamma_k)$, $d_k = \exp(\delta_k)$, $e_k = \exp(\varepsilon_k)$

In practice the 24 items groups are divided into seven homogeneous groups consisting of a varying number of item groups. The hedonic coefficients for all the item groups within such a group are estimated simultaneously, letting the outlet coefficients b_k, k=1-8 and the time coefficients e_k, k=1-4 be the same for all item groups. Instead one dummy variable for each of the item groups but one must be added, representing different mean prices. Say that we have a group of G item groups and that variety i belongs to item group g. Denote the new parameters a_j, j=1-G and $a_1 = 1$ as there is an overall mean a in the model. Let the variable $U_{gi}=1$ and the variables $U_{ji}=0$ for all $j \neq g$ for all varieties in the g:th item group. The price for item i belonging the g:th item group can then be formulated as

$$\mathrm{Price}_i = a\prod_{j=1}^{G} a_j^{U_{ij}} \prod_{k=1}^{7} b_k^{V_{ik}} \prod_{j=1}^{G}\left[\prod_{k=1}^{7} c_{jk}^{X_{ijk}} \prod_{k=1}^{7} d_{jk}^{Y_{ijk}} \right] \prod_{k=1}^{3} e_k^{Z_{ik}} \qquad (4)$$

U_{ij}, X_{ijk} and Y_{ijk} equal 0 for all $j \neq g$. The parameters c and d and variables X and Y are now indicated by j for item group and k for explanatory variable k.

This technique to reduce the number of different parameters is also used for estimating common effects of origin for women's winter coats and women's winter jackets for which there are fewer data and seemingly problems to find a good set of explanatory variables. Common origin effects are also postulated for two groups of women's jumpers and three groups of children's clothing.

6.5 The type of hedonic index

Generally there are a number of possible methods for utilising the information of the estimated regression coefficients in a price index. Dalén (1992b) gives an outline of methods. In this application there are three circumstances that limit the number of alternatives. The elementary aggregate is a so called RA-index, see formula (6). The regression model is multiplicative with dummy variables only. The regression coefficients cannot be estimated each month, at most once a year for cost reasons. Due to these conditions an adjustment index is the natural choice. This means that either the price reference period or comparison period price shall be adjusted, both methods giving the same result here. In the clothing application the price from the comparison period is adjusted to be comparable to the price reference period price.

Let p_i^t denote the actual price for a variety i in item group g. We have observed seven variables of the variety at time t concerning its origin, $x_{ig1}^t - x_{ig7}^t$ and seven variables concerning physical properties $y_{ig1}^t - y_{gi7}^t$. The variety at time 0 (December last year), that corresponds to or is identical to this one, has the observed values $x_{ig1}^0 - x_{ig7}^0$ and $y_{ig1}^0 - y_{ig7}^0$.
The adjusted price at time t is

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

$$p_i^{t*} = p_i^t \prod_{k=1}^{7} c_k^{(x_{igk}^0 - x_{igk}^t)} \prod_{k=1}^{7} d_k^{(y_{igk}^0 - y_{igk}^t)} \qquad (5)$$

The coefficients b_k, k=1-7, are not used as we compare prices for varieties within an outlet. The coefficients e_k, k=1-3 are not used as they were included in the regression models only to get a good fit of the model to data. Note also that $x_{ijk}^0, x_{ijk}^t, y_{ijk}^0$ and y_{ijk}^t are 0 for all j≠g.

The elementary aggregate for an item group is computed as

$$I = \frac{\sum_i p_i^{t*} / (p_i^0 + p_i^{t*})}{\sum_i p_i^0 / (p_i^0 + p_i^{t*})} \qquad (6)$$

After quality adjustments all substitutes within the item specification are considered comparable to the price reference period variety. An analysis of estimated variances for the price indices has shown that the variances are smaller when all substitutions are included rather then omitting for example the substitutes that are most dissimilar to the price reference period varieties.

7 The hedonic method in practice

7.1 Price collection

The Swedish production of CPI used to be divided into three parts; the daily necessities system, DNS[3], the local price system, LOPS, and central prices, CP. DNS has the special feature of a random pps-sample of varieties. All substitutions are decided centrally and so is quality adjustments. In LOPS the price collectors shall choose the most sold variety of each item group and the price collectors also have to do subjective quality evaluations for many item groups. The CP is a medley of methods handled by the central staff of Statistics Sweden.

Shoes are examples of LOPS, and clothing was within that system until 1991. The new price collection forms used from 1991 and the special way to handle quality differences in forced substitutions qualified clothing to be a separate system, the fourth.

The sample design is different compared to the design for other item groups. In each outlet four or five varieties are selected for each item. To choose more than one variety was expected to be efficient as the price collectors could collect more prices without walking around in the stores. An initial idea was to collect prices for very few item groups and ten to twenty varieties per item group, but the price collectors said that it would be a problem to remember exactly what varieties they had in their samples. An optimal number of varieties seems to be four or five. The price collectors are happy with the new system as they need not do any subjective quality evaluations.

[3] In Dalén's papers this system is referred to as List price system, LIPS, as prices were collected from price lists until 1992.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

7.2 Data registration and editing

Data are collected in shuttle forms, sent between the central office and the price collectors. Before sending the forms to the price collectors the 2000 forms must be sorted by person and before data registration they have to be sorted by item group. This is a monotonous work and a mistake leads to displeasure by the price collectors or the key puncher.

So far Statistics Sweden have used one software system for key punching, another for data editing and a third for index computation and analysis. This gives rise to extra work and in a near future the number of systems will be reduced. Centrally there is a workload for one clerk about seven days each month. For security reasons a couple of persons need to be familiar with the production system.

7.3 Index computation

In practice less than eight characteristics of origin and less then seven physical properties are used in the hedonic quality adjustment. The adjustment is here illustrated by an example:

This is a forced substitution of a woman's dress in the comparison month. As all explanatory variables are dummy variables they can be expressed as no or yes. Notice that there is at most one "yes" among the dummyvariables concerning origin, while there could be four "yes" for the five variables concerning physical properties. The dummyvariable for medium status brands is omitted, being the reference.

Hedonic variable	Hedonic coef-ficient	Properties of variant in price ref. period	Values of variant in price ref. period	Properties of variant in compa-rison period	Values of variant in compa-rison period	Values of quality differences
No brand or low status brands	0,68	yes	0,68	no	1,00	0,6800
High status brands	1,06	no	1,00	yes	1,06	0,9434
Unknown status, Europe 1	1,07	no	1,00	no	1,00	1,0000
Unknown status, Europe 2 and US	0,86	no	1,00	no	1,00	1,0000
Unknown status, other countries	0,71	no	1,00	no	1,00	1,0000
Lining	1,24	no	1,00	yes	1,24	0,8065
Knitted	0,88	yes	0,88	no	1,00	0,8800
30-100 % wool/flax	1,19	yes	1,19	no	1,00	1,1900
65-100 % cotton	0,81	no	1,00	yes	0,81	1,2346
Two pieces	1,42	no	1,00	no	1,00	1,0000
Sleeveless	0,88	no	1,00	yes	0,88	1,1364
Total quality			0,8092		0,8245	0,9815

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

The quality is higher for the substitute than for the price reference period variety and therefore the price at the actual month is multiplied with 0,9815. The quality difference between the new variety and the one which it supersedes, if this is not the price reference period variety, is not explicitly estimated.

7.4 Updating hedonic coefficients etc.

The hedonic coefficient are updated annually. There is a parallel updating of hedonic coefficients and weights in the CPI. The weights in the long-term-index (December) are based on consumption values during the same year as the price index concerns. The hedonic coefficients are estimated with data collected during the same year.

The strategy is to analyse old data to find the best explanatory variables. There is plenty of time for this explorative analysis. In February, when weights and hedonic coefficients must be computed on new data it is in principle only a question of estimation, without testing different models and sets of independent variables. This is also a question of credibility, not being free to choose how to compute the index numbers. This estimation and the necessary analysis is done during a week or two in the beginning of February.

As the hedonic method has been in use for only a couple of years in Sweden it still is worth while to estimate the hedonic coefficient annually to make possible the analysis of stability over time. If there really was such a stability over time it would not be necessary to do the estimation every year.

8 Adjustment for the Sales-price-effect

The Swedish CPI is a chain index with annual links. Each annual link is almost statistically independent of last year's link as 30% of the outlets are substituted and the price collectors are instructed to select new varieties also in old outlets in the new link's base period if the old varieties are not as representative as other varieties are. The instruction for the price collectors also says that all varieties, regardless of type of price, should be considered when choosing the "most sold" in the price reference period. One can assume, however, that the price collectors prefer varieties not having a sales price, i.e. with a regular price, as these varieties can be expected to be found in subsequent months. For each of the comparison months the price collectors must collect prices for the varieties as long as they can be found, and only a proportion of some 1-30% of the varieties are substituted each month. This means, for December 12 months after the price reference period, that the price collector cannot avoid varieties with sales prices in the same way as in the price reference period. This would lead to a downward bias in the CPI.

Statistics Sweden compute an adjustment coefficient in the following way: December is the last month of last year's link and the price reference period for the next year's link. The price collection is done during the same week. Price collection is done with two samples, partially overlapping. Based on the two samples we make two estimates of the effect of sales and campaigns on the mean price level. Each of the estimates is computed in the same way as an index for price change but ordinary prices are put in as price reference period prices and actual prices are used instead of comparison period prices in the index formula. If there is a purposive avoidance of sales prices in price reference period, the index of the sales-price-effect will be higher for the price reference period sample than for the final December sample.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

The price index for each month is now adjusted with the ratio of the sales-price-index for the price reference period sample and the sales-price-index for the final December sample. This is a simple example with a few varieties:

All prices are collected in December some year. There are four varieties in the survey of the present year and five varieties to be measured the following year beginning in December. With two samples we compute two indices in December for the same phenomena, the effect of sales. We use the same formula as in the elementary aggregate of CPI-computation.

Final month of last year's link			Price ref. month of actual year	
Regular price	Actual price		Regular price	Actual price
199	199		199	199
199	149		239	239
98	78		398	349
595	495		599	599
			899	799
Index (Actual/Regular) = 83,9			Index (Actual/Regular) = 95,1	

The formula for the elementary aggregate in the Swedish CPI is

$$I = \frac{\sum_i p_i^t / (p_i^0 + p_i^t)}{\sum_i p_i^0 / (p_i^0 + p_i^t)}$$

Put the actual prices in the place of the comparison period prices p_i^t and the regular prices in the place of the price reference period prices p_i^0. In this example three out of four varieties had sales prices in the final month of last year's link and the actual prices were on average 83,9% of the regular prices. Analogously the actual prices in the new price reference period sample were on average 95,1% of the regular prices.

The difference between these two estimates of the sales price effect is not only stochastic, but systematic and would lead to a bias in the price index if we link one chain to the other. The reason is that the price collectors tend to chose new varieties, not on sale, in the new price reference period sample. In this example, the price index for every month of the new year must be adjusted up-wards with the ratio 95,1/83,9.

To be very careful, in the Swedish long term index for December, the annual link, we use only those data in the calculation of the sales-price-effect that correspond to prices used in price index calculation for December twelve months later. If there are some missing data (for any reason) in the comparison month December the corresponding data are taken away from the above calculation. Say, in the example above, that the fifth item, costing 799 in the price reference period sample, has no observation of price in December twelve months later because no substitute could be found, then the sales-price-effect calculated on the remaining four observations is 96,8 and the long term index (for December) is adjusted with the ratio 96,8/83,9.

STATISTICS SWEDEN
Economic Statistics
Anders Norberg

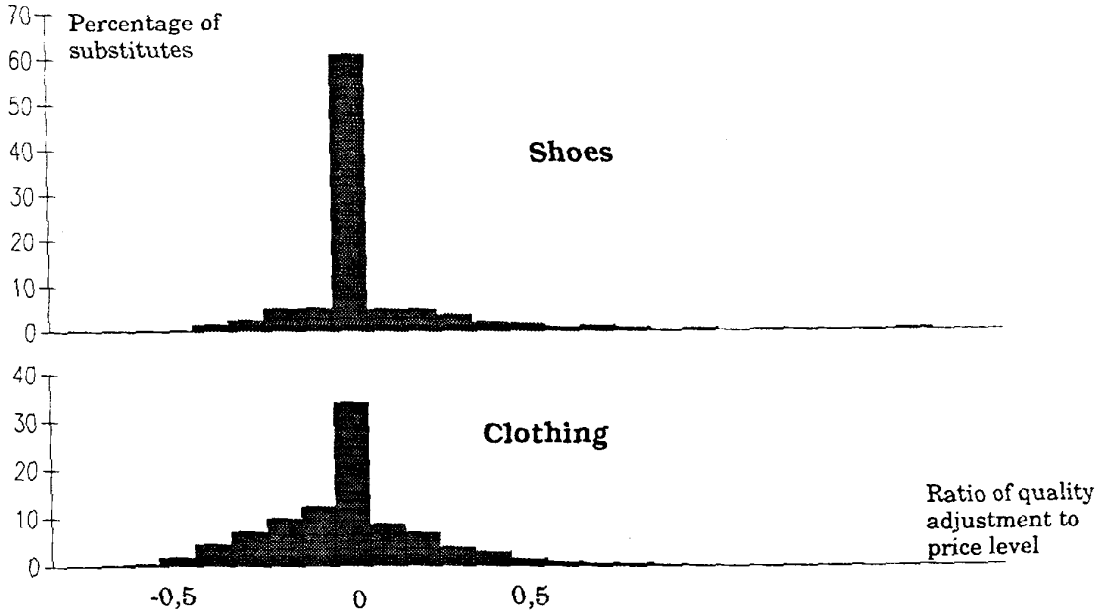
9 Results

Hedonic adjustment for quality differences increased the CPI for clothing with 1,2 ± 1,4[4] percentage points in December 1994, increased the index in September 1995 with 0,6 ± 1,3 percentage points and would have increased the index with 1,6 percentage points in December 1993 if used, compared to an index based on all data regardless of quality differences. In earlier studies with women's clothing 1991 and 1992 the hedonic index was 0,9 and 2,0 higher than an index without quality adjustment. The methods for quality adjustment used until 1990 and 1991-93 generally gave lower indices than an index without any quality adjustment.

Analyses by splitting the estimated variance of the index into two components show that the variances in the estimated regression coefficients give a relatively small contribution to total variance of the index.

The diagrams below show the dispersions of quality adjustments in proportion to a price levels. For shoes the ratio is the quality difference in Swedish kronor evaluated by the price collector to the regular price of the superseded variety. For clothing this ratio is given directly by the hedonic adjustment factor. For about 57% of the substitutions of shoes the quality differences are evaluated as zero, compared to 21% for clothing. The regular price for 10% of the substitutes are equal to the regular price of the substituted shoes, compared to 12% for clothing. There are different asymmetries in the distributions, for shoes there are 17% negative and 24% positive evaluations, but for clothing there are 45% negative and 26% positive evaluations. The predominance of negative adjustments for clothing means that the hedonic index is higher than an index computed on all data without quality adjustment.

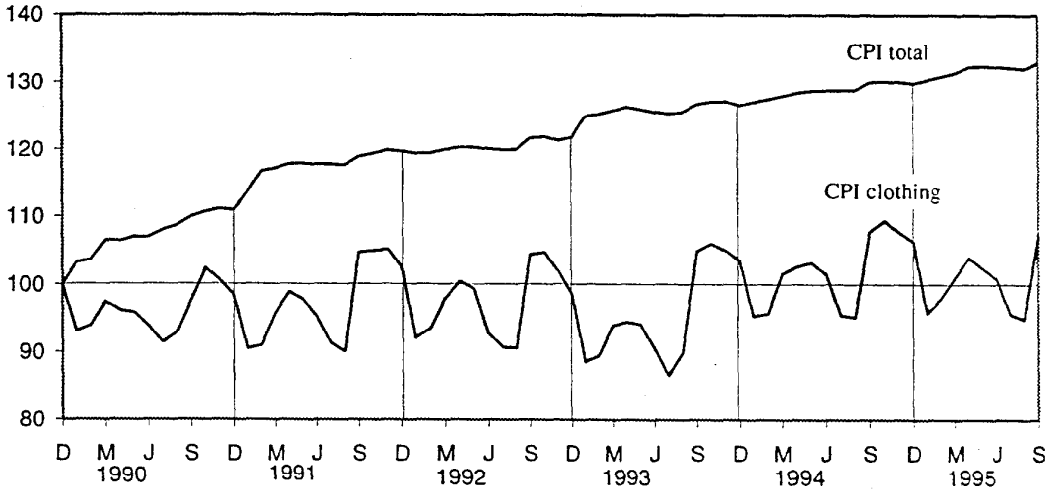
Diagram 1 The distributions of the ratio of quality adjustment to price level for shoes 1993-1995 and clothing 1994



[4] 95 % confidence interval

Empirical studies 1993-1995 have shown that the adjustments for different sales-price-effects have had a significant effect on the price indices. The adjustment increased the Swedish CPI for clothing with 1,3, 2,8 and 2,1 ± 1,7 percentage points respectively for the three years, which means about a 0,1 percentage point increase on total CPI. Similar adjustments have been introduced for shoes in the CPI.

Diagram 2 The Swedish Consumer Price Index for all goods and services and clothing. December 1989 = 100.



10 References

Armknecht, P.A., (1991): "Replication and regressions: Modern Solutions to age old Problems in Price Index Construction", a paper presented at the Conference of European Statisticians in Geneva, 18-21 November 1991

Armknecht, P.A. & Weyback, D., (1989): "Adjustments for Quality Change in the U.S. Consumer Price Index", Journal of Official Statistics, 1989:5 107-123

Dalén, J., (1992a): "Computing Elementary Aggregates in the Swedish Consumer Price Index", Journal of Official Statistics Vol 8 No 2, 1992, pp 129-147

Dalén, J., (1992b): Operationalising a Hedonic Index in an official Price Index Program: Personal Computers in the Swedish Import Price Index, SCB R&D Report 1992:15

Liegey, P.R., (1990): "Quality Adjustment of Apparel Commodities in the CPI", paper presented at Conference on Research in Income an Wealth, Workshop on Price Measurement and Their Uses, Washington 22-23 March 1990

Norberg, A (1993): "Considerations on computing a Hedonic Price Index for Clothing in the Swedish Consumer", a paper presented at the Conference of European Statisticians in Geneva, 25-28 October 1993