# Quantifying Errors in the Swedish Consumer Price Index

Jörgen Dalén

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1993:8

INLEDNING

TILL

**R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.**
**Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.**

## Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

## Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 1993:8. Quantifying errors in the Swedish consumer price index / Jorgen Dalén.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1993-X101OP9308

# Quantifying Errors in the Swedish Consumer Price Index

Jörgen Dalén

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1993:8

| | |
|---|---|
| Från trycket | Oktober 1993 |
| Producent | Statistiska centralbyrån, utvecklingsavdelningen |
| Ansvarig utgivare | Lars Lyberg |
| Förfrågningar | Jörgen Dalén, tel 08-783 44 94 |



ISSN 0283-8680

| | |
|---|---|
| Printed | October 1993 |
| Producer | Statistics Sweden,<br>Department of Research and Development<br>S-115 83 Stockholm |
| Publisher | Lars Lyberg |
| Inquiries | Jörgen Dalén, telephone +46 08 783 44 94 |



ISSN 0283-8680

## QUANTIFYING ERRORS IN THE SWEDISH CONSUMER PRICE INDEX

*by*
*Jörgen Dalén, Department of Economic Statistics, Statistics Sweden*

**Abstract:** In this report a structure is proposed for evaluating, comparing and aggregating the most important errors in the Swedish Consumer Price Index. The structure has the form of a mean square error model, where the error components are given explicitly as well as their estimators. Its application to CPI data for 1981-1992 is given and discussed.

**Contents**

## 1. Introduction

The attempts to measure and aggregate errors of a Consumer Price Index (CPI) are probably as old as the index itself. The first work of which we are aware is in the collected papers of Edgeworth (1925) which is a work first published in 1889. He discusses three types of errors: 1) errors in weights, 2) errors in price relatives and 3) errors resulting from unrepresented product categories. Edgeworth discusses these errors theoretically based on a mathematical model as well as empirically and gives an advice that has since been passed on to new generations of index practitioners:

"Take more care about the prices than the weights." (p. 320)

Since then a couple of papers on this topic have emerged - see Biggeri & Giommi (1987) for a fuller reference list. Biggeri & Giommi themselves provide a detailed error analysis based on the mean square error model. They also give an extensive list of all types of errors present in a CPI. However, when it comes to actual numerical estimation of the error components, only sampling errors are considered and low level sampling variation is disregarded.

In Sweden there is also a long tradition of dealing with the KPI precision problem. Early works are Ruist (1953) and Malmquist (1958), both in Swedish. Ruist approached problems of regional, outlet and item sampling as well as errors due to quality change. Malmquist's main approach was to study the formula error of a Laspeyres index but he also continued Ruists approach with later data. Both reports contain many numerical estimates. Andersson, Forsman & Wretman (1987) studied the variance contribution from outlet sampling for some commodity groups in the KPI.

Dalén & Ohlsson (1993) gives a more complete list of later papers dealing with the CPI variance estimation problem.

In this report a framework for analysing errors in the Swedish Consumer Price Index (Konsumentprisindex, KPI for short) is presented. Our aim is to provide a coherent structure for discussing, comparing and aggregating important errors in a CPI. The analysis is geared to the particular structure and procedures of the KPI and the details of our model may therefore not be generally applicable. But we believe that with suitable modifications our general approach may be applicable to the often similar circumstances that are prevailing in, e.g., other European countries.

A CPI is computed based on several levels of aggregation. From a technical statistical point of view it is composed of a number of surveys covering different areas of private consumption. In the case of the KPI where no regional levels are involved, these surveys have independent designs.

A first division of the KPI errors is therefore to separate those between surveys from those within surveys. Errors between surveys are first those resulting from the necessary operationalisations of the abstract economic concept underlying the CPI. These errors, due to e.g. imperfect index formulae, are not considered here. They are probably best estimated with econometric methods, see e.g. Manser & McDonald (1988) for a study of the substitution bias of a fixed weight Laspeyres price index formula with data from the U.S. CPI.

We will incorporate two types of **errors between surveys** in our model.

1) **Errors in consumption weights**, i.e. the weights used in the aggregation process from item group indices to the all-item KPI. The error considered here is the difference between an index computed with correct weights for each survey and an index computed with estimated survey weights. This error will be modelled as a bias component. Effects of estimated weights within surveys will belong under category 4 below.

2) **Errors due to non-coverage of item groups**. Certain products, mainly services, are not included in the KPI, notably (1992) financial services, public child care and care of the elderly and certain international transport services. This fact will give rise to errors which, in this report, are modelled as a random error giving rise to a variance component.

**Errors within surveys** are also of two basic types.

3) **Sampling errors in the price surveys**. These errors are in practice a result of various sampling procedures, probability sampling as well as purposive selection. In both cases we model the error as a random error giving rise to a sampling variance. In a major part of the KPI the sampling design is of a cross-classified type. Sampling errors are treated extensively in another report - see Dalén & Ohlsson (1993).

4) **Non-sampling errors in the price surveys**. Under this category fall many different kinds of errors. What they have in common is that we will try to assess their likely size by means of various forms of sensitivity analyses. They will be further exemplified in Section 3.2 below. We will model all these errors as being systematic and giving rise to bias components or bias risks.

We will use a mean square error model with both variance and bias components in two stages. We consider four types of errors in our error model.

These four types of errors in principle cover all errors in a CPI once basic definitions and specifications are laid down. Note how closely they correspond to the types of errors that Edgeworth (1925) analysed.

We will also present procedures for measuring these errors. These procedures will be of varying quality - from the scientifically "hard" variance estimators for probability sampling to "softer" measures based on various types of sensitivity analyses for determining the likely size of biases. A major purpose of this report is to try to give a framework for comparing and aggregating these different types of measures. We will, however, not try to actually determine the total error of the KPI - some of the important errors will not be numerically estimated.

We like to stress that our error model is really a **model** and not an attempt to determine the true nature of the errors. The way you build an error model depends on the way the errors are actually measured. Also, what might be naturally seen as a systematic error in a short-term comparison might be better viewed as a random error in a long-term comparison.

In Section 2 we give our formal error model and in Section 3 its application. Proofs will be deferred to appendices 1-2.

## 2. Formal error model

### *2.1 The conceptual model*

We start by introducing a number of index concepts corresponding to different steps in the estimation process. The first logical concept is that of the ideal goal of the index , which would be some kind of aggregated cost-of-living (constant utility) price index for all households. This is, however, an all together abstract and non-operational concept and hence we will not be concerned with it in this analysis.

The next step is the defined goal, which could also be viewed as the operationalized true value of the KPI. This concept is defined in terms of true weights, $w^*_{hj}$, and true subindices, $I_{hj}$, for product groups:

$$I^* = \sum_{h\in U}\sum_{j\in U_h} w^*_{hj} I_{hj} \quad \text{with} \quad \sum_{h\in U}\sum_{j\in U_h} w^*_{hj} \equiv 1 \tag{1}$$

where $h \in U$ is a division into superstrata (more about this concept below) $U_h$ of product groups $j \in U_h$.

Next, we replace the $w^*_{hj}$ with estimated weights $w_{hj}$, with their sum also standardised to unity and obtain:

$$I = \sum_{h\in U}\sum_{j\in U_h} w_{hj} I_{hj} \quad \text{with} \quad \sum_{h\in U}\sum_{j\in U_h} w_{hj} \equiv 1 \tag{2}$$

The difference between (2) and (1) is called the **weight error**. In countries where weights are estimated from a sample survey of household expenditures, the natural thing would be to model this error as a variance component - see e.g. Biggeri & Giommi (1987) or Balk & Kersten (1986) for approaches along this line. In the Swedish KPI, however, the weights are estimated in a different manner by the National Accounts. We therefore choose to model the weight error as a bias component, since we will use a "true value"-technique (Section 3.4 below) for evaluating its size. Thus we define our first error component as:

$$B_1 = I - I^* . \tag{3}$$

The process of estimating **I** could be viewed as consisting of two steps. Step 1 is the selection of product groups reflecting the fact that there are some non-covered groups. This selection is of course highly purposive - the reason that some product groups are not included is usually difficulties in finding reliable and cost-effective measurement methods. For this step we will introduce a quasi-randomisation model in which the covered groups are considered to be a stratified simple random sample of item groups from each superstratum h . In practice we will have a certainty superstratum h=C , where all item groups are covered and a set of sampled superstrata $h \in H$, so that $C \cup H = U$. The "sample" of covered item groups in superstratum h is denoted $S_h$ and $S = \bigcup_{h\in H} S_h$ .

In step 2 we estimate the item group indices $I_{hj}$ with $\hat{I}_{hj}$. In this step we take samples of price observations within the product group $j \in h$ according to various sampling designs. We end up with the estimator

$$\hat{I} = \sum_{h\in U}\sum_{j\in S_h} \tilde{w}_{hj}\, \hat{I}_{hj} \text{ , where} \tag{4}$$

$$\tilde{w}_{hj} = \begin{cases} w_{hj} & \text{if } h \in C \\ w_{hj}\dfrac{w_h}{w_h^S} & \text{if } h \in H, \end{cases}$$
$$w_h = \sum_{j\in U_h} w_{hj} \text{ and } w_h^S = \sum_{j\in S_h} w_{hj} . \tag{5}$$

The expected value of $\hat{I}$ is denoted $I^E$. Expectation is taken with respect to both the quasi-sampling of product groups and to the sampling of price observations. Since $\hat{I}$ has the form of a sum of stochastic ratios (with $w_h^S$ in the denominator) there is no exact expression for $I^E$. We denote by $I^{ER}$ the usual ratio approximation of $I^E$.

In step 2 we allow for a non-sampling bias so that

$$E(\hat{I}_{hj}) = I_{hj} + b_{hj} = I^{B}_{hj} \tag{6}$$

where $b_{hj}$ is the non-random bias induced by the measurement procedure for item group j in stratum h due to, e.g., item and outlet substitutions, imperfect quality adjustments or errors in the price capturing process. We also introduce the symbol $\hat{I}^{B}_{hj} = \hat{I}_{hj} + b_{hj}$ .

The aggregate bias in step 2 is called $B_2$ and we have

$$B_2 = \sum_{h\in U}\sum_{j\in U_h} w_{hj} b_{hj} = I^{ER} - I \approx I^{E} - I \tag{7}$$

Our interest now focuses upon the total error of the estimate in (4). As our measure of total error we choose the mean square error, i.e.:

$$E(\hat{I} - I^*)^2 = E_1 E_2(\hat{I} - I^*)^2 \tag{8}$$

where $E_1$ and $E_2$ stand for expectation in step 1 and 2, respectively. We may decompose (8) into

$$E(\hat{I} - I^*)^2 = V(\hat{I}) + B^2(\hat{I}), \tag{9}$$

V being total variance and B total bias. Further decomposition gives:

$$\begin{aligned} V(\hat{I}) &= V_1 + V_2, \text{ where} \\ V_1 &= V_1(E_2(\hat{I}|S)) \text{ and} \\ V_2 &= E_1(V_2(\hat{I}|S)) \end{aligned} \tag{10}$$

and

$$B(\hat{I}) = B_1 + B_2 \text{ (+ ratio estimator bias)} \tag{11}$$

After some calculations - summarised in Appendix 1 we obtain the following expressions for our four error components:

$$V_1 = \sum_{h\in H} \frac{N_h(N_h - n_h)}{n_h}\sigma_h^2 \quad \text{with } \sigma_h^2 = \frac{1}{N_h - 1}\sum_{j\in U_h} w_{hj}^2 (I^{B}_{hj} - I^{B}_{h.})^2$$
$$\text{and } I^{B}_{h.} = \frac{\sum_{j\in U_h} w_{hj} I^{B}_{hj}}{w_h} \tag{12}$$

$$V_2 = \sum_{h\in H} \frac{N_h}{n_h} \sum_{j\in U_h} w_{hj}^2 V_{hj} \tag{13}$$

$$B_1 = \sum_{h\in U}\sum_{j\in U_h} (w_{hj} - w^*_{hj}) I_{hj} \tag{14}$$

$$B_2 = \sum_{h \in U} \sum_{j \in U_h} w_{hj} b_{hj} \tag{15}$$

where

$N_h$ is the total number of item groups in superstratum h,

$n_h$ is the number of covered (quasi-sampled) groups in superstratum h and

$V_{hj} = V(\hat{I}_{hj})$ is the sampling variance in step 2 for item group $j \in h$.

*2.2 Estimators of error components*

Next we seek the best possible estimates of the four error components defined in (12)-(15) above. In the practical applications we will use the following four estimates.

$$\hat{V}_1 = \sum_{h \in H} \frac{w_h (w_h - w_h^S)}{\left(w_h^S\right)^2} \sum_{j \in S_h} w_{hj}^2 (\hat{I}_{hj} - \hat{\hat{I}}_{h.})^2 \ \text{ with } \ \hat{\hat{I}}_{h.} = \frac{\sum_{j \in S_h} w_{hj} \hat{I}_{hj}}{\sum_{j \in S_h} w_{hj}} \tag{16}$$

$$\hat{V}_2 = \sum_{h \in U} \sum_{j \in S_h} \tilde{w}_{hj}^2 \hat{V}_{hj}, \text{where } \hat{V}_{hj} \text{ is an unbiased estimate of } V_{hj} \tag{17}$$

$$\hat{B}_1 = \sum_{h \in U} \sum_{j \in S_h} (\tilde{w}_{hj} - \tilde{w}_{hj}^*) \hat{I}_{hj} \quad \text{with} \quad \begin{array}{l} \tilde{w}_{hj}^* = \begin{cases} w_{hj}^* & \text{if } h \in C \\ w_h^* w_{hj}^* / w_h^{S*} & \text{if } h \in H, \end{cases} \\ w_h^* = \sum_{j \in U_h} w_{hj}^* \ \text{ and } \ w_h^{S*} = \sum_{j \in S_h} w_{hj}^*. \end{array} \tag{18}$$

$$\hat{B}_2 = \sum_{h \in U} \sum_{j \in S_h} \tilde{w}_{hj} b_{hj} \tag{19}$$

In appendix 2 the properties of these estimated error components are derived. It turns out that $\hat{V}_2$ and $\hat{B}_2$ are approximately unbiased (first order Taylor linearisation). $\hat{B}_1$ and $\hat{V}_1$ are, however, biased and we have the following relations:

$$E(\hat{B}_1) = B_1 + R_B, \text{ where the discrepancy term}$$
$$R_B = \sum_{h \in U} \sum_{j \in h} \left(w_{hj} - w_{hj}^*\right) b_{hj} \text{ and} \tag{20}$$

$$E(\hat{V}_1) = V_1' + R_V, \tag{21}$$

where $V'_1$ differs from $V_1$ only by stratumwise factors $(N_h-1)/N_h$ and $R_V$, a discrepancy term, depends on the step 2 sampling variances in the quasi-sampled superstrata. See Appendix 2 for an exact expression for $R_V$.

We see that $\hat{B}_1$ also depends on the survey biases $b_{hj}$. In principle this dependency would make it possible to adjust $\hat{B}_1$ to obtain an unbiased estimate. This is not practicable, however, for reasons discussed below. For $\hat{V}_1$ to be useful it is necessary that $R_V$ is either small or estimable. We will demonstrate below that both these facts are at hand.

For our purposes of obtaining crude error estimates we will, in Section 3 below, demonstrate empirically that $\hat{B}_1$ and $\hat{V}_1$ are useful, although it is of course necessary to keep the biases in (20) and (21) in mind.

## 3. The practical application of the error model

### *3.1 Sampling errors*

Sampling errors are discussed in Dalén & Ohlsson (1993). Here we will only give a summary of the results from this report.

KPI consists of many price surveys done with different methods and sampling designs including purposive sampling. It is possible to make a decomposition such that these surveys are (at least approximately) independent of each other. Making variance estimates for each one of these we could in the end use (17) in order to arrive at an estimate of total sampling error. In practice such variance estimates have been done indicating a total sampling error for an annual long term link of about

$$V_2 \approx 0{,}04$$

which is equivalent to a 95% confidence interval of about 0.4. Actual computations have been done for all the larger surveys.

### *3.2 Non-sampling errors in price surveys*

As mentioned above there are several sources of bias in price surveys. We mention here the most important types:

1) **Procedural bias**. The procedure used for setting up a price measurement system is always biased to a smaller or larger extent compared with an ideal index definition. Some examples are:
- Procedures for linking in new items and outlets where comparisons are done only within and not between close substitutes. This error is often referred to as (low level) substitution bias.
- The methodological (conceptual) choice on owner-occupied housing according which differ to alarge extent between different countries.
- Biases resulting from using inferior index formulae at the elementary aggregate level.
- Selection bias due to purposive sampling.

2) **Quality adjustment errors**. These errors result from the implicit or explicit adjustments done when an item substitution is done.

3) **Low level weighting errors**. Often accurate weights are lacking at low levels in a CPI aggregation system. We don't know what quantities are bought to bargain prices as opposed to ordinary prices or how many train or air tickets that are bought to reduced prices. Weight used in practice are cruder and sometimes equal weights are assumed.

4) **Errors in the recorded price.**. There could be various reasons for errors and ambiguities in the actual recorded price, from direct mistakes to the use of list prices, discounts, coupons, bonus systems, and other kinds of price differentiation which result in the recorded price being more or less irrelevant to a large part of the consumers.

5) **Traditional forms of non-sampling errors.** Here we refer to non-response and undercoverage in those price surveys where probability sampling is used. Non-response exists for outlets but is usually very small. Undercoverage in sample surveys is a close relative to the selection bias due to purposive sampling above and should normally be a much smaller problem as long as the rate of undercoverage is not extremely large.

In principle we would like to use a formal bias estimation procedure and aggregating biases as in (19) above. In practice this is not often possible but the above formalisation is always useful as a conceptual basis for an intelligent error analysis. Let us start, however, with a case where (19) could be applied in a direct manner.

**Example 1 - formula bias.** In Dalén (1992, Table 1) the effect of using different index formulae for elementary aggregates is demonstrated. If one of these, such as the geometric mean, is taken as an ideal definition , the errors of the others could be interpreted as formula biases . It was shown that the difference between the worst of those index formulae (which was actually used and a better one was as large as 0.65 for a 9 month comparison.

A few other examples of a more indirect use of (19) would be:

**Example 2 - owner-occupied housing.** In the KPI the index for owner-occupied one-family homes are calculated according to an asset cost approach (in the sense of Early, 1990) covering mortgage interest, property tax , maintenance cost etc, while as the index for owner-occupied multi-family dwellings simply follows the index for (multi-family) rented dwellings. Other countries have made entirely different methodological choices in this area. A natural thing is therefore to make a simple sensitivity analysis according to (19) applying two alternative methods instead of the one used. In the following table we show the sensitivity of the KPI for all items of
**ALT1**: letting the index for multi-family owner-occupied housing follow that for one-family owner-occupied housing instead of that of rented dwellings,
**ALT2**: letting all indices for owner-occupied housing follow that for rented dwellings (a crude application of a rental equivalence principle).

The result would be the following for the years 1981-1991. We present the differences betwwen the two two alternative calculations and the actual KPI figure.

*TABLE 1: Bias risks for housing*

| YEAR | KPI | ALT1-KPI | ALT2-KPI |
|---|---|---|---|
| *1981* | 109,4 | -0,10 | 0,46 |
| *1982* | 109,9 | -0,22 | 1,36 |
| *1983* | 109,3 | -0,20 | 1,24 |
| *1984* | 108,1 | -0,04 | 0,25 |
| *1985* | 105,7 | 0,12 | -0,63 |
| *1986* | 103,2 | -0,11 | 0,83 |
| *1987* | 104,9 | -0,14 | 0,86 |
| *1988* | 106,2 | 0,01 | -0,06 |
| *1989* | 106,7 | 0,01 | -0,08 |
| *1990* | 110,7 | 0,03 | -0,16 |
| *1991* | 108,0 | -0,49 | 2,29 |
| *1992* | 101,9 | -0,11 | 0,27 |
| *MEAN* | | -0,10 | 0,55 |

These figures should be interpreted as **bias risks** rather than bias estimates, since there is not a theoretical consensus on which of these alternatives is the most correct. What they show is the sensitivity of the index estimate to different, but not unreasonable, index definitions. The large bias risks for certain years are noteworthy; for 1991 this was due to changed Government rules of taxing and subsidising housing.

**Example 3 - quality changes for clothing items.** In the latter part of the 1980's the KPI methods of measuring price changes for clothing items were reviewed, since there were suspicions of considerable biases. In the old method the interviewers made a subjective evaluation of quality differences in monetary terms whenever (the very frequent) substitutions occurred as the old variety was no longer marketed. It was found that these evaluations to a large extent implied quality improvements and in particular for women's clothing, so that the items with the highest frequency of substitutions also had the smallest price increases. As a measure of the bias risk of the method then used was taken the difference between the index calculated with and without the quality evaluation (QE) of the interviewer. For the years 1984-86 this gave the following result:

*TABLE 2: Price indices and bias risks for apparel items 1984-86*

| YEAR | WITH QE | WITHOUT QE | DIFFER-ENCE | WEIGHT | BIAS RISK FOR KPI |
|---|---|---|---|---|---|
| *1984* | 105.32 | 112.78 | -7.46 | 0.066 | -0.49 |
| *1985* | 105.08 | 109.42 | -4.34 | 0.067 | -0.29 |
| *1986* | 101.80 | 105.00 | -3.20 | 0.072 | -0.23 |
| *MEAN* | 104.07 | 109.07 | -5.00 | 0.068 | -0.33 |

**Example 4 - discounts for petrol.** Most petrol dealers apply some kind of discounts for customers with special cards. These discounts are currently not taken into consideration in the CPI. Provided that we consider inclusion of discounts to be the theoretically preferred procedure the bias entailed by not including them would be (at the time of the introduction of the discount system) the KPI weight of petrol (about 0.04) times the average discount in %. Assuming the discount to be 3% on average would entail a (one-time) bias of 0.12.

### *3.3 Errors due to non-coverage of item groups*

Errors due to non-measured item groups are in our system evaluated according to (16). Of course, in reality there is no random mechanism in the selection of groups to be measured and one possibility might therefore be to model the error as a bias and expose it to various forms of sensitivity analyses. For a certain time period it may be possible to use expert judgement to assess the possible error, but in order to get at the likely long-run error we prefer the more mechanic assumption of quasi-sampling. What we are aiming at, for one thing, is a quantitative measure of "how much it is worth" to develop new measurement systems for these groups and the answer to that question requires a measure relevant for many different years.

Given (16) the next problem is the choice of superstratum structure to represent the non-covered groups which, in 1992, covered about 3.6% of private consumption together. This ought to be done so that the design is "non-informative" in the intuitive sense that the actual outcome of the quasi-sampling procedure is as likely as any other outcome. Since the non-selected item groups are rather small (each has a weight of less than 1%) and we have assumed simple random sampling as our quasi-sampling procedure, the selected item groups of the same superstratum should also be small.

We have tried two solutions in this regard. **Method A** is to divide the whole CPI into two superstrata, one self-representing category of very large item groups (rents, interest for home-owners, petrol, cars and a few more) and one superstratum which covers all non-covered item groups and all small covered item groups. Within method A we have used three different cutoff points in terms of the weight of the item group, 2%, 1% and 0.5% , above which we have considered the group to be self-representing. One drawback of Method A is that some of the representing item groups will exhibit dependencies due to a common outlet sample. **Method B** is to select a particular set of representing item groups for each non-covered group. These sets generally consist of independently sampled item groups. In table 3 we show how this was done:

*Table 3: The representation of non-covered item groups in Method B*

| **Represented group (KPI weight)** | **Foreign air travel (0.49%)** | **Child care and care of the elderly (0.99%)** | **Renters' repairs (0.08%)** | **Furniture repair (0.09%)** | **Other services (1.97%)** |
|---|---|---|---|---|---|
| Representing groups | 1 Domestic air travel<br>2 Packaged tours | Other item groups in the public sector:<br><br>1 Local transports<br>2 Pharmaceuticals<br>3 Medical care<br>4 Dental services | 1 Homeowners' repairs, services<br>2 Homeowners' repairs, goods | 1. Repairs of washing machines<br>2 Domestic services<br>3 Auto repairs (5 services)<br>4 Moving<br>5 TV repair<br>6 Photo developing<br>7 Barbers' services<br>(3 services) | 1. Vehicle inspection<br>2. Driver's education<br>3. Garage rent<br>4. Train fares<br>5. Domestic boat fares<br>6. Telephone services<br>7. Postage |

The result of the computations for the years 1980-91 according to (16) is the following:

*Table 4: Estimates of variance contribution for non-covered item groups*

| **YEAR** | **METHOD A:** | | | **METHOD B** |
|---|---|---|---|---|
| | **CUTOFF POINT** | | | |
| | **0.5%** | **1%** | **2%** | |
| *1980* | 0.0067 | 0.0116 | 0.0159 | 0.0232 |
| *1981* | 0.0079 | 0.0119 | 0.0135 | 0.0062 |
| *1982* | 0.0033 | 0.0054 | 0.0090 | 0.0102 |
| *1983* | 0..0038 | 0.0060 | 0.0075 | 0.0016 |
| *1984* | 0.0038 | 0.0036 | 0.0051 | 0.0064 |
| *1985* | 0.0029 | 0.0038 | 0.0044 | 0.0045 |
| *1986* | 0.0077 | 0.0207 | 0.0184 | 0.0009 |
| *1987* | 0.0115 | 0.0088 | 0.0207 | 0.0022 |
| *1988* | 0.0050 | 0.0044 | 0.0062 | 0.0021 |
| *1989* | 0.0046 | 0.0055 | 0.0058 | 0.0112 |
| *1990* | 0.0081 | 0.0118 | 0.0170 | 0.0106 |
| *1991* | 0.0325 | 0.0289 | 0.0278 | 0.0248 |
| *MEAN* | 0.0082 | 0.0102 | 0.0126 | 0.0087 |

We see, that although results for a certain year vary quite a lot between different methods, the long-run means are not much different.. As a rough figure we could take $V_1$ =0.01 as a long-term order-of-size estimate. Of course, the differences between years reflect genuine changes in the sense that item group variations are larger in certain years. For example, in 1991 the high figures reflect differential changes in value-added taxes and other taxes and subsidies, which of course increase errors from leaving certain item groups out of the index.

What about the discrepancy term $R_V$ in (21)? First we note that $R_V$ will in practice always be positive which means that (16) is likely to be an overestimate of $V_1$. $R_V$ is easiest to evaluate for Method B, for which crude, but fairly reliable calculations indicate that it is of a smaller order of size than $V_1$ and not more than 10%. Less reliable, but not unreasonable, calculations indicate something similar for Method A.

*3.4 Bias due to preliminary weight estimates.*

Bias estimates according to (18) were made by taking the final National Accounts (NA) estimates for private consumption for the years 1986, 1987 and 1988 to be the "true" weights. These final estimates are made about three years after the reference year with the use of final estimates from a multitude of primary statistical sources while as the preliminary NA figures underlying the KPI weights use preliminary estimates and projections of various kinds. The bias estimates obtained in this way were: -0.23 for 1986, -0.14 for 1987 and -0.12 for 1988 with an average of -0.16.

Here the discrepancy term $R_B$ in (20) is potentially disturbing. It might be that the estimates of weight bias is in fact contaminated by step 2 biases. However, for $R_B$ to have any long run influence on the bias estimates there must be a strong correlation between the step 2 biases and the weight differences. With the crude use that we make of these estimates, we believe that calculations according to (18) does not generally lead to erroneous conclusions. It is, however, necessary to look in particular at the weight differences for those products for which the step 2 bias risks are known to be significant.

*3.5 Total error*

It would be nice to be able to sum up the errors according to (9)-(11) to a total error measure. Based on the empirical presentation above this is not advisable, however. It might be possible to suggest $V \approx 0{,}04$ as a rough measure of total variance for a one-year link, corresponding to a 95% confidence interval of $\pm 0{,}4$, since the variance components seem to be fairly stable over time. But the bias risks are extremely variable from year to year and could even change sign. It is also evident that these risks are often larger than the random error. This is particularly true if the effect of the choice of method for owner-occupied housing is taken into account.

*3.6 Errors for longer periods*

This analysis has, so far, been concerned with errors in the one-year link. Errors for shorter time periods are, in addition, influenced by seasonal variations in consumption and prices and the way these variations handled in the KPI and an analysis of short term errors must therefore take account of this aspect also. We will not attempt such an analysis here.

Errors for longer periods of, say, several years are of considerable interest. Sampling errors are affected by the correlation structure between years. This structure is very complicated and only a crude assessment is possible. Both the product sample and the outlet sample have large degrees of overlap between years. But for large covariances to occur it is also necessary either that outlets which increase the prices more than average a certain year also increases their prices more another year or that a certain product which shows a price increase more than the average of its product group a certain year also tends to have the same kind of price movement the next year. Neither of these tendencies are likely to be at hand.

If there are no large positive correlations between years in the sampling system, sampling errors will tend to be less important for long-term comparisons. The interest will therefore focus upon the existence of systematic errors with a tendency to have the same sign over a long period. Possible such errors are

- substitution biases due to the procedures for linking in new products and outlets,
- errors because of incomplete or imperfect quality adjustment,
- conceptual difficulties such as those for owner-occupied housing or

- failure to capture the price actually paid, due to the use of list prices, discounts, coupons etc.

An important task for applied index research in the future will be to determine the likely sizes of these errors.

## 4. References

Andersson, C., Forsman, G., & Wretman, J. (1987). On the Measurement of Errors in the Swedish Consumer Price Index. *Bulletin of the International Statistical Institute*, 52, Book 3, 155-171.

Balk, B., & Kersten, H.M.P. (1986). On the Precision of Consumer Price Indices Caused by the Sampling Variability of Budget Surveys. *Journal of Economic and Social Measurement*, 14, 19-35.

Biggeri, L., & Giommi, A. (1987). On the Accuracy and Precision of the Consumer Price Indices, Methods and Applications to Evaluate the Influence of the Sampling of Households. *Bulletin of the International Statistical Institute*, 52, Book 3, 137-154.

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

Dalén, J.& Ohlsson, E. (1993). Variance estimation in the Swedish Consumer Price Index. *R&D Reports, Statistics Sweden.*

Early, J. (1990). Improving Consumer Price Indexes. *Journal of Official Statistics*, 6, No 2, 179-204..

Edgeworth, F.Y. (1925). Papers relating to Political Economy, Vol 1, 298-341. *London: The Royal Economic Society.*

Malmquist, S. (1958). Precisionsproblemet vid beräkning av konsumentprisindex I och II (in Swedish). *Statistisk Tidskrift*, 7, 492-504 and 612-628.

Manser, M.E. & McDonald, R.J. (1988). An Analysis of Substitution Bias in Measuring Inflation, 1959-85. *Econometrica,* 56, No 4, 909-930.

Ruist, E. (1953). Några synpunkter på socialstyrelsens prismaterial. In *SOU 1953:23 Konsumentprisindex*, Stockholm.

## Appendix 1

We will demonstrate the following two results:

$$1)\ V_1 = V_1\left(E_2(\hat{I})|S\right) \approx \sum_h \frac{N_h(N_h - n_h)}{n_h} \sigma_h^2$$

$$2)\ V_2 = E_1\left\{V_2(\hat{I})|S\right\} \approx \sum_h \frac{N_h}{n_h} \sum_{j \in U_h} w_{hj}^2 V_{hj}$$

The approximations are always of the ratio approximation type. We disregard the certainty superstrata in the following:

$$1)\ V_1\left(E_2(\hat{I})|S\right) = V_1\left( E_2(\sum_h \sum_{j \in S_h} \tilde{w}_{hj} \hat{I}_{hj}) \right) = V_1(\sum_h \sum_{j \in S_h} \tilde{w}_{hj} I_{hj}^B) = \sum_h w_h^2 V_1 \left\{ \frac{\sum_{j \in S_h} w_{hj} I_{hj}^B}{\sum_{j \in S_h} w_{hj}} \right\} \approx$$

$$\left\{ \begin{array}{l} \text{According to the usual ratio} \\ \text{approximation, see e.g. Cochran} \\ \text{(1977, formula 2.39) with} \\ y_{hj} = w_{hj} I_{hj}^B \ \text{and} \ x_{hj} = w_{hj} \end{array} \right\} \approx \sum_h w_h^2 \frac{1 - \frac{n_h}{N_h}}{n_h \frac{w_h^2}{N_h^2}} \frac{\sum_{j \in U_h} (w_{hj} I_{hj}^B - w_{hj} I_{h.}^B)^2}{N_h - 1} =$$

$$= \sum_h \frac{N_h(N_h - n_h)}{n_h} \frac{\sum_{j \in U_h} w_{hj}^2 (I_{hj}^B - I_{h.}^B)^2}{N_h - 1} \qquad \text{with } I_{h.}^B = \frac{\sum_{j \in U_h} w_{hj} I_{hj}^B}{\sum_{j \in U_h} w_{hj}}$$

$$2)\ E_1\left(V_2(\hat{I})|S\right) = E_1\left( V_2(\sum_h \sum_{j \in S_h} \tilde{w}_{hj} \hat{I}_{hj}) \right) = \left\{ \begin{array}{l} \text{Because of the} \\ \text{independence between} \\ \text{surveys in step 2} \end{array} \right\} = E_1\left( \sum_h \sum_{j \in S_h} \tilde{w}_{hj}^2 V_{hj} \right) =$$

$$\sum_h w_h^2 \frac{N_h}{n_h} E_1 \left\{ \frac{\frac{N_h}{n_h} \sum_{j \in S_h} w_{hj}^2 V_{hj}}{\left( \frac{N_h}{n_h} \sum_{j \in S_h} w_{hj} \right)^2} \right\} \approx \left\{ \begin{array}{l} \text{First order} \\ \text{Taylor} \\ \text{approximation} \end{array} \right\} \approx \sum_h w_h^2 \frac{N_h}{n_h} \frac{\sum_{j \in U_h} w_{hj}^2 V_{hj}}{w_h^2} = \sum_h \frac{N_h}{n_h} \sum_{j \in U_h} w_{hj}^2 V_{hj}$$

## Appendix 2

We prove the bias properties of the estimators of the error components in (16) - (19).

We will use the following additional symbol: $\hat{I}^{B}_{h.} = \dfrac{\sum_{j\in S_h} w_{hj} I^{B}_{hj}}{\sum_{j\in S_h} w_{hj}}$

1) $\hat{V}_I$: We look for $E(\hat{V}_I) = E_I E_2(\hat{V}_I)$. In step 2 the only stochastic element is the squared bracket and we therefore start by looking for

$$E_2\left(\hat{I}_{hj} - \hat{\hat{I}}_{h.}\right)^2 = E_2\left\{\left(I^B_{hj} - \hat{I}^B_{h.}\right) + \left(\hat{I}_{hj} - I^B_{hj}\right) - \left(\hat{\hat{I}}_{h.} - \hat{I}^B_{h.}\right)\right\}^2 = \left\{\begin{matrix}\text{We have} \\ E_2(\hat{I}_{hj}) = I^B_{hj} \\ \text{and} \\ E_2(\hat{\hat{I}}_{h.}) = \hat{I}^B_{h.}\end{matrix}\right\} =$$

$$\left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + E_2\left(\hat{I}_{hj} - I^B_{hj}\right)^2 + E_2\left(\hat{\hat{I}}_{h.} - \hat{I}^B_{h.}\right)^2 - 2E_2\left\{\left(\hat{I}_{hj} - I^B_{hj}\right)\left(\hat{\hat{I}}_{h.} - \hat{I}^B_{h.}\right)\right\} =$$

$$\left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + V_{hj} + E_2\left\{\sum_{k\in S_h} \frac{w_{hk}}{w^S_h}\left(\hat{I}_{hk} - I^B_{hk}\right)\right\}^2 - 2E_2\left\{\left(\hat{I}_{hj} - I^B_{hj}\right)\sum_{k\in S_h}\frac{w_{hk}}{w^S_h}\left(\hat{I}_{hk} - I^B_{hk}\right)\right\} =$$

$$\left\{\begin{matrix}\text{Because of} \\ \text{independence} \\ \text{between} \\ \text{surveys } k \in h\end{matrix}\right\} = \left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + V_{hj} + \frac{\sum_{k\in S_h} w^2_{hk} V_{hk}}{\left(w^S_h\right)^2} - \frac{2 w_{hj} V_{hj}}{w^S_h}$$

Next we take the step 1 expectation of

$$E_2\left(\hat{V}_I\right) = \sum_{h\in H} \frac{w_h(w_h - w^S_h)}{\left(w^S_h\right)^2}\left[\sum_{j\in S_h} w^2_{hj}\left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + \sum_{j\in S_h} w^2_{hj} V_{hj} + \frac{\sum_{j\in S_h} w^2_{hj} \sum_{k\in S_h} w^2_{hk} V_{hk}}{\left(w^S_h\right)^2} - \frac{2\sum_{j\in S_h} w^3_{hj} V_{hj}}{w^S_h}\right]$$

By a first order Taylor approximation we obtain

$$E_I\left(E_2\left(\hat{V}_I\right)\right) \approx \sum_{h\in H} \frac{w_h\left(w_h - \frac{n_h}{N_h} w_h\right)}{\left(\frac{n_h}{N_h} w_h\right)^2}\left[\frac{n_h}{N_h}\sum_{j\in h} w^2_{hj}\left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + \frac{n_h}{N_h}\sum_{j\in h} w^2_{hj} V_{hj} + \frac{\frac{n_h}{N_h}\sum_{j\in h} w^2_{hj}\frac{n_h}{N_h}\sum_{k\in h} w^2_{hk} V_{hk}}{\left(\frac{n_h}{N_h} w_h\right)^2} - \frac{2\frac{n_h}{N_h}\sum_{j\in h} w^3_{hj} V_{hj}}{\frac{n_h}{N_h} w_h}\right] =$$

$$\sum_{h\in H} \frac{N_h - n_h}{n_h}\left[\sum_{j\in h} w^2_{hj}\left(I^B_{hj} - \hat{I}^B_{h.}\right)^2 + \sum_{j\in h} w^2_{hj} V_{hj} + \frac{N_h}{n_h}\frac{\sum_{j\in h} w^2_{hj}\sum_{k\in h} w^2_{hk} V_{hk}}{w^2_h} - \frac{N_h}{n_h}\frac{2\sum_{j\in h} w^3_{hj} V_{hj}}{w_h}\right] =$$

$$\sum_{h\in H}\frac{N_h-n_h}{n_h}\sum_{j\in h}w_{hj}^2\left(I_{hj}^B-\hat{I}_{h.}^B\right)^2+\sum_{h\in H}\frac{N_h-n_h}{n_h}\sum_{j\in h}w_{hj}^2V_{hj}\left[1+\frac{N_h}{n_h}\frac{\sum_{j\in h}w_{hj}^2}{w_h^2}-\frac{N_h}{n_h}\frac{2w_{hj}}{w_h}\right]=V_1'+R_V$$

where $V'_1$ differs from $V_1$ only by stratumwise factors $(N_h-1)/N_h$. The discrepancy term $R_V$ depends on the step 2 variances in the superstrata where quasi-sampling was done.

$$2)\ \hat{V}_2:\ E_1E_2\left(\hat{V}_2\right)=\sum_{h\in U}E_1\left\{w_h^2\frac{\sum_{j\in S_h}w_{hj}^2V_{hj}}{\left(\sum_{j\in S_h}w_{hj}\right)^2}\right\}=\sum_{h\in U}w_h^2\frac{N_h}{n_h}E_1\left\{\frac{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}^2V_{hj}}{\left(\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}\right)^2}\right\}\approx$$

$$\left\{\begin{array}{l}\text{First order}\\ \text{Taylor}\\ \text{approximation}\end{array}\right\}\approx\sum_{h\in U}\frac{N_h}{n_h}\sum_{j\in h}w_{hj}^2V_{hj}$$

$$3)\ \hat{B}_1:\ E_1E_2(\hat{B}_1)=E_1\left\{\sum_{h\in U}\sum_{j\in S_h}\left(\tilde{w}_{hj}-\tilde{w}_{hj}^*\right)I_{hj}^B\right\}=E_1\left\{\sum_{h\in U}\sum_{j\in S_h}\left(\frac{w_hw_{hj}}{\sum_{j\in S_h}w_{hj}}-\frac{w_h^*w_{hj}^*}{\sum_{j\in S_h}w_{hj}^*}\right)I_{hj}^B\right\}=$$

$$\sum_{h\in U}\left\{w_hE_1\left(\frac{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}I_{hj}^B}{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}}\right)-w_h^*E_1\left(\frac{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}^*I_{hj}^B}{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}^*}\right)\right\}\approx\sum_{h\in U}\left\{\sum_{j\in h}w_{hj}I_{hj}^B-\sum_{j\in h}w_{hj}^*I_{hj}^B\right\}=\sum_{h\in U}\sum_{j\in h}\left(w_{hj}-w_{hj}^*\right)I_{hj}^B$$

$$=B_1+\sum_{h\in U}\sum_{j\in h}\left(w_{hj}-w_{hj}^*\right)b_{hj}$$

$$4)\ \hat{B}_2:\ E_1E_2\left(\hat{B}_2\right)=E_1\left(\sum_{h\in U}\sum_{j\in S_h}\tilde{w}_{hj}b_{hj}\right)=\sum_{h\in U}E_1\left\{w_h\frac{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}b_{hj}}{\frac{N_h}{n_h}\sum_{j\in S_h}w_{hj}}\right\}\approx\sum_{h\in U}\sum_{j\in h}w_{hj}b_{hj}=B_2$$

**R & D Reports** är en för U/ADB och U/STM gemensam publikationsserie, som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna.

**R & D Reports Statistics Sweden** are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers.

Reports published 1993:

| | |
|---|---|
| 1993:1 (grön) | Kvalitetsfonden 1991/92 - Projekt, handläggning och utvärdering (Roland Friberg) |
| 1993:2 (grön) | Översyn av Undersökningen av lastbilstransporter i Sverige (UVAV) (Bengt Rosén, Mahnaz Zamani) |
| 1993:3 (grön) | Bortfallsbarometern nr 8 (Mats Bergdahl, Pär Brundell, Jan Hörngren, Håkan Lindén, Peter Lundquist, Monica Rennermalm) |
| 1993:4 (gul) | Guidelines on the Design and Implementation of Statistical Meta-information Systems (Bo Sundgren) |
| 1993:5 (grön) | Kvalitetsrapporten 1995 - Utveckling av kvaliteten för SCBs statistikproduktion (Jan Eklöf, Per Nilsson) |
| 1993:6 (grön) | Metoder att få konsistens mellan olika skattningar och samtidigt öka precisionen (Sixten Lundström) |
| 1933:7 (grön) | Variance Estimation in the Swedish Consumer Price Index (Jörgen Dalén, Esbjörn Ohlsson) |

Tidigare utgivna **R&D Reports** kan beställas genom Ingvar Andersson, SCB, U/LEDN, 115 81 STOCKHOLM (telefon 08-783 41 47, telefax 08-783 45 99).

**R&D Reports** listed above as well as issues from 1988-92 can - in case they are still in stock - be ordered from Statistics Sweden, att. Ingvar Andersson U/LEDN, S-115 81 STOCKHOLM, SWEDEN (telephone nr 46 08 783 41 47, telefax nr +46 08 783 45 99).