Variance Estimation in the Swedish Consumer Price Index

Jörgen Dalén och Esbjörn Ohlsson

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1993:7

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 1993:7. Variance estimation in the Swedish consumer price index / Jorgen Dalén; Esbjörn Ohlsson.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1993-X101OP9307

Variance Estimation in the Swedish Consumer Price Index

Jörgen Dalén och Esbjörn Ohlsson

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1993:7

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,
	Department of Research and Development
	S-115 83 Stockholm
Publisher	Lars Lyberg
Inquiries	Jörgen Dalén, telephone +46 08 783 44 94

ISSN 0283-8680

VARIANCE ESTIMATION IN THE SWEDISH

CONSUMER PRICE INDEX

by

Jörgen Dalén, Department of Economic Statistics, Statistics Sweden
Esbjörn Ohlsson, Mathematical Statistics, Stockholm University

Abstract: In large parts of the Swedish Consumer Price Index, products and outlets are sampled independently, yielding a two-dimensional, cross-classified sample. In this paper, a variance formula for this index is derived by exploiting the general theory for cross-classified sampling. Numerical variance estimates are given, along with a description of some of the detailed procedures involved in their computation. Finally, the implications of these estimates on the actual allocation is discussed.

Key words: Cross-classified sampling, two-dimensional sampling, sampling error.

Table of Contents

1. Introduction

This report is concerned with estimating the sampling error of the Swedish Consumer Price Index (in Swedish "Konsumentprisindex", KPI for short). Large parts of the KPI are based on price quotations from two-dimensional samples, each of which is the cross-classification of a sample of *outlets* (shops, restaurants, etc.) and a sample of *products* (items, commodities). In the sequel, such a procedure for sampling from a two-dimensional population will be called *Cross-Classified Sampling*. Ohlsson (1992) gave general results on the variance of an estimator based on a cross-classified sample. In this report we will use these general results to derive estimators of the sampling variance of the KPI. Numerical results for 1991 and 1992 will also be given.

The problem of estimating the sampling variance of a CPI has received quite some interest during the last years. An early reference is Banerjee (1956), who addresses the problem of optimal allocation of the item sample. The first one to give numerical variance estimates for a CPI is, as far as we know, Wilkerson (1967). Recent papers include Leaver & Valliant (1993) on the United States CPI; the two-dimensional sample in the US is, however, a two-stage sample and not a CCS. Balk & Kersten (1986) and Biggeri & Giommi (1987) use balanced half-samples and similar methods to estimate the variance due to the use of weights from a household expenditure survey. Other important papers are Leaver et al (1991) and Leaver & Swanson (1992), which give numerical estimates of the variance for the US CPI for the years 1978-1986 and 1987-1991, respectively, conditional on the expenditure weights.

In Sweden Ruist (1953) and Malmquist (1958), both in Swedish, are early attempts to estimate the KPI sampling error. The (one-dimensional) variance due to the sampling of outlets in the KPI, conditioning on the item sample and making the simplifying assumption of with replacement sampling was computed by Andersson, Forsman & Wretman (1987).

If simple random sampling without replacement (srswor) is used in both dimensions, a variance formula can be obtained from Vos (1964). The connection between the KPI samples and Vos' theory was first noted by Dalén (1992). Since the outlet and product samples are typically drawn with probability proportional to size (pps), Vos' results cannot be applied to the KPI. The generalization to the pps case is given in Ohlsson (1992).

This paper starts by giving an overview of the KPI sampling system in Section 2 . In Section 3 the index definition and estimator are given and in Section 4 the theoretical variance formulas are presented. Section 5 gives the variance estimators and Section 6 discusses the actual procedures in applying these estimators to the actual KPI surveys. Section 7 presents and discusses numerical results for 1991-1992 and Section 8 deals with the problem of optimal allocation.

To the best of our knowledge, the present paper is the first to give numerical variance estimates based on a design-based variance formula for a CPI.

2. The KPI sampling structure

A distinctive feature of the Swedish KPI is that it is composed of many (50-60) independent price surveys for different product groups. Some of these are large in terms of the total weight covered but several are very small.

The all item KPI could be written as a sum of weighted one-survey indices:

$$I = \sum_k w_k I_k \ , \qquad (2.1)$$

where I_k is the one-survey index and w_k the aggregate weight of all items in that survey. Based on sample data the I_k are estimated independently, by $\hat{I}_k$ say, giving rise to the estimated KPI, $\hat{I}$. Hence, the aggregate variance can be computed simply as

$$V(\hat{I}) = \sum_k w_k^2 \, V(\hat{I}_k) \tag{2.2}$$

Hence, an overall variance estimator can be derived by estimating the $V(\hat{I}_k)$ for each survey separately.

The real problem is how to estimate variances for each individual survey. In the search for such estimates we have given priority to large surveys. For surveys covering about 45% of the weight, thorough variance estimates have been done. For surveys covering an additional 35% of the weight we have made less thorough estimates of the order of size of the variances. Based on general knowledge of the price data, the many independent, small surveys making up for the remaining 20% of the weight are not believed to influence a measure of total KPI sampling error significantly. Thorough variance estimates for these surveys are yet to be done.

There are two major sampling dimensions in a CPI for measuring price change - of outlets and of products. In many countries purposive sampling is used in one or both of these dimensions. As far as we know, the only country attempting an all out probability design for its CPI is the United States. In Sweden, outlet sampling is mainly done by probability. Product sampling is done by probability for product groups covering about 18% of the KPI weight while as purposive selection is used for about 40%. For other product groups there is either no product sampling at all (all products are covered) or the distinction between outlets and products is not clear. In the latter cases various mixtures of probability-based and purposive procedures are used.

There is also a third sampling dimension in a CPI - for the estimation of weights. In the KPI weights are mainly taken from the National Accounts that in its turn use retail trade surveys, household budget surveys and other sources. The weight dimension is not further discussed in this report. See Dalén (1993) for a discussion of this and other KPI errors.

The approach in this paper is basically design-based. From this point of view there is no such thing as a variance from a purposive sample. But there is an obvious need to compute some measure of the contribution to the KPI error, due to purposive sampling, at the very least for the purpose of obtaining a rational allocation with respect to the size of the outlet sample vs. the size of the product sample. Our approach on this issue is to calculate variances from purposive samples as if they were drawn by probability, with a design similar to the actual probability samples in some other parts of the KPI. This requires setting up a "dummy design" for the purposive selection reflecting as closely as possible how this selection was actually done. This approach corresponds to that of *quasi-randomization* when modelling the non-response distribution in surveys, see e.g. Särndal, Swensson and Wretman (1992, page 574). Below (in Section 6.2) we describe the construction of a postulated design in the largest KPI price survey.

With this approach we get a unified framework for estimating the variance of several important KPI surveys. An alternative would of course be to change to a model-based approach for the entire KPI. This alternative has not been investigated.

3. Index definition and estimator

The general theory of variance calculation in cross-classified sampling (CCS) is given in Ohlsson (1992). In CCS there are two independent sampling dimensions, rows and columns. In each dimension the samples may be stratified and/or drawn with unequal inclusion probabilities.

The two-dimensional sample S is the cross-classification of the row sample (S^R) with the column sample (S^C), i.e. $S = \{(i,j): i \in S^R \text{ and } j \in S^C\}$.

In this section we shall prepare for the application of this theory to the KPI environment. We consider a single, arbitrary KPI survey. The population is two-dimensional with products as rows and outlets as columns, and a CCS design is used. We start by defining the target parameter, called I, and its estimator, called $\hat{I}$. (For simplicity, we suppress the subindex k, so that I here corresponds to I_k in Section 2.) The I and $\hat{I}$ given here are generalizations of the parameters and estimators of some of the most important KPI surveys. In Section 6 we will go into more details on the actual KPI surveys.

Both the products and the outlets are stratified (by product similarity and type of retail trade, respectively). The crossing of product stratum g with outlet stratum h is called *cell* (g,h). Note that the cells are not used as a strata. Let v_{gh} be a weight for the cell (in the KPI, v_{gh} is the turnover for the products in group g traded in the outlets of type h). The weights are normalized so that

$$\sum_{g=1}^{G}\sum_{h=1}^{H} v_{gh} = 1 \tag{3.1}$$

where G and H are the number of row strata and column strata, respectively. The overall index I is a weighted average of cell indices I_{gh},

$$I = \sum_{g=1}^{G}\sum_{h=1}^{H} v_{gh} I_{gh}. \tag{3.2}$$

We next give the structure of the I_{gh}'s. Let f_{ij} be some function of the price of product i in outlet j in one or several points in time and let g_{ij} be another such function; cf. Example 1 below.

Introduce the indicator

$$1_{ij} = \begin{cases} 1 \text{ if } i \text{ is traded in } j \\ 0 \text{ else} \end{cases} \tag{3.3}$$

For each i let w_i^R be a (row marginal) weight of product i and for each j, let w_j^C be a (column marginal) weight of outlet j with $\sum_i w_i^R = \sum_j w_j^C = 1$. Next let

$$y_{ij} = 1_{ij} w_i^R w_j^C f_{ij}, \quad x_{ij} = 1_{ij} w_i^R w_j^C g_{ij}. \tag{3.4}$$

The cell totals of the y's and x's are denoted by Y_{gh} and X_{gh}, i.e.

$$Y_{gh} = \sum_{i \in g}\sum_{j \in h} 1_{ij} w_i^R w_j^C f_{ij}, \quad X_{gh} = \sum_{i \in g}\sum_{j \in h} 1_{ij} w_i^R w_j^C g_{ij}. \tag{3.5}$$

Here $i \in g$ indicates that the summation is restricted to the i's in stratum g and likewise for $j \in h$. Of course, the functions f_{ij} and g_{ij} may also include some weights - the marginal weights w are given explicitly above only to simplify some of the following formulas. We define the cell index I_{gh} as the ratio

$$I_{gh} = \frac{Y_{gh}}{X_{gh}} = \frac{\sum_{i \in g}\sum_{j \in h} 1_{ij} w_i^R w_j^C f_{ij}}{\sum_{i \in g}\sum_{j \in h} 1_{ij} w_i^R w_j^C g_{ij}}. \tag{3.6}$$

This completes the general definition of I. Let us look at some special cases.

Example 1. Suppose the index is a measure of change in the price level from time 0 to time 1. Let p^t_{ij} be the price of i in j at time t; t=0,1. Upon putting $f_{ij} = p^1_{ij}$ and $g_{ij} = p^0_{ij}$, I_{gh} becomes a ratio of (weighted) mean prices. If instead we let $f_{ij} = p^1_{ij} / p^0_{ij}$ and $g_{ij} = 1$, then I_{gh} becomes a (weighted) mean of price ratios. In the KPI we put

$$f_{ij} = \frac{p^1_{ij}}{(p^0_{ij} + p^1_{ij})/2} \quad and \quad g_{ij} = \frac{p^0_{ij}}{(p^0_{ij} + p^1_{ij})/2} \tag{3.7}$$

and the weights w^R_i and w^C_j are (rough) measures of turnover of the product and outlet, respectively. See Dalén (1992) for the reasons for using (3.7).

We conclude that, by proper choice of f_{ij} and g_{ij}, we can make (3.6) cover many common index formulas, with a notable exception for the geometric mean of price ratios. ♦

We now turn to the definition of the estimator $\hat{I}$ of the index defined above. We shall assume that the indicators 1_{ij} and the prices (and hence f_{ij} and g_{ij}) are known only for the sample, while the weights w^R_i and w^C_j are known for the entire population. This corresponds to the actual KPI situation. In product stratum g, the sample S^R_g is assumed to be drawn with probabilities proportional to the w^R_i, i.e., for some predetermined sample size m_g

$$\pi^R_{gi} = m_g w^R_i . \tag{3.8}$$

In outlet stratum h, the sample S^C_h is assumed to be drawn with probabilities

$$\pi^C_{hj} = n_h w^C_j , \tag{3.9}$$

for some sample size n_h. Here we must assume that the quantities on the right-hand side of (3.8) and (3.9) do not exceed 1. In practice this is achieved by, whenever necessary, forming separate complete enumeration strata for large units. Taking care of such strata is a straight-forward task, but it makes the formulas rather involved. For simplicity we will, in the theoretical exposition, assume that no large unit strata are necessary.

The Horvitz-Thompson estimator of Y_{gh} and X_{gh} are, by (3.5), (3.8) and (3.9).

$$\hat{Y}_{gh} = \frac{1}{m_g n_h} \sum_{i \in S^R_g} \sum_{j \in S^C_h} 1_{ij} f_{ij} \quad \text{and} \quad \hat{X}_{gh} = \frac{1}{m_g n_h} \sum_{i \in S^R_g} \sum_{j \in S^C_h} 1_{ij} g_{ij} , \tag{3.10}$$

respectively. As an estimator of the ratio in (3.6) we take

$$\hat{I}_{gh} = \frac{\hat{Y}_{gh}}{\hat{X}_{gh}} = \frac{\sum_{i \in S^R_g} \sum_{j \in S^C_h} 1_{ij} f_{ij}}{\sum_{i \in S^R_g} \sum_{j \in S^C_h} 1_{ij} g_{ij}} . \tag{3.11}$$

Note that $\hat{I}_{gh}$ is "self-weighting"; this property is lost by the introduction of large unit strata, though. Finally the estimated I is, by (3.2),

$$\hat{I} = \sum_{g=1}^{G} \sum_{h=1}^{H} v_{gh} \hat{I}_{gh} . \tag{3.12}$$

4. The variance of the index

We shall apply the general CCS theory to find an expression for the variance V(Î). In doing so we must specify the sampling procedures that are used to generate samples with inclusion probabilities as in (3.8) and (3.9). The actual sampling procedures used in the surveys of interest are: sequential Poisson sampling (of outlets), random systematic PPS sampling (of products in one system) and purposive sampling (of products in one system). For a description of random systematic sampling and ordinary Poisson sampling, see e.g. Brewer & Hanif (1983, p.22); for a description of sequential Poisson sampling see Ohlsson (1990 and 1993). For simplicity, we will postulate random systematic PPS sampling in both dimensions, cf. the discussion on purposive sampling in Section 2.

As usual when estimating a ratio, Î is not unbiased. As argued in Appendix 2, Î is approximately unbiased, though. Hence, we measure the sampling error by V(Î), as an approximation to the mean square error of Î. The simulation studies in Section 7 of Ohlsson (1992) indicate that the bias is indeed small as compared to the variance.

Before presenting the formula for V(Î), we introduce some further notation. For $i \in g$ and $j \in h$, let

$$e_{ij}^{gh} = I_{ij}(f_{ij} - I_{gh}g_{ij}) \;, \tag{4.1}$$

and put

$$\bar{e}_{i.}^{gh} = \sum_{j \in h} w_j^C e_{ij}^{gh} \text{ and } \bar{e}_{.j}^{gh} = \sum_{i \in g} w_i^R e_{ij}^{gh} . \tag{4.2}$$

Proposition 4.1: The variance of the index Î can be approximated as follows

$$V(\hat{I}) \approx \quad V_{PRO} + V_{OUT} + V_{INT} \tag{4.3}$$

where

$$V_{PRO} = \sum_{g=1}^{G} \frac{1}{m_g} \sum_{i \in g} \left(1 - (m_g - 1) w_i^R\right) \left(\sum_{h=1}^{H} \frac{v_{gh}}{X_{gh}} \bar{e}_{i.}^{gh} \right)^2 w_i^R , \tag{4.4}$$

$$V_{OUT} = \sum_{h=1}^{H} \frac{1}{n_h} \sum_{j \in h} \left(1 - (n_h - 1) w_j^C\right) \left(\sum_{g=1}^{G} \frac{v_{gh}}{X_{gh}} \bar{e}_{.j}^{gh} \right)^2 w_j^C \tag{4.5}$$

and

$$V_{INT} = \sum_{g=1}^{G} \sum_{h=1}^{H} \frac{1}{m_g n_h} \frac{v_{gh}^2}{X_{gh}^2} \Bigg[\sum_{i \in g} \sum_{j \in h} \left(1 - (m_g - 1) w_i^R\right)\left(1 - (n_h - 1) w_j^C\right)\left(e_{ij}^{gh}\right)^2 w_i^R w_j^C \\ - \sum_{i \in g} \left(1 - (m_g - 1) w_i^R\right)\left(\bar{e}_{i.}^{gh}\right)^2 w_i^R - \sum_{j \in h} \left(1 - (n_h - 1) w_j^C\right)\left(\bar{e}_{.j}^{gh}\right)^2 w_j^C \Bigg] \tag{4.6}$$

The derivation of this proposition from the general results in Ohlsson (1992) is postponed to Appendix 2.

The approximation in (4.3) is partly due to the use of a standard type linearization of the ratio estimator, (see Appendix 2, formula A.1). and partly due to the use of approximations for the second order inclusion probabilities of random systematic sampling, (see Appendix 2, formula A.8). V_{PRO} will be called the "variance between products", V_{OUT} the "variance between outlets", while V_{INT} is "outlet and product interaction". A reason for this terminology is that V_{PRO} is the variance that would result if outlets were completely enumerated, and similarly for V_{OUT}. (This fact can not be seen in the above formulas, because of their approximative nature, but it is easily seen in the exact variance formulas in Ohlsson, 1992).

The V_{INT} component does not have the interaction structure one might expect from Ohlsson (1992, formula 1.9) and the fact that srswor is a particular case of random systematic sampling. Presumably this is caused by the approximation of the second order inclusion probabilities (A.8). Suppose, however, that the finite population corrections $1-(m_g-1)w_i^R$ and $1-(n_h-1)w_j^C$ are all close to 1, and hence can be omitted. Then it is readily seen that V_{INT} can be rewritten as

$$V_{INT} = \sum_{g=1}^{G}\sum_{h=1}^{H}\frac{1}{m_g n_h}\frac{v_{gh}^2}{X_{gh}^2}\sum_{i\in g}\sum_{j\in h}\left(e_{ij}^{gh} - \bar{e}_{i.}^{gh} - \bar{e}_{.j}^{gh} + \bar{e}_{..}^{gh}\right)^2 w_i^R w_j^C \tag{4.7}$$

Here $\quad \bar{e}_{..}^{gh} = \sum_{i\in g}\sum_{j\in h} w_i^R w_j^C e_{ij}^{gh} = 0 \qquad (4.8)$

$\bar{e}_{..}^{gh}$ is inserted into (4.7) only to reveal the similarities with the interaction term in a two-way analysis of variance table; cf. also Ohlsson (1992, formula 1.9).

5. Variance estimators

In deriving the variance estimators presented below we will basically use Proposition 4.1 which is valid for random systematic sampling. Piecewise Horvitz-Thompson estimation is used for unknown population quantities. The finite population corrections (fpc) are, however, adjusted on heuristic grounds so that they become zero for sampling units included with probability one. This can also be viewed as a "calibration to srswor", meaning that the fpc's have been adjusted so that the variance formulas are exact in case all selection probabilities are equal. In such a case, random systematic sampling reduces to srswor; hence we would like our variance formulas to be identical to the ones for srswor. The fact that this is not the case before adjustment can be attributed to the approximation of the second-order inclusion probabilities in (A.8) which is used to derive Proposition 4.1. The adjustment is important in cases of inclusion probabilities equal (or close) to one where we want the fpc's to be equal (or close) to zero.

Similarly, the variance estimators were inflated by factors $m_g/(m_g-1)$ and $n_h/(n_h-1)$ to "calibrate" to the unbiased variance estimator in srswor.

When trying to estimate (4.6) we ended up with negative variance estimates in certain cells. The reason is presumably that the approximation procedure in (A.8) does not work well in this situation. We therefore turned to a quadratic structure of the V_{INT} component as in (4.7). Fortunately, as we shall see, the contribution of V_{INT} to $V(\hat{I})$ is much smaller than that of V_{PRO} and V_{OUT}.

The following formulas were thus used for variance estimation:

$$\hat{V}_{PRO} = \sum_{g}\frac{1}{m_g(m_g-1)}\sum_{i\in S_g^R}(1-\pi_{gi}^P)\left\{\sum_{h}\frac{v_{gh}}{\hat{X}_{gh}}\hat{e}_{i.}^{gh}\right\}^2, \tag{5.1}$$

$$\hat{V}_{OUT} = \sum_{h}\frac{1}{n_h(n_h-1)}\sum_{j\in S_h^C}(1-\pi_{hj}^C)\left\{\sum_{g}\frac{v_{gh}}{\hat{X}_{gh}}\hat{e}_{.j}^{gh}\right\}^2 , \tag{5.2}$$

$$\hat{V}_{INT} = \sum_g \sum_h \frac{v_{gh}^2}{\hat{X}_{gh}^2} \frac{1}{n_h(n_h-1)} \frac{1}{m_g(m_g-1)} \left\{ \sum_{i \in S_g^R} \sum_{j \in S_h^C} (1-\pi_{hj}^C)(1-\pi_{gi}^P)(\hat{e}_{ij}^{gh} - \hat{\bar{e}}_{i.}^{gh} - \hat{\bar{e}}_{.j}^{gh})^2 \right\} \quad (5.3)$$

$\hat{e}_{ij}^{gh}$ is as in (4.1) but with $\hat{I}_{gh}$ replacing I_{gh}.

$$\hat{\bar{e}}_{i.}^{gh} = \frac{1}{m_g} \sum_{i \in S_g^R} \hat{e}_{ij}^{gh} \text{ and } \hat{\bar{e}}_{.j}^{gh} = \frac{1}{n_h} \sum_{j \in S_h^C} \hat{e}_{ij}^{gh} \quad (5.4)$$

6. Variance estimation in practice

The variance estimation procedures above were applied to two of the major price measurement systems in the KPI. These are the List Price System (LIPS) and the Local Price System (LOPS).

6.1 The List Price System (LIPS)

In LIPS probability sampling is used for products as well as outlets. Prices are taken from wholesalers' price lists for each month except December when interviewers collect actual prices in the shops. (From 1993 price lists are no longer be used at all.) LIPS accounts for about 18% of the total CPI weight.

LIPS covers most food products (not fresh food such as fruits and vegetables, bread and pastries or fish) and other daily commodities such as those typically found in a supermarket. Sampling of products is based on historic sales data from the three major wholesalers in Sweden using systematic PPS sampling of products, stratified into about 60 product strata.

Sampling of outlets is done by pps sampling, viz. sequential Poisson sampling (Ohlsson, 1990 and 1993). The size measure is number of employees plus 1, used as a rough measure of turnover. The outlets are first stratified by region and (to a certain extent) wholesaler. The product sample is matched to the sampled outlets so that the outlet is given the product sample of the wholesaler that delivers its goods.

There were about 1200 products in the sample in 1991-1992. The number of outlets was 58 in 1991 and 59 in 1992.

Our postulated design (which is a slight simplification of the actual one) uses an asymmetric stratification structure. We start by stratifying the whole product-outlet population into three wholesaler strata. In each of these **primary strata** cross-classified sampling is done independently. In each primary stratum an unstratified PPS sample of outlets is cross-classified with a stratified PPS sample of products. This gives the following variance structure

$$V(\hat{I}) = \Sigma_h w_h^2 V_h(\hat{I}) \quad (6.1)$$

where the w_h are market share weights of the three wholesalers and the $V_h(\hat{I})$ are variances in primary strata which in turn have three components computed as in (5.1)-(5.3). Note that there is now only one outlet stratum within each primary stratum.

6.2 The Local Price System (LOPS)

In LOPS interviewers collect prices each month. The system accounts for about 21% of the total CPI weight.

Outlets are divided into 25 strata according to the SNI code (Swedish Code of Industrial Classification which closely follows the International Standard of Industrial Classication - ISIC) of the outlet. Within a stratum sequential Poisson sampling is used as in LIPS.

For products, however, purposive sampling is used in several steps. The products are divided into product groups according to the National Accounts and other sources of information . Within a product group one or more products are chosen, often only one. All in all there are some 140 "representative products" in LOPS. For each of these 140 products a commodity specification is done at the central office. The interviewer is then asked to find the particular variety according to the specification that is the most sold within the surveyed outlet.

The actual sampling procedure in LOPS is thus CCS with probability sampling of outlets, but purposive sampling of products. Our postulated design is, however, probability sampling of products, too.

There are 20 (1991) or 21 (1992) outlet strata (some strata are collapsed) and 48 (1991) or 43 (1992) product strata.

The forming of the product strata was based on the actual procedures used in the selection of the representative products. There the starting point is often information on the consumption value of a rather narrow product group like "bananas", "dish washers", or "towels". The final sample is then one or several representative products in each group. In our variance estimation procedures we consider these final products as randomly chosen from their respective product groups. For imputing a subjective "inclusion probability" into the variance formulas we basically ask the question. How much of the consumption value in this product group is accounted for by the product finally selected by the interviewer? For example for the product group *bananas* there is typically only one brand and one price in an outlet and we therefore set the inclusion probability to one for the representative product *one kg of bananas* within the product group *bananas*. On the contrary, a particular brand of the representative product *towel, terry cloth, 100% cotton, hemmed, about 50x70 cm* within the product group *towels* is likely to have a rather small share of an outlet's total sales value of towels and in this and many similar cases we set the "inclusion probability" to 0.1.

We believe that, although there are subjective elements in our procedure, it gives a fairly realistic picture of the random error arising from product sampling. The greatest problem is probably our procedures for collapsing product strata in those cases where only one product was selected in a product group. This may lead to some overestimation since products in the actual products strata could be expected to have more similar price movements than those in the collapsed strata.

6.3 Other price surveys

Variance computations were done for other price surveys too. The methods used were cruder and more simplified. The exact procedures are not of general interest so a short summary will be sufficient.

The **apparel survey** (covering about 6% of the total KPI weight) uses sequential sampling of outlets and a purposive sample of 24 garments. But in each outlet several (up to 8) varieties of each garment were priced. According to various comparability criteria only a certain portion of the varieties are included in the index. Here we used a simple one-dimensional variance estimation procedure based on the effective sample size for each garment.

The **rental survey** (directly 10% and with imputation about 13.5% of the weight) is based on a random sample of about 1000 apartments and is thus one-dimensional. The estimator is post-stratified into different size groups crossed with newly built versus old ones. The estimated index is a kind of unit-value index comparing the average rent per m^2 for all apartments at two points of time, with corrections for differences in quality (equipment etc.) Here, our variance computations are simplified in so far that they do not take account of the changing population.

The **interest survey** (8.5% of the weight) estimates the amount of interest paid or foregone for owner-occupied homes by multiplying their total capital investment (defined as total purchase values) with the average interest rate. For estimating capital investment a stratified sample of some 650 homes was drawn in 1973. To this sample annual supplementary samples of newly built homes are added which gives a total sample of about 800 homes in 1992. Conditional variance calculations reflecting only the sampling for and estimation of capital investment have been done. This variance component is judged to dominate the sampling error of the interest survey.

The **car survey** (3%) was, in 1991, based on 31 brands of cars (60 in 1992). The design could be interpreted as stratified with one purposively selected unit in each stratum. Variances have been computed based on a crude assumption of simple random sampling. There is some evidence that this is a fairly reasonable assumption in this case.

The **petrol survey** (4%) covers five types of petrol (almost all existing) at 120 petrol stations, sampled with sequential Poisson sampling so variance estimation was done in a straight-forward way.

The **survey of alcoholic beverages** (2.6%) relies on information from the Swedish state monopoly for its total sales. It thus has zero sampling error (leaving aside taxfree sales at airports etc.).

Other product groups also have zero or very small sampling error. This is the case for, e.g., **gambling and lotteries, TV license fees, hospital care** and **dental care** (2.9% together).

For other surveys weights are so small that they could not influence a measure of total KPI sampling error significantly. Direct variance estimates have not yet been done, however.

7. Numerical results for 1991 and 1992

Up to now we have computed sampling errors for two years for the major price surveys. In Table 1-3 in Appendix 1 we give December-to-month variance estimates (within an annual link) for those surveys obtained with formulas (5.1)-(5.3) for the two largest surveys. In table 4-5 in Appendix 1 some other variance estimates are given.

The interpretation of these results is the following with LIPS as our example. V_{TOT} for June 1991 is 0.0904. If the LIPS part of KPI for June 1991 with December 1990 as reference period were 104 an appropriate 95% confidence interval for its sampling error would be $104 \pm 1.96 \times \sqrt{0.0904} = 104 \pm 0.6$. If LIPS had been the only source of sampling error in the KPI and the KPI figure for June 1991 was 105 then a 95% confidence interval for it would be $105 \pm 1.96 \times \sqrt{0.00283} = 105 \pm 0.1$. The contribution to the KPI variance is calculated according to (2.2) remembering that the weight for the whole LIPS system was 0.177 in 1991.

In LIPS only minor changes in the samples were done between 1991 and 1992. Still there were quite significant changes in the estimated variances. In early 1991, as well as for earlier years for which we have cruder estimates (not presented here), the V_{PRO} and V_{OUT} components are about the same size although, as mentioned in 3.1 above, there are 20 times more products than outlets. This means that the underlying between-product variation was much larger than the underlying between-outlet variation and was the basic reason for the seemingly extreme allocation of the sample. We also note that the V_{INT} component is of a much smaller order of size than V_{PRO} and V_{OUT}. In 1992 and particularly its last months, however, this relation changed with V_{PRO} decreasing and V_{OUT} increasing. A possible explanation of this is that some outlets have changed their pricing behaviour, moving from quickly changing campaign prices to more stable "always-low-prices". It remains to be seen if this tendency will remain, in which case the allocation should be adjusted.

For LOPS (table 2) the V_{PRO} component is always larger than V_{OUT}. We have here some 800 outlets and 140 products. Since it costs less to include one more product than one more outlet in the sample, this means that we have had a poor allocation. These results have generated a successive movement towards a more efficient allocation with more products and fewer outlets in this survey which is the most expensive in the KPI, cf. Section 8.

The Apparel survey (table 3) shows increasing variances from 1991 to 1992 and is now the greatest concern in our KPI sampling design. Unfortunately, we have not so far been able to produce a good decomposition of this variance. A major problem here is that our criteria for comparability are sometimes so restrictive that the result is a very small effective sample. But more liberal comparability criteria lead to a risk for increasing biases instead. For 1993, measures have been taken to increase the efficiency of the comparability criteria.

For all these three surveys there is a tendency for the variances to increase over the year. This is of course to be expected since we measure changes from December and changes are generally small in the beginning of the year. In LOPS and the Apparel Survey, but not in LIPS, variances are also high during the summer. A feature which all three surveys have in common is the high variability of the variance estimates which complicates allocation decisions.

The Interest survey gave the conditional variance estimates for 1990-1992 presented in table 4.

For other price surveys than those discussed in some detail above, some crude estimates are given in table 5 only to show that their contributions to KPI variance is of a smaller order of size than for those surveys where we have used more careful procedures.

8. The allocation problem

In order to give a survey an optimal sampling allocation two things are neccessary: a variance function and a cost function. Above we have concentrated on the variance function. Here the cost function is first discussed .

In the Swedish KPI practice there are two levels of the allocation problem - between surveys and within surveys. So far we have not been able to give direct cost estimates to different surveys. As proxies we could use sample sizes which gives the following picture, with reference to 1991, for the surveys discussed above.

TABLE 6

Survey	Variance estimate, on average	Contribution to KPI variance	Effective sample size	Mode of data collection
List Price System	0.09	0.003	18000	Price Lists (Visits)
Local Price System	0.19	0.009	8500	Visits, Some Telephone
Apparel Survey	2.06	0.009	1000	Visits
Interest Survey	1.97	0.013	800	Mail
Rental Survey	0.1	0.002	1000	Mail
Car Survey	1	0.0015	31	Telephone
Petrol Survey	0.001	0.000002	600	Telephone

Now, of course, costs are not generally proportional to sample size. In LIPS, for example, a mode of data collection (mainly price lists) was used which makes it much less expensive than LOPS (where interviewers to a large extent visit the outlets for price measurement) despite its larger sample size.

Some obvious misallocation is readily seen from table 5, however. The sample sizes of the car and the petrol survey should rather be reversed, for example. Also there is an obvious need for increasing the sample sizes for homes (for which the large variance was discovered only recently) and for apparel items.

When it comes to allocation within surveys, we take our most expensive survey, LOPS, as our example. Here we use the following cost function:

$$C = C_0 + \sum_h n_h \left\{ a_h + b_h \sum_g m_g \left(\sum_{k \in g} r_{ghk} \right) \right\}, \tag{8.1}$$

where
C_0 is a fixed cost for administration etc.,
n_h is the number of outlets in outlet stratum h,
m_g is the number of items in item stratum g,
a_h is the fixed cost per outlet in stratum h, mainly due to travel time,
b_h is the cost of measuring one item in outlets of stratum h and

r_{ghk} is the relative frequency of item $k \in g$ in outlets of stratum h.

In practice a_h depends on the extent to which telephone interviews could be used for price measurement in stratum h. This could be done in most months, if there are few and simply defined items in the outlet.

The b_h are the same for most strata but food items are generally simpler to measure and so b_h is smaller for the daily commodity stores.

Now if we try to combine (8.1) with (5.1)-(5.3) we run into a non-linear optimization problem for which it seems impossible to find an explicit solution. A direction to go would therefore be to try some kind of numerical optimization procedure. So far we have not made attempts in this direction since the necessary work could be expected to be large.

Although a formal optimization has not been done, the variance and cost functions have made sizeable improvements in the LOPS allocation possible by simple inspection of the numerical results combined with calculation of marginal changes. We have increased the number of products (in particular in the highly variable product group *fresh fruit and vegetables*) in the survey while cutting down on the number of outlets. All together this has led to lower costs and smaller variance at the same time!

9. References

Andersson, C., Forsman, G. & Wretman, J. (1987). On the measurement of errors in the Swedish Consumer Price Index. *Bull. Int. Stat. Inst.* **52**, Book 3, 155-171.

Balk, B.M. & Kersten, H.M.P. (1986). On the precision of Consumer Price Indices caused by the sampling variability of budget surveys. *Journal of Economic & Social Measurement* **14**, 19-35.

Banerjee, K.S. (1956). A note on the optimal allocation of consumption items in the construction of a cost of living index. *Econometrica* **24**, 294-295.

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.

Brewer, K.R.W. & Hanif, M. (1983). *Sampling with unequal probabilities.* Springer, New York.

Connor, W.S. (1966). An exact formula for the probability that two specified sample units will occur in a sample drawn with unequal probabilities and without replacement. *J. Amer Statist. Assoc.***61**, 384-390.

Dalén, J. (1992). Computing elementary aggregates in the Swedish Consumer Price Index, *Journal of Official Statistics* **8**, 129-147.

Dalén, J. (1993). Quantifying errors in the Swedish Consumer Price Index. *R&D Reports*, Statistics Sweden.

Hartley, H.O. & Rao, J.N.K. (1962). Sampling with unequal probabilities and wihout replacement. *Ann. Math. Statist.***33**, 350-374.

Leaver, S., Johnstone, J., & Archer, K. (1991). Estimating unconditional variances for the U.S. Consumer Price Index for 1978-1986. American Statistical Association, *Proceedings of the Section on Survey Research Methods*, 614-619.

Leaver, S. & Swanson, D. (1992). Estimating variances for the U.S. Consumer Price Index for 1987-1991. American Statistical Association, *Proceedings of the Section on Survey Research Methods*, forthcoming.

Leaver, S. & Valliant, R. (1993). Statistical problems in estimating the U.S. Consumer Price Index. To appear in *Survey Methods for Businesses, Farms and Institutions*. Wiley, New York.

Malmquist, S. (1958): Precisionsproblemet vid beräkning av konsumentprisindex, *Statistisk Tidskrift* **7**, 492-504 and 612-628.

Ohlsson, E. (1990): Sequential Poisson sampling from a Business register and its application to the Swedish Consumer Price Index. *R & D Report 1990:6*, Statistics Sweden.

Ohlsson, E. (1992). Cross-classified sampling for the Consumer Price Index. *R & D Report 1992:7*, Statistics Sweden.

Ohlsson, E. (1993). Coordination of several samples using permanent random numbers. To appear in *Survey Methods for Businesses, Farms and Institutions*. Wiley, New York.

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

Särndal, C.E., Swensson, B., and Wretman, J. (1992). *Model assisted Survey Sampling*. Springer Series in Statistics.

Vos, J.W.E. (1964). Sampling in space and time. *International Statistical Review* **32**, 226-241.

Wilkerson, M. (1967). Sampling error in the Consumer Price Index. *Journal of the American Statistical Association* 62, 899-914

Appendix 1: Tables

TABLE 1: Variance estimates in LIPS 1991-1992

Month	$\hat{V}_{PRO}$	$\hat{V}_{OUT}$	$\hat{V}_{INT}$	$\hat{V}(\hat{I})$	Contribution to KPI Variance
1991					
January	0.0135	0.0225	0.0025	0.0385	0.0012
February	0.0313	0.0427	0.0034	0.0774	0.0024
March	0.0311	0.0467	0.0038	0.0816	0.0026
April	0.0408	0.0456	0.0053	0.0917	0.0029
May	0.0514	0.0442	0.0040	0.0996	0.0031
June	0.0450	0.0416	0.0038	0.0904	0.0028
July	0.0448	0.0457	0.0041	0.0946	0.0030
August	0.0458	0.0415	0.0044	0.0917	0.0029
September	0.0480	0.0633	0.0045	0.1158	0.0036
October	0.0432	0.0647	0.0041	0.1120	0.0035
November	0.0390	0.0628	0.0042	0.1060	0.0033
December, stix	0.0373	0.0649	0.0042	0.1061	0.0033
December, ltix	0.0393	0.0685	0.0097	0.1175	0.0037
1992					
January	0.0106	0.0134	0.0016	0.0256	0.0007
February	0.0117	0.0144	0.0018	0.0279	0.0008
March	0.0121	0.0146	0.0019	0.0286	0.0008
April	0.0108	0.0241	0.0020	0.0369	0.0010
May	0.0167	0.0299	0.0022	0.0488	0.0013
June	0.0171	0.0303	0.0026	0.0500	0.0014
July	0.0200	0.0299	0.0026	0.0525	0.0014
August	0.0173	0.0400	0.0027	0.0600	0.0016
September	0.0207	0.0357	0.0027	0.0591	0.0016
October	0.0192	0.0307	0.0028	0.0527	0.0014
November	0.0151	0.0599	0.0030	0.0780	0.0021
December, stix	0.0164	0.0721	0.0031	0.0916	0.0025
December, ltix	0.0177	0.1076	0.0065	0.1318	0.0036

TABLE 2: Variance estimates in LOPS 1991-1992

Month	$\hat{V}_{PRO}$	$\hat{V}_{OUT}$	$\hat{V}_{INT}$	$\hat{V}(\hat{I})$	Contribution to KPI variance
1991					
January	0.0490	0.0322	0.0126	0.0938	0.0044
February	0.0502	0.0353	0.0160	0.1015	0.0048
March	0.0968	0.0388	0.0189	0.1545	0.0072
April	0.1142	0.0443	0.0218	0.1803	0.0085
May	0.1343	0.0454	0.0243	0.2040	0.0096
June	0.2648	0.0510	0.0274	0.3432	0.0161
July	0.1423	0.0581	0.0359	0.2363	0.0111
August	0.0874	0.0623	0.0333	0.1830	0.0086
September	0.1164	0.0636	0.0323	0.2123	0.0100
October	0.0982	0.0571	0.0339	0.1892	0.0089
November	0.1296	0.0577	0.0342	0.2215	0.0104
December	0.1029	0.0630	0.0370	0.2029	0.0095
1992					
January	0.0394	0.0258	0.0125	0.0777	0.0032
February	0.0744	0.0294	0.0172	0.1210	0.0049
March	0.0487	0.0335	0.0178	0.1000	0.0041
April	0.0603	0.0442	0.0215	0.1260	0.0051
May	0.0678	0.0387	0.0244	0.1309	0.0053
June	0.0797	0.0447	0.0276	0.1520	0.0062
July	0.1452	0.0531	0.0317	0.2300	0.0094
August	0.0991	0.0627	0.0334	0.1952	0.0080
September	0.0718	0.0529	0.0338	0.1585	0.0065
October	0.1450	0.0587	0.0484	0.2121	0.0086
November	0.1476	0.0600	0.0533	0.2609	0.0106
December	0.1416	0.0705	0.0561	0.2682	0.0109

TABLE 3 Variance estimates in the Apparel Survey 1991-1992

Month	Total variance		Contribution to KPI total variance	
	1991	*1992*	*1991*	*1992*
January	0.98	1.28	0.0034	0.0041
February	1.47	2.10	0.0051	0.0067
March	1.63	3.38	0.0056	0.0109
April	2.09	1.94	0.0072	0.0088
May	2.68	3.01	0.0093	0.0109
June	4.50	7.82	0.0156	0.0176
July	4.49	11.06	0.0155	0.0356
August	4.01	11.19	0.0139	0.0360
September	2.07	3.33	0.0072	0.0107
October	2.01	2.01	0.0069	0.0065
November	1.49	3.62	0.0051	0.0122
December	1.84	3.68	0.0064	0.0124

TABLE 4 Conditional variance estimates in the interest survey

	1990	*1991*	*1992*
Variance estimate	0.79	1.97	2.52
Contribution to KPI total variance	0.005	0.013	0.016

TABLE 5 Crude variance estimates for some price surveys

Survey	Variance Estimate	KPI Weight 1992	Contribution to KPI Variance
Rental Survey	0.1	0.1348	0.002
Car Survey	1	0.0298	0.0015
Petrol Survey	0.001	0.0402	0.000002

Appendix 2: Derivation of the variance formula

In this appendix we shall derive Proposition 4.1 from Theorem 2.1 and Theorem 3.1 in Ohlsson (1992).

Recalling (3.2), (3.6) and (3.11)-(3.12), we have the following (Taylor series) linear approximation.

$$\hat{I} - I = \sum_g \sum_h v_{gh} \frac{\hat{Y}_{gh} - I_{gh}\hat{X}_{gh}}{\hat{X}_{gh}} \approx \sum_g \sum_h v_{gh} \frac{\hat{Y}_{gh} - I_{gh}\hat{X}_{gh}}{X_{gh}} \qquad \text{(A.1)}$$

From the unbiasedness of the Horvitz-Thompson estimator and (A.1) we conclude that $\hat{I}$ is approximately unbiased. We can now concentrate on deriving the variance of the right-hand side of (A.1). Set

$$z_{ij} = v_{gh} I_{ij} w_i^R w_j^C \left(f_{ij} - I_{gh} g_{ij}\right) / X_{gh}, \qquad \text{(A.2)}$$

and note that by (3.5) and (3.6) the cell total Z_{gh} is

$$Z_{gh} = \sum_{i \in g} \sum_{j \in h} z_{ij} = v_{gh}\left(Y_{gh} - I_{gh} X_{gh}\right) / X_{gh} = 0. \qquad \text{(A.3)}$$

Let $\hat{Z}_{gh}$ be the Horvitz-Thompson estimator of Z_{gh}. By (3.8), (3.9) and (4.1) we have

$$\hat{Z}_{gh} = \frac{v_{gh}}{m_g n_h X_{gh}} \sum_{i \in S_g^R} \sum_{j \in S_h^C} e_{ij}^{gh} \quad , \qquad \text{(A.4)}$$

Let $\hat{Z}$ be the sum of $\hat{Z}_{gh}$ over all cells. From (A.1) and (3.10) we get

$$V(\hat{I}) \approx V\left(\sum_g \sum_h \hat{Z}_{gh}\right) = V(\hat{Z}) \qquad \text{(A.5)}$$

We shall apply the results in Ohlsson (1992) to the right-hand side of (A.5). From Theorem 2.1 there we get a decomposition of $V(\hat{Z})$ into three parts. We shall show, one by one, that these parts equal the three components in Proposition 4.1. We start with V_{PRO}. In accordande with the mentioned theorem we first take the expectation over the outlet sample, conditional on the product sample, in (A.4) and get

$$\frac{v_{gh}}{m_g X_{gh}} \sum_{i \in S_g^R} \sum_{j \in S_h^C} \bar{e}_{i.}^{gh}. \qquad \text{(A.6)}$$

Let $\hat{Z}_{g.}$ be the sum of (A.6) over all column cells h, i.e.

$$\hat{Z}_{g.} = \frac{1}{m_g} \sum_{i \in S_g^R} \sum_h \frac{v_{gh}}{X_{gh}} \bar{e}_{i.}^{gh}. \qquad \text{(A.7)}$$

According to Theorem 2.1 in Ohlsson (1922), the first component in $V(\hat{Z})$ is $V\left(\sum_g \hat{Z}_{g.}\right)$. Next note that $\hat{Z}_{g.}$ is the (one-dimensional) Horvitz-Thompson estimator of the total $Z_{g.} = \sum_h Z_{gh}$ in case the outlets are completely enumerated. Further note that the $\hat{Z}_{g.}$ are independent, since they are based on different strata. Recalling the well-known approximation formula for the variance of the Horvitz-Thompson estimator in random systematic sampling (formula 1.8.4 in Brewer & Hanif, 1983), we see that $V\left(\sum_g \hat{Z}_{g.}\right)$ is equal to the expression for V_{PRO} in Proposition 4.1.

The expression for V_{OUT} in the proposition is obtained from V_{PRO} by noting the symmetry of rows and columns in a CCS design. Our final concern is the interaction term V_{INT}. According to Theorem 2.1 this term (called VRC there) can be found by adding all the within cells interaction terms. We therefore concentrate on a single cell and drop the cell indices g and h in the rest of the proof.

We shall apply Theorem 3.1 in Ohlsson (1992), which gives V_{INT} in terms of inclusion probabilities. Connor (1966) supplied exact expressions for the second order inclusion probabilities in pps random systematic sampling. These expressions are unmanageable in practice, though. Instead, we shall use an approximation due to Hartley & Rao (1962), see also Brewer & Hanif (1983, p.14). Omitting all th terms except the first one in this approximation we get

$$\pi_{ii'}^R = \frac{m-1}{m}\pi_i^R\pi_{i'}^R,\ i \neq i' \text{ and } \pi_{jj'}^C = \frac{n-1}{n}\pi_j^C\pi_{j'}^C,\ j \neq j' . \qquad \text{(A.8)}$$

By inserting (A.8) into the expression for V_{INT} (VRC) given in Theorem 3.1, and using the fact that Z=0, we find

$$\begin{aligned} V_{INT} = &\sum_i\sum_j \frac{1}{\pi_i^R\pi_j^C}\left(1-\frac{m-1}{m}\pi_i^R\right)\left(1-\frac{n-1}{n}\pi_j^C\right)z_{ij}^2 \\ &-\sum_i \frac{1}{n\pi_i^R}\left(1-\frac{m-1}{m}\pi_i^R\right)z_{i.}^2-\sum_j \frac{1}{m\pi_j^C}\left(1-\frac{n-1}{n}\pi_j^C\right)z_{.j}^2 . \end{aligned} \qquad \text{(A.9)}$$

By inserting (3.8), (3.9) and (A.2) into (A.9) we get (4.6) which completes the proof of Proposition 4.1.

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)

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).