The Swedish Consumer Price Index

A handbook of methods

Throughout this Handbook, the Swedish CPI is referred to as the KPI (Konsumentprisindex).
SCB stands for Statistiska Centralbyrån (Statistics Sweden).

Statistiska centralbyrån
Statistics Sweden

The Swedish Consumer Price Index
A handbook of methods

Producent SCB, Ekonomisk statistik,
Priser
Producer Box 24300, 104 51 STOCKHOLM

Förfrågningar priser@scb.se
Inquiries

ISBN 91-618-1097-5

Printed in Sweden
SCB-Tryck, Örebro 2001. 03

Foreword

The Consumer Price Index (CPI) of a country is closely watched by many qualified users both nationally and internationally. The methods used in practical CPI work are of concern to index users as well as to students, index statisticians in other countries, etc.

The handbook presented here describes the methods used in the Swedish CPI, known in Swedish as *Konsumentprisindex (KPI).* It is hoped that the handbook will be useful for several purposes and for several categories of readers:

- Users interested in details of CPI methods
- CPI statisticians in other countries
- Students in Statistics and in Economics
- The Statistics Sweden staff and the CPI Board

The methods of a CPI is constantly exposed to changes. This book primarily describes the situation in year 2000. For earlier or later years some modifications may apply, to some extent indicated in the text.

The author Jörgen Dalén, PhD, has a long experience in price index methodology, previously employed by Statistics Sweden and now working as a consultant.

Statistics Sweden, March 2001

Staffan Wahlström

Roger Pettersson

Contents

General background

Legal acts

The KPI forms part of Sweden's Official Statistics. These are statistical series produced and presented on a continuous basis and funded by Central Government. Participation in the surveys that constitute the KPI is compulsory (SCB's directives on information for the KPI, SCB-FS 1995:9). The information collected for the KPI is protected according to the Secrecy Law (1980:100, ch. 9, §4).

The Net Price Index *(Nettoprisindex, NPI)* also forms part of Sweden's official statistics.

KPI data are also used for calculating the *Harmonised Index for Consumer Prices (HICP)*. The HICP is subject to EU regulation in great detail, which is not presented here.

In a Government Instruction for SCB (1988:137) it is provided that there shall be a KPI Board, consisting of Members from Government Agencies including Sveriges Riksbank (the central bank of Sweden) and of five members with scientific competence in the fields of economics and statistics. Its tasks are specified as follows (§17):

> ” The KPI Board shall handle matters concerning the computation of the KPI and decide on matters of principle concerning the application of the basic principles of the index. It shall also promote the development of methods for the computation of the Consumer Price Index.”

The Board, which meets 3-4 times a year, both examines and decides matters of principle. Its power to make decisions is unique in Government Statistics. In practice, its decision-making is fairly far-reaching, including, for example, setting statistical targets for many complex subindexes such as utility tariffs and banking and insurance, adapting them to reflect changes in market structure, e.g. those due to deregulation, which make the standard fixed basket treatment inappropriate.

Guiding principles of the KPI

There are two schools of thought when it comes to formulating the fundamental principles as to what a CPI should measure. One school starts from the concept of standard of living and states that an index should monitor the change in the cost of maintaining a fixed standard of living. The economic theory of a True Cost of Living Index is relevant here. Exhibit 1 provides a brief account of this theory. A full account can be found, for example, in Diewert (1976).

Exhibit 1: The theory of the cost-of living index

The economic approach to index numbers rests on the assumption of optimising behaviour on the part of economic agents. In the case of consumers it assumes utility maximisation or cost minimisation.

Konüs (1924) was the first to define a *true cost-of-living index for an individual (household)*. When defining Konüs' price index one starts from a utility function ***f(Q)*** with the consumed quantities as arguments and with a cost function

$$C(u,P) = \min_{Q}\{PQ : f(Q) \geq u\}$$

i.e. ***C(u,P)*** is the solution to the problem of minimising the cost $PQ = \sum_{j=1}^{N} P_j Q_j$ for attaining at least the utility ***u*** where ***P*** is the price vector that the individual is faced with.

Konüs' price index is now defined as

$$I_K(P^0, P^1, Q^0, Q^1) = \frac{C(u,P^1)}{C(u,P^0)}$$

The other school simply declares that the index should measure the average price change for a fixed "basket" of consumer products and services. This index philosophy is sometimes referred to as proposing a "pure" or "Laspeyres-type" index.

In the Swedish KPI there has been a long tradition of adherence to the first point of view, taking the theory of the true cost of living index as

its guiding-star, though without this being explicitly laid down. More specific proposals in line with this approach have been put forward by the 1997 Commission but decisions on them have not yet been made. The HICP, on the other hand, follows the second philosophy, its main purpose being to reflect a central bank concept of inflation.

An important concept of economic theory, guiding the functional form of the index, is that of a "superlative" index. It can be proved that certain index formulas, which are symmetric with respect to prices and quantities, provide excellent approximations to a True Cost of Living Index, assuming any one of a number of alternative flexible expressions for the utility function underlying consumer behaviour. Various constraints make the estimation of such formulas difficult in practice, but they nonetheless provide useful criteria when it comes to evaluating the bias of the index.

Uses of the index

The areas for which the KPI is used can be divided into the following three main groups:

For compensation and as a general measure of change in the cost of living of households.

Compensation is declared to be the main purpose of the index. This could involve using it for determining adjustments of

- pensions, social allowances and other income transfers to households from the public sector
- tax rates
- transfers within the private sector, e.g. alimony
- prices in long-term contracts
- interest on index-linked bonds (realränteobligationer)

The KPI is not, however, designed with any one of these particular compensation purposes in mind. It is rather intended to be a general measure of the change in the cost of living. For particular compensation purposes, the choice is to use either the All Items KPI or to design a tailor-made measure based on KPI subindexes and / or other information.

For transforming nominal value changes into volume changes.
These uses include

- Calculating and analysing the purchasing power or changes in real income of households
- Providing a basis for wage negotiations and the analysis of changes in real salaries and wages

Deflating retail trade turnover and private consumption in the National Accounts

For macro-economic policy.
The most important examples are the use of the KPI

- As a general measure of the change in the domestic purchasing power of the Swedish Krona.
- For comparisons with consumer price changes in other countries.
- As a target variable for macro-economic policy, especially the monetary policy of the central bank of Sweden.

History

The 1914 cost of living index

Sweden's first official index compilations were made in 1912 by the Social Welfare Board. They dealt with changes in the prices of foodstuffs since 1905. In 1916 and 1917 changes were made in the methods of calculation reflecting the preliminary findings of the 1914 household budget survey. July 1914 was made the base period of a new index, which covered the full budget of a "normal family" of workers consisting of two adults and two children. The index was calculated with an unchanged budget up to 1932, when the weights were changed to those derived from a 1923 household budget survey, price updated to 1931 and the new index series was linked to the old one. The same was done in 1939, using the 1933 budget survey.

In these early years, the *cost-of-living index* (the name of the index) was a fixed base index based on the Laspeyres formula. It was compiled four times a year except in the very early years.

The Consumption Price Index of Sveriges Riksbank

In 1931 the central bank of Sweden started to compile a new index. This index was developed in co-operation with the Social Welfare Board. It was intended to show, over short intervals of time, the average price changes for the entire consumption of goods and services by the private sector of the economy. The index was adjusted to eliminate effects of changes in indirect taxes and subsidies. Until late 1943, price data were collected from 15 local areas and from then up to 1949 from 25 local areas. From time to time the central bank of Sweden undertook special consumption surveys for the purpose of the index. Between 1937 and 1949 the index was calculated using a geometric mean in accordance with the principle of a Törnqvist price index. It was chained by multiplying week to week and month to month indexes. However, at the end of each year (at first in each quarter), adjustments were made based on direct comparisons between the current period and some earlier period.

From 1949 the Social Welfare Board took over the responsibility for this index. From then on (and before 1937) the index was calculated according to an arithmetic mean formula.

The 1943 Commission

During the second world war, certain difficulties arose in calculating the index because of the considerable changes in the war-time expenditure patterns and even the disappearance of some commodities. This motivated the appointment of the 1943 Index Commission (SOU, 1943), which undertook the first review of the foundations of the index . The idea that, as a matter of principle, the index should refer to the same standard of living in two time periods was then explicitly laid down.

In 1943, following the report of the 1943 Commission, the method of calculating the index, still termed "Cost of Living Index", was revised. From then on, it was computed as a chain index with annual links in which the weights applied to the current year. Two indexes were compiled - one including and one excluding direct taxes. Moreover, the number of local areas and of representative commodities were increased considerably.

The 1952 Commission

Up to June 1954 the Cost of Living Index was computed quarterly and new weights according to eq. (4) below, referring to the current year, were introduced in the December index. In July 1954 the new monthly index, now called the "Consumer Price Index" (*Konsumentprisindex*) - as proposed by a Government Commission appointed in 1952 (SOU, 1953) - was initiated and the distinction between a short-term and a long-term index was formulated. The practical implementation of the proposals in the 1943 and 1952 Government Commissions is carefully described in Socialstyrelsen (1961).

The 1955 report on the Housing subindex

A special investigation into the problems of estimating price change for both rented and owner-occupied housing was undertaken by a special commission, appointed in 1955. A kind of quality adjusted unit value index was proposed for rented apartments, with newly

built apartments being introduced into the numerator. A special sample survey of rented apartments was proposed and initiated soon thereafter.

For owner-occupied housing, an expenditure approach was proposed, where the following components were included: i) depreciation, ii) mortgage interest, iii) maintenance and repairs, iv) insurance, and v) water and sewage. The detailed implementation of this approach has varied somewhat over the years. The method used currently is described below.

The 1997 Commission

In 1997, a new Government Commission was set up to review the basic principles of the Consumer Price Index, (*KPI*), a long time having elapsed since the preceding review. It published its final report in late 1999 (KPI COMMISSION, 1999). Among its proposals are:

- A new chain index construction, involving full year-to-year index links according to the superlative Walsh index formula, with a final Laspeyres year-to-month link covering the last 2 years.
- Partially new principles for the choice of elementary aggregate index formulae, taking account of differences in price elasticities of demand at different levels of the index.
- A new treatment of owner-occupied housing based on the user-cost principle and on an assumed constant relation within each index link between the long-term nominal interest rate and the (expected) long-term capital gain, with a constant depreciation rate. The resulting index for capital cost, excluding property tax, would then move according to an index of dwelling prices, the weight corresponding to capital cost adjusted for tax effects.
- Inclusion of income-related charges for child care and care of the elderly based on the prices that the households actually pay for the services, regardless of whether changes of those prices are caused by rising incomes or by changed fee structures.
- Production of an official constant-tax index as a measure of underlying inflation. This index corresponds to the UND1X index produced today and is described below.

The Commission proposals have been reviewed by organisations and agencies with an interest in consumer price measurement. Its

proposals gained general acceptance except for those concerning owner-occupied housing and to some extent income-related charges. A Government decision on implementation is still not taken, (March 2001), however.

Historical data (Myrdal-Bouvins index)

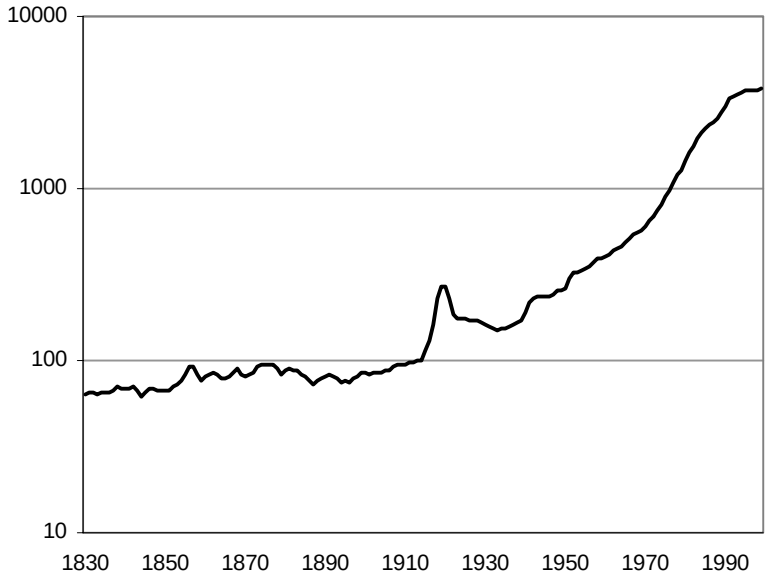
The first Swedish price index series starts from 1830 and continues up to 1914. It was developed in the early 1930's as a retrospective study under the supervision of Gunnar Myrdal of the University of Stockholm. It is known under the name of the Myrdal-Bouvin index .

The sources used were estimated average market prices, using purchase prices paid by official institutions and price quotations appearing in the press. In addition, data were obtained from business files and accounts; in certain cases prices recorded by private persons were used.

This index used the ten-year period 1861–1870 as its reference period. The principle behind the calculation was to follow the cost of an adult's annual consumption.

Inflation from 1830 to today.
Price Index excluding direct taxes and social benefits

14


10000
1000
100
10
1830
1850
1870
1890
1910
1930
1950
1970
1990

Sources of information

Information on current KPI results can be obtained from a variety of sources, both electronically and in paper format. Here we give an overview over the most important such sources

Press release

Each month a press release is published, with information on the KPI change since the previous month and over the last 12 months. Data on the NPI, HICP and on underlying inflation are also presented. In addition to being published in paper form it is posted on the SCB website at the same moment as the new data are made public.

The press release is also published in the SCB monthly publication *SCB Indikatorer*.

A short text in Swedish and English is released on the computer screens of a number of international news agencies (e.g. Reuter) at the moment of publication.

Statistical Report PR 14, (monthly)

This printed report is published every month on the day the index is released. It contains the same information as the press release. It also presents index numbers and rates of change for the All Items KPI and for the main groups for the most recent month and monthly for the last 12 months. There is also a table for main groups and subgroups showing index numbers and rates of change for the last two months. In this table, the weights for the main groups and subgroups are also presented.

The publication also provides brief information, rates of change and, where appropriate, index numbers for the NPI, HICP (both for Sweden and for other EU countries) and the underlying inflation as well as the latest Basic Amount (see below) figure. Diagrams show rates of KPI change and international inflation figures (according to the HICP).

In every issue there is also a short English summary.

Statistical Report, PR 15, (annually)

This publication gives more information about KPI calculation, content and use. More elaborate presentations are provided concerning the NPI, HICP, underlying inflation, Basic Amount (see below), old index series and on Purchasing Power Parities.

Time series of index numbers from 1980 for the All Item KPI and the main groups are presented. For subgroups, index numbers are also provided from 1980. The publication also presents time series on the long-term index and on weights. Rates of change for the last few years are given. Average prices for representative products that have been in the index during the last two years are presented for most COICOP groups.

The Basic Amount is given for all years in which it has existed and the publication also presents an older time series from 1830, which can be used as an approximate measure of price changes over a longer period.

For the NPI, HICP and underlying inflation, figures corresponding to those for the KPI are given, but in somewhat less detail.

Brief English versions of the Swedish texts are also published in this report.

SCB's website

On the SCB website (www.scb.se), statistical series on the KPI, HICP, NPI, underlying inflation, inflation rate and the basic amount are available. Time series are usually from 1980 or earlier (from 1995 for the HICP).

Press releases are also published on the website and the most recent inflation figure is given on the home page directly.

Sweden's Statistical Databases

There is free access to SCB's databases through the website. The following data can be obtained:

Consumer Price Index (1980=100), fixed index numbers

Consumer Price Index (1980=100), Shadow Index Numbers and Indices for Main Groups

Consumer Price Index (1980=100), Shadow Index Numbers and indices for Main Groups, Annual Averages

Consumer Price Index (1949=100)

Consumer Price Index, yearly changes, per cent (Inflation Rate)

Net Price Index (1980=100)

Net Price Index (1959=100)

Net Price Index, yearly changes, per cent

Harmonised Index of Consumer Prices (1996=100)

Harmonised Index of Consumer Prices, yearly changes, per cent

Consumer Price Index, yearly changes, per cent (Inflation Rate)

Inflation Rate according to CPI, NPI and HICP (Graph)

Underlying inflation rate according to UND1X, monthly changes

Underlying inflation rate according to UND1X, annual changes

Underlying inflation rate according to UNDINHX, monthly changes

Underlying inflation rate according to UNDINHX, annual changes

Basic Amount, annual data since 1960

General principles and methodology of the KPI today

The basic principle as formulated by the 1943 Commission

The 1943 Commission was the first to formulate the foundations of the KPI in some depth. It derived a unique expression for a Consumer Price Index by viewing price change over a year as a series of stepwise changes according to the theoretical concept of a Divisia index. In each of these steps, value or quantity weights were held constant. This led to two possible formulations of the desired index, as a geometric or an arithmetic mean. The arithmetic mean became the Commission's final recommendation. At that time the index was only published quarterly but applied to the present monthly index, this corresponds to the following formulation of a single index link from December (12) year y-1 to December year y:

$$I_{y-1,12}^{y,12} = \frac{\sum_k P_k^{y,12} Q_k^y}{\sum_k P_k^{y-1,12} Q_k^y} \qquad (1)$$

What this means is that the basic link in the index is defined as the price change between two successive Decembers, with quantity weights representing the whole calendar year between them. Longer term comparisons would be obtained by multiplying successive links like (1) together into a *chain index*.

The chain index today

The present index construction still follows the principles of formula (1). The KPI is a chain index with annual links going from December one year to December in the following year which are multiplied to calculate the long index series. This at present has 1980 as its index reference year. For each successive new link, weights are recalculated based on new information.

We make a distinction between this long-term link (***L***), which uses quantity weights Q_y from year y and the short-term link (***K***) which uses quantity weights Q_{y-1} from year ***y-1***. The definitions of the links are

$$L_{y-1,12}^{y,12} = \frac{\sum_k P_k^{y,12} Q_k^y}{\sum_k P_k^{y-1,12} Q_k^y} \tag{2}$$

and

$$K_{y-1,12}^{y,m} = \frac{\sum_k P_k^{y,m} Q_k^{y-1}}{\sum_k P_k^{y-1,12} Q_k^{y-1}}, \tag{3}$$

where summation is over ***N*** products with subscripts ***k*** (subscripts will later be dropped when there is no risk for misunderstanding).

The chained index from base year ***0*** to a month ***m*** in year ***Y*** will now be[1]:

$$KPI_0^{Y,m} = I_0^{0,12} * L_{0,12}^{1,12} * \ldots * L_{y-1,12}^{y,12} * \ldots * L_{Y-2,12}^{Y-1,12} * K_{Y-1,12}^{Y,m} \tag{4}$$

This means that in the long run the KPI series only depends on the long-term links; the short-term links (which use weights for the previous year) are successively replaced by their long-term counterparts (which use weights for the year just ending) for December each year.

The KPI according to (4) is presented in a series with 1980 as the index reference year (=100). This long-term movement is used as a measure for government compensation purposes -- e.g. for pensions and child support rules for divorced parents (see also the Section below on the basic amount). Also many private contracts have an index clause under which payments follow the long-term KPI movement.

---

[1] The extra link from 0 to 0,12 is needed in order to use a full year as the index reference period. Its exact definition is: $I_0^{0,12} = L_{-1,12}^{0,12} \Big/ \frac{1}{12} \sum_m K_{-1,12}^{0,m}$

12-month changes

When 12-month changes are published, the long-term index component is replaced by the corresponding short-term component instead of looking at the changes in the index series according to (4) directly. One reason for this procedure is that the long-term component includes substitution effects that do not represent the current 12-month period, another reason is the need for international comparability. The price change from ***Y-1,m*** to ***Y,m*** is thus calculated as:

$$KPI_{Y-1,m}^{Y,m} = K_{Y-1,12}^{Y,m} K_{Y-2,12}^{Y-1,12} \Big/ K_{Y-2,12}^{Y-1,m} = \frac{KPI_0^{Y,m}}{KPI_0^{Y-1,m}} \frac{K_{Y-2,12}^{Y-1,12}}{L_{Y-2,12}^{Y-1,12}} \quad (5)$$

The 12-month changes according to (5) are often used for inflation monitoring. For example, the inflation target of the central bank of Sweden follows expression (5).

Index aggregation at the higher level

In practice, of course, no quantities can be directly observed at the higher KPI aggregation levels. The National Accounts (NA) are instead used to provide value weights, ***V***. These are defined as price times quantity. In our notation ***V=P*Q***. Up-to-date NA consumption values today exist for more than 100 consumption categories. During the annual weight revision which takes place in January and early February, new NA values are brought into the index. For example, early in year ***y*** values for year ***y-1*** – denoted V^{y-1} are used both to replace the short term link of December ***y-1*** with a long term link *and* for the new short term links for year ***y***. These values are price updated (or ”backdated”) to the price reference period of the link in question.

In precise mathematical terms, and remembering that $\boldsymbol{V=PQ}$, we obtain:

Long-term weight for $\boldsymbol{y\text{-}1}$: $$W^{y-1,L} = \frac{V^{y-1} \Big/ \frac{P^{y-1}}{P^{y-2,12}}}{\sum V^{y-1} \Big/ \frac{P^{y-1}}{P^{y-2,12}}} \quad (6)$$

to be plugged into $$L_{y-2,12}^{y-1,12} = \sum W^{y-1,L} \frac{P^{y-1,12}}{P^{y-2,12}} \quad (7)$$

and short-term weight for year $\boldsymbol{y}$: $$W^{y,K} = \frac{V^{y-1} \frac{P^{y-1,12}}{P^{y-2,12}} \Big/ \frac{P^{y-1}}{P^{y-2,12}}}{\sum V^{y-1} \frac{P^{y-1,12}}{P^{y-2,12}} \Big/ \frac{P^{y-1}}{P^{y-2,12}}} \quad (8)$$

to be plugged into $$K_{y-1,12}^{y,m} = \sum W^{y,K} \frac{P^{y,m}}{P^{y-1,12}} . \quad (9)$$

The indexes $\frac{P^{y-1}}{P^{y-2,12}}$ are computed as arithmetic averages of the 12 monthly short-term links from January $\boldsymbol{y\text{-}1}$ up to December $\boldsymbol{y\text{-}1}$.

Index aggregation at lower levels

Below the level where National Accounts weights are available, procedures necessarily vary somewhat. They depend on the price structures, market situations and access to weight information that exist in different product areas.

Intermediate computation levels in the CPI are those between the highest aggregation level, where up-to-date information is available, and the lowest aggregation level, where prices of *product-offers* (a unique product in a unique outlet at a certain point in time) are combined into so called *elementary aggregates*. At this level, various kinds of information are used for weights, much of which in practice relate to earlier years.

At the lowest, *elementary aggregation level,* a special formula is normally used for computing a subindex. It is called the *RA formula.* In some other cases the simpler ratio of average prices - the *A formula* - is used. These two formulas are shown in Table 1, in their weighted and unweighted forms.

Table 1: Elementary aggregate formulas

For-mula	*Unweighted form*	*Weighted form*
RA	$I_{y-1,12}^{y,m} = \dfrac{\sum_k p_k^{y,m} / (p_k^{y-1,12} + p_k^{y,m})}{\sum_k p_k^{y-1,12} / (p_k^{y-1,12} + p_k^{y,m})}$ (10)	$\boldsymbol{I_{y-1,12}^{y,m} = \dfrac{\sum_k w_k p_k^{y,m} /(p_k^{y-1,12} + p_k^{y,m})}{\sum_k w_k p_k^{y-1,12} /(p_k^{y-1,12} + p_k^{y,m})}}$ (11)
A	$I_{y-1,12}^{y,m} = \dfrac{\frac{1}{n}\sum_k p_k^{y,m}}{\frac{1}{n}\sum_k p_k^{y-1,12}}$ (12)	$I_{y-1,12}^{y,m} = \dfrac{\sum_k w_k p_k^{y,m}}{\sum_k w_k p_k^{y-1,12}}$ (13)

The motivation for the RA formula is partly that it can be seen as an approximation to the basic formula (1) above. See Dalén (1992) for details concerning this approximation. Also, it does not suffer from the upward bias that is present in the simpler, so called ”average of relatives” formula. The RA formula can also be shown to approximate a geometric mean index quite well and is thus consistent with unitary price elasticity of demand[2]. (In Dalén, 1992, a more detailed discussion of this issue is provided.)

The A formula is primarily used when quantity weights are available, i.e., when the w_k can be interpreted as volumes. Its unweighted form is appropriate, when probability proportional to size (pps) sampling (discussed later under sampling) using volume correlated size measures has been used.

In certain subindexes, special calculation methods are used. These are mentioned in the text below.

[2] When price elasticity is equal to one, the consumer changes her purchased quantities in the same proportion but opposite direction as a change in relative price levels, so that the relative purchasing values of different products remain constant.

Central and local price collection

There are basically two different modes of price collection. For most services and some goods the central staff collects the prices, either by telephone or by a small-scale mail survey using a shuttle form[3], and enters them directly into the computer, usually into an Excel Workbook where the subindex computation is done. This procedure is usually referred to as *central price collection*. For rents there is a large-scale mail survey carried out and compiled by a separate bureau of the SCB.

For many goods and services there is *local price collection*. In these cases the product specifications are established centrally. In some cases the specifications are loose and the price collectors then choose the most sold product within the specification in a selected outlet. In other cases the specifications are tight and the price collectors then seek the corresponding products in the sampled outlet or, if one is not found, omit the product.

Products are divided into product groups according to the National Accounts and other sources of information. Within each product group one or more centrally defined product specifications are set out, often only one. All in all there are some 200 such *representative products* in the index. For each of these a commodity specification is created at the central office. The price collector is then asked to find that particular variety fitting the specification which is the most sold within the sampled outlet.

This means that prices are collected by price collectors by visiting the outlets directly. (The price collectors are located all over Sweden and are employed on a full or part-time basis as interviewers for all surveys managed by the SCB.) The outlets are visited on an optional day in the week in which the 15th of the month occurs. In outlets where few prices are to be collected and no complex quality adjustment is needed, price collectors may use the telephone.

In each outlet, from one up to as many as 500 prices are observed. In all some 20,000 prices in 700 outlets are observed in the local price

[3] In a shuttle form, prices are entered in each month and the form is mailed back and forth between the central office and the respondent (outlet).

collection. Prices are entered into forms which are later scanned; for clothing shuttle forms are used.

Before product-wise price indexes are calculated, data are checked for large deviations compared with last month's price. They are edited if errors are discovered, contacts with price collectors being made where necessary. When all subindexes have been finalised in their preliminary versions, a general meeting of the KPI staff is held where the results are carefully scrutinised.

Sources of weights

At the aggregate level, National Accounts data on Household Consumption are used. These data are divided into more than 100 categories and are very early estimates of consumption in the year just ended. They use preliminary data for the first three quarters of that year and add a projection for the fourth quarter, made in December. Retail trade data are the most important input into these estimates.

At lower levels of estimation a variety of sources is used, some of which are mentioned in more detail below. Two particularly important sources deserve mentioning here.

Household Budget Surveys (HBS) are used for decomposing many NA categories into smaller groups. Examples are clothing into various garments and furniture into tables, chairs, beds etc. Due to large error margins, several years of HBS data have to be aggregated. They are updated annually with the best information that is available at the time. Classification problems add to the difficulties in this work.

For food, the primary source for subweights is data from the National Agricultural Agency, which are published with a one-year lag in relation to the CPI requirements. For subgroups with large recent price changes (coffee being a major example), adjustments are made based on other information, for example from wholesalers. In other cases where the reliability of these data can be questioned, complementary information is collected, whenever possible.

Sampling methods

With respect to sampling, the KPI exhibits a mixed picture. Probability sampling is a preferred method and the predominant technique here is *order pps sampling*. Particularly for products, however, probability sampling is often not practicable and other, purposive, methods are then used. Cutoff sampling and quota sampling are the most common among these.

Where distinguishable subgroups with weights within a product group exist which are known or can be estimated and are of considerable size, a stratification is often carried out before sampling. Strata may be by subgroups of products, by outlets of a certain type or by providers of a certain service. However, unlike many other countries there is generally no regional stratification level in the KPI.

In Annex (p. 79) an overview is given over the different sampling techniques used, by product group. The various methods are now presented in greater detail.

Outlet sampling in local price collection

In product groups, where local price collection is used, outlets are divided into 38 retail trade and service strata according to SNI code (Swedish Standard Industrial Classification which closely follows NACE, Rev. 1, the EU standard). In each stratum a sample of outlets is drawn from the Central Business Register by an *order pps technique* (see Exhibit 2). This first gross sample is drawn about 6 months before the year in which the sample is to be used. This sample is then screened, in October-November, both in the central office and by the price collectors visiting the outlets, and some of the outlets initially drawn are excluded for various reasons. For example, they may be head offices rather than outlets, or they may not sell any of the sampled products. In the case of clothing, a purposive allocation of particular garments to the sampled outlets takes place. In some outlets only a few prices are collected, whereas in others (for example supermarkets and department stores) the number is very large. The end result of this sampling process is a net sample of outlets whose size is fixed in advance.

Positive co-ordination of outlet samples between years is obtained through the use of random numbers permanently associated with every outlet in the sampling frame. See, e.g., Ohlsson (1995) for a

description of this technique. Sampling rotation is performed so that 20 per cent of the random numbers are changed every year. Combined with changes in the sampling frames, this results in some 70-75 per cent of outlets remaining in the sample from one year to the next.

Product sampling in local price collection, pps method

For foodstuffs and other everyday commodities, product sampling is made by order pps using sampling frames provided by the three major wholesaler / retailer chains in Sweden. These sampling frames cover all products sold in one year prior to the index year which are shipped through the wholesaler. These are estimated to be some 80 per cent of all goods sold in supermarkets. The sampling frame covers products at a detailed level, where a unique price normally exists, for example ”Coca-Cola, plastic bottle, 2l”.

Three different product samples of 400 items each are created, one for each chain. The product sample is then matched to the outlet sample according to the chain to which a sampled outlet belongs. Only product-offers in the sampled outlet are thus included. This reduces the effective product sample size in each outlet to some 250-300 product-offers.

Exhibit 2: Order pps sampling

Where sampling frames exist, some measure of size is usually attached to the units in the frame which can be used as a weight in the index formula that we want to estimate. Examples are number of employees in an outlet, sales values of different brands of a product or number of cars purchased of a certain model. This situation favours the use of *pps* (probability proportional to size) *sampling*. A particular variant of pps sampling that has been found useful for the KPI is *order pps sampling*.

The theory behind order pps sampling is given by Rosén (1997a and 1997b). Here we will only describe the special case of its application in the KPI. A uniform random number U_i between 0 and 1 and a variable $z_i=nx_i/\Sigma_i x_i$, where x_i is a size measure, are associated with each sampling unit i and a *ranking variable* Q is constructed as a function of U and z. The units in the universe are then sorted in ascending order and the n units with the smallest value of the ranking variable are included in the sample. Two important examples of such ranking variables are:

$$Q_i = \frac{U_i}{z_i} \qquad (14)$$

$$Q_i = \frac{U_i(1-z_i)}{z_i(1-U_i)} \qquad (15)$$

Units, where $z_i \geq 1$ are first included with certainty and excluded from the frame. The procedure is then repeated until there are no such units in the frame after which the sampling procedure takes place according to (14) or (15).

Order sampling procedures are not exactly pps, but in samples of sufficient size they can be shown to be approximately pps. The second variant of ranking variable Q_i (sometimes referred to as Pareto pps) is a marginally better choice and is therefore normally used.

An important practical advantage of the order sampling technique is that it is able to cope with out-of-scope units (for example an outlet that does not sell any sampled product) in a statistically sound way, which allows a fixed pre-determined sample size. In Table 2, we give an example of how this works for a sample of 3 outlets out of 10. We have ordered a sampling frame of outlets, where X_i is the number of employees in the outlet, in ascending order with respect to the ranking variable Q_i and our first sample turns out to consist of outlets 6, 1 and 8. However, say that we now discover that outlet 1 is out of scope. We then turn to the fourth unit in order – outlet 9 -

and include that one instead, so that the final sample consists of outlets 6, 8 and 9.

Table 2: Order pps sample of 3 out of 10 outlets with out-of-scope units. The frame is arranged in ascending order with respect to the ranking variable.

Outlet	X_i	U_i	Q_i	Sample
6	25	0.755509	0.036943	X
1	13	0.198082	0.207721	(X)
8	6	0.915131	0.310666	X
9	11	0.277131	0.346024	X
10	8	0.834138	0.380468	
7	10	0.709046	0.412599	
4	9	0.46373	0.580264	
3	5	0.500162	1.25	
5	1	0.067941	1.836435	
2	2	0.297524	2.926051	

Product sampling in local price collection, purposive method

In other cases of local price collection, a product specification is drawn up in the central office. The price collector is then instructed to choose the variety that is *most sold* in terms of volume within this specification in the sampled outlet. This method could be viewed as a special case of cutoff sampling (see below). Outlet staff are often asked to assist in the judgements that have to be made when applying this criterion. For new varieties it is sometimes difficult to know in advance which items will sell best and it is thus not certain that the selection will in practice always catch the variety that will be most sold. The ”most sold” rule also applies to replacements, when an item disappears from the market.

Exhibit 3: Item specification and selected variety, two examples

- **Paint**. Central specification: *For outdoor painting of windows. White, 1 litre.* The price collector selects the most sold variety in the outlet.
- **Sofas**. Central specification: *None.* The price collector selects the most sold variety in the outlet.

Sampling in central price collection

In those product groups where prices are not collected locally, a wide spectrum of both probability and non-probability methods is used. These methods are described below for particular product groups in connection with descriptions of their other methodological aspects. Here we present two purposive approaches that are quite widely employed in KPI surveys.

Cutoff sampling. When good sampling frames are not available, probability sampling is generally impracticable. Also, for surveys relating to products with a very small weight, small samples of perhaps only 1-10 units are wanted and it may then be more practical to cover only the largest.

Cutoff sampling means selecting only units with the largest subweights (sales, turnover, population etc.). This method is reasonable if these units together cover a large share of the total and / or if they are more stable so that it is easier to obtain information from them.

In some cases the sizes of the sampling units are only known approximately. In these cases one can speak of a "judgmental cutoff" method.

Exhibit 4: Example of cutoff sampling

Five companies out of 30 in Sweden cover 95 per cent of the total market for heating oil used in single-family houses. These five companies are included with certainty and the other 25 companies have zero probability of inclusion.

Quota sampling. Quota sampling means that the sample is selected judgementally but in such a manner that the proportions in the sample with respect to a number of important price-determining characteristics are approximately the same as those in the relevant universe. For an example of quota sampling, see the Section on package holidays below. Quota sampling is considered a reasonable option, when the number of stratifying variables is large in relation to the sample size.

Complete coverage. Finally, in some cases it is possible to cover virtually the whole universe. Usually this is the case for areas with Government owned monopolies. In the KPI, this is for example the case with alcoholic beverages, where the Systembolaget makes monthly price index calculations covering its entire sales which are directly entered as KPI subindexes for those products. For medicines sold by the Apoteksbolaget there is a similar situation.

New and disappearing products and outlets, replacements

There are two ways of introducing new products and outlets into the KPI sample:

Re-sampling is done when a new index link is started up in December, when both the old and the new product and outlet samples are measured. In this way an *overlap* is created, so that the old sample is used for backwards and the new sample for forwards comparisons, without any explicit quality adjustment. To the extent that a) the market is in equilibrium so that the price differentials between the old and the new sample in December reflect genuine consumer valuation of quality differences and b) both samples adequately represent the universe, the estimate of price change is unbiased. Where an outlet remains in the sample from one year to the next, product-offers are normally not resampled in December.

A *replacement* of a particular product-offer with another one, usually in the same outlet, is caused by the disappearance or reduced importance of a product-offer. In this case a quality adjustment is normally done (see below).

Outlets are not replaced. New outlets are only introduced in the course of resampling in December, in connection with the start-up of a new link. In those very few cases where an outlet is closed down or price measurements cease to be possible for some other reason, that outlet's product-offers are deleted.

The rules governing replacements are best described by quoting the instructions given to price collectors, shown in Exhibit 5. Note especially the requirement to replace product-offers, which are not longer among the most sold ones, throughout the year.

Exhibit 5: Excerpts from instructions to price collectors regarding replacements

"It is important that you try to measure the price for the same variety throughout the year. At least once a year, however, you should check with your contact person that the selected variety is still among the most sold. For many product groups this should be done more often. When the selected variety is no longer among the most sold you should choose another one. But you should not change variety during the year if the selected variety is slightly less sold than before. It has to sell significantly less than the most sold one for a replacement to be made during the year.

You should avoid making these checks in December. If the outlet remains in the sample also next year, it is an advantage if the selected varieties are used across the turn of the year."

Quality change

Several methods are used in the KPI to allow for quality change in different product categories. We first describe two main approaches which are used for local price collection products other than clothing. These are divided into two categories - quality adjustment products and non-quality adjustment products.

Quality adjustment products in local price collection

Here, the price collectors perform the adjustments. The method is again best described by quoting the instructions to the price collectors.

Exhibit 6: Excerpts from instructions to price collectors on quality adjustment

When should a quality adjustment be made?
A quality adjustment when replacing a product-offer shall be made for so called quality adjustment products, as given by (..list..).
A quality adjustment shall not be done when you replace a product in the reference period[4]. Also, if a price quote does not exist for the product earlier in the year or in the reference period, no quality adjustment can be made, since you do not then have information of any old variety.

Principles for quality adjustment
By quality difference is understood a difference in material or design. Differences due to fashion changes are not counted as quality differences. Factors in the products that you should consider are e.g. function, comfort, durability, security, guarantees and easiness of handling.

Differences in quality are to be valued from the viewpoint of the consumer. Differences in e.g. production or distribution costs should not be considered. You should instead try to assess how the average consumer experiences differences in material and design. This is difficult and in practice it means that you will have to use your own assessment of the differences.
Please note that the quality of a variety can also change by changing the service provided to the consumer. Examples are rules concerning guarantees and home delivery.

Differences in material and design are often difficult to detect. Also, the old variety is rarely available for inspection. Ask for help from a shop assistant in detecting the differences. But remember that it is your own valuation and not that of the assistant that you are to report. The valuation of the assistant should only guide your own valuation. This is because it is difficult for the assistant, due to her role as a seller, to make a true personal valuation.

Two varieties differ in a number of characteristics. One way of making the valuation could therefore be characteristic by characteristic, in money terms, summing the valuation over the characteristics.

You should ignore the price differences between the varieties. If the price difference is large, it does not necessarily mean that the quality difference is large. There can also be a quality difference although the price is the same. Companies often compensate themselves for changes in cost in connection with the introduction of a new model. Therefore, a higher price does not necessarily mean that the new model is of higher quality than the old one.

Quality adjustments according to these principles are difficult to make. If you feel uncertain about your valuation, you should note your views and observations in the bottom part of the form. Remember too that you will gather a considerable knowledge of the products through your CPI work and that by and by you will be in a better position to make the adjustments from a consumer point of view.

When you specify a quality adjustment you state how much better or worse the new variety is, in money terms. When no quality difference xists between the new and the old variety, you simply state the adjustment as 0 (zero)."

[4] December.

The price collector's adjustment is added to the reference price, resulting in an adjusted reference price. This is according to the following recursive formula:

$$p_{0(m),i} = p_{0(m-1),i}\left[1 + \frac{K_{m,i}}{p^{reg}_{m-1,i}}\right] \qquad (16)$$

$K_{m,i}$ denotes the quality adjustment (positive or negative) for a replacement from month m-1 to m, as stated in kronor by the price collector. $p_{0(m),i}$ denotes the adjusted reference price in month m (note that the reference price may be adjusted several times in a year and that $p_{0(0),i}=p_{0i}$ is the actual, observed reference price). $p^{reg}_{m-1,i}$ denotes the *regular price* in month m-1. (This term is not rigorously defined but can, if it differs from the actual price, usually be interpreted as a "normal" price as opposed to a promotional price, or as a price existing prior to a sales price. If it is stated on the price label alongside the current price, the price collectors are instructed to note it as part of their price collection work.)

Non-quality adjustment products in local price collection

For products where quality adjustments are not made, but package size has been altered, only new package sizes where the quantity change is less than 50 per cent are accepted as replacements. A proportional adjustment is then made so that the price effectively becomes a price per quantity unit.

In other cases, where a product-offer can no longer be found, it is deleted and price change is computed over the rest of the product-offers in the product group.

Other product groups

Quality adjustment methods for other product groups are described in the Section *KPI by product group* below. Here we shall only give a short summary.

Clothing: A hedonic method is used for garments.

New cars: An expert panel takes the decisions on the positive or negative value of changes in specification for new car models.

Used cars: A regression model, developed by a private company, is used for holding mileage constant, when comparing prices.

PCs: Monthly chained indexes consisting of models that are found in both months are multiplied. The method could be labelled as a matched model index, using a monthly overlap method. Or a *monthly chained overlap method*, for short.

Overlap method when chaining in December

In many product groups samples are renewed in December at the start-up of a new index link. This means that in December prices are collected for both the old and the new sample. An *overlap* method is then applied for effectively evaluating the quality differences between the two samples.

In many products groups this annual overlap is in practice the most influential method for dealing with product and outlet dynamics. For unbiasedness, it rests on the assumption that the quality differences between the old and the new samples are on average equal to the price differences between the two samples.

The KPI by product group

The KPI uses the same product classification as the HICP. This classification is denoted COICOP / HICP.

In the Annex, a brief summary account, by Coicop group, is given of the content and methods used in different subindexes, referring back to the general description above. Here some further details are provided to complement the information in the Annex.

Food and non-alcoholic beverages, tobacco

Stratified sampling of outlets is done by order pps from the Business Register, stratified into retail trade categories by 5-digit code. The pps size measure is the number of employees (plus 1), which is correlated with the total sales value in an outlet (but not necessarily with the total sales value in a certain product stratum). Out-of-scope units are eliminated from a gross sample list so as to obtain an exact predetermined net sample size. Outlet strata are weighted, using retail trade survey results relating to the distribution of revenue over different product groups. This distribution is usually more aggregated than the Coicop groupings used in the KPI, so additional information and judgement is needed.

Weights for detailed subgroups (strata) are obtained from statistical information produced by the National Agricultural Agency. Order pps sampling of products is applied within most product strata. The measures of size are historic sales data from the three major wholesalers, who supply the major part of the goods sold in the outlets. In this way we obtain three different probability samples, one for each wholesaler, each consisting of about 400 commodities.

The commodity sample is finally matched to the outlet sample so that each sampled outlet is allocated a commodity sample according to the wholesaler to which it belongs and according to the products which it actually sells. This means that the effective size of the product-offer sample varies between outlets. If a product-offer disappears, it is usually not replaced by another one but is instead deleted, price change being calculated over the remaining product-offers in the same stratum. If only package size changes and does so by less than 50 per cent, then the reference price is changed

proportionally. A new sample is normally drawn for each December so that an overlap is established.

For fresh food (bread, fruit, fish and to some extent meat), sampling by the price collector is used – within a central specification. Here the ”most sold” criterion applies where more than one product-offer is found that matches the central specification.

No explicit adjustment is made for seasonal products, e.g. fruits and vegetables. Instead, similar products are considered essentially equivalent (new and old potatoes, small citrus fruits, different varieties of apples etc.). Products with strongly seasonal availability in the market are excluded, e.g., fresh strawberries, raspberries, cherries, peaches and plums. Grapes are excluded on account of their marked quality variability.

Resampling, both of outlets and products is done annually, so that an overlap is created in December, when prices for both the old and new sample are obtained.

Alcoholic beverages

Alcoholic beverages, with an alcoholic content of more than 3.5 per cent, are sold only in special stores belonging to *Systembolaget*, a Government-owned monopoly chain. This company computes monthly price indexes for the KPI covering its total sales in this month compared to the reference month. The formula is equivalent to the formula used at the upper level of the index (equations 8–9 above). However, only varieties sold in both of these months are included in the calculation, so no replacements are made. This means that some 10 per cent of the turnover is excluded from the calculation.

Low alcoholic beer (3.5 per cent or less alcohol) and some other low-alcohol beverages are mainly sold in supermarkets and are treated according to the same methods as for food above. Alcohol sold in airports or boats (tax-free) is excluded.

Clothing and footwear

The rapid changes of items in the clothing market make advanced and complex methods for dealing with replacements and quality adjustments necessary. For this reason *hedonic models* of the

relationship between garment prices and characteristics (Coicop 03.1.2) are formulated and estimated. (The other Coicop clothing and footwear subgroups are treated by traditional methods.) Due to the needs of the hedonic method, there are also specific sampling methods in the garments product group.

There are 26 central representative products (specifications) for clothing, for 23 of which there are hedonic models. These specifications are purposively selected to represent the whole clothing area. In each outlet 2-10 variants within a specification are chosen by the price collectors. In all, about 70 outlets are sampled with the order pps method but only about 25 outlets are surveyed for each specification. The distribution of the specification within the random outlet sample is done purposively, in the light of information about the range of products sold in each sampled outlet. Outlets not covered are those where clothing sales forms only a small proportion of total sales, e.g. some large supermarkets. Mail order is presently included on a trial basis but not represented according to its real market share.

In the late 1980's there was found to be a serious bias in the clothing index that was due to the purposive sampling used in the outlets. In making the base period selection, price collectors tended to choose variants sold at regular, and thus relatively high prices, so under-representing the lower sales prices. In the following months of the index year, with these, unchanged, samples, the share of sales prices in the sample data largely reflected their true share in the whole universe. The result was a serious underestimation of clothing inflation. The problem was dealt with in 1993 . Starting in that year, two measures were taken. Firstly, the initial selection of product-offers is now done in November of the year before the index year. Secondly, the index is adjusted by a factor compensating for the imbalance in the importance of sales prices. A sales price adjustment is made which is equal to the ratio between the sales price effect in the final month (December) of the old sample to that of the same (first) month in the new sample. This sales price effect is defined as an index computed for actual prices compared to one calculated for "regular" prices (see above in the section on Quality change p. 29) as reported by the price collectors. This adjustment is applied during the following year. The average adjustment is around 2 per cent (upwards), somewhat smaller for footwear.

Seven different hedonic models are estimated, with 2-6 product specifications in each model. The dependent variable is the log of the regular price of the variety. The independent variables are dummy variables which fall into four different classes:

1. The first class relates to outlet types which tend to use different mark-up percentages and also correspond to the preferences of different customer segments.

2. The second class refers to the origin of the garments, i.e. their brand and country of production.

3. The third class relates to physical characteristics of the garments that are supposed to affect their relative valuations by consumers. Usually, characteristics are chosen that correspond to varying raw material and labour cost.

4. The fourth class reflects time. Since data for more than one month are used, dummy variables describing the month to which the prices refer, have to be included.

Coefficients are thus estimated at the start of every year, based on price and characteristics data from the clothing survey itself in the year before. The coefficients are then used to adjust current prices each month, when replacements occur. The following formula, where the * variables represent adjusted prices, conveys the general idea:

$$I_{mv} = \frac{\sum_{i \in mv} p^{*}_{mit} \,/\, (p_{im0} + p^{*}_{imt})}{\sum_{i \in mv} p_{im0} \,/\, (p_{im0} + p^{*}_{imt})} \qquad (17)$$

More details on the methods applied for clothing can be found in Norberg (1999).

Housing: rented apartments

A *rental survey*, taking up almost 15 per cent of the CPI weight, (including co-operative apartments and garages for which the rental index is imputed) is based on a random sample of about 1000

apartments, drawn as a subsample from a larger, annual SCB survey. A questionnaire is sent each quarter to the landlords of the sampled apartments. Re-sampling is done each year. A ratio of average rents between the current month and December of the preceding year for a fixed universe of identical units is estimated. (Before 1998 a different method was used in which the average rent was compared for samples of unmatched units.)

The rental subindex includes heating costs, hot and cold water etc., since in Sweden these are covered by the rent paid to the landlord. (Electricity for non-heating purposes is not included.) The so called housing allowance (bostadsbidrag) received by low income families is not taken into account, since it is considered to be part of income rather than a rent subsidy. Separate additions for garage and cable TV are not included in the KPI rental concept. Possible discounts, which exist in some cases (e.g. for the first months) are not considered.

The first-phase sample (drawn for the annual housing and rent survey) is a stratified sample, by region, year of construction, and floor size. The sampling design for the subsample (second phase) is an order pps sample drawn directly from the first phase sample, where the apartment rent for January last year, divided by the apartment's first phase inclusion probability, is used as the size variable. This annual subsampling applies the same methods as are used in many other areas of Swedish economic statistics. Consecutive samples are co-ordinated through the use of permanent random numbers (Ohlsson, 1995). In this way some 70–80 per cent of the old sample remains in the new sample, although the new one is a proper random sample from the updated universe of apartments.

Re-sampling is done each year, for the new index link. Within a year, an apartment is deleted from the sample if a major change has occurred, such as conversion or major renovation.

Housing: owner-occupiers

General method

In Sweden the owner-occupier housing sector consists only of single-family houses (including row houses). A user-cost approach has been

adopted. In the present version of this approach, several components are included such as *mortgage interest*, *depreciation*, *repairs*, *insurance*, *water and sewage* and *property tax*. (Heating and energy costs are added as separate items, see below.) Any capital gain (or loss) resulting from home-ownership does not enter the index. The total weight for owner-occupied housing (excluding electricity and fuel) is based on the National Accounts value for imputed rents. However, from 2001 the total weight will be based on direct estimates of different components of user costs, from a survey of owner-occupiers. At present such estimates are only used for distributing the total National Accounts weight between these components.

Mortgage interest

Most of the subindexes for owned housing are fairly straight-forward, but the calculation of the mortgage interest subindex is more complex. This subindex takes up more than half of the weight for owner-occupiers, in all around 15 per cent. It is composed of two indexes which are multiplied together: an index of average interest rates and an index of purchase values, i.e. the amounts paid for the properties, when they were last purchased.

The mortgage interest index monitors the interest, paid or foregone, on the capital invested in the home, so both borrowed and own capital are covered by the weight for this index.

The mortgage interest index is thus computed as a product of two separate indexes. The *interest rate index*, RS_{01}, measures the change of the average nominal interest rate paid for borrowed capital and the *capital stock index*, KS_{01}, measures the change of the total capital invested in homes. The mortgage interest index from period 0 to period 1 is thus defined as:

$$I_{01} = RS_{01} * KS_{01} \qquad (18)$$

KS_{01} is further defined as:

$$KS_{01} = \frac{K_S^1 + K_N^1}{K_S^0 + K_N^1 / BPI} \tag{19}$$

where K_S^0 is the total invested capital (the sum of purchase values) in all homes existing in period 0,
K_S^1 is the total invested capital in period 1 in all homes existing in period 0 (purchase values are updated according to sales between 0 and 1),
K_N^1 is the total purchasing value of homes that are new in period 1 and
BPI is obtained from SCB's construction price index for newly built single-family homes and thus "backdates" K_N^1 to the price level in the previous year.

In RS_{01}, the change in average nominal interest rate, the average is defined over all outstanding loans taken for the purpose of financing single-family homes. The average interest rate is weighted over different types of housing loans made by banks and mortgage institutions. (Different *types* refer primarily to loans with fixed rates of different lengths and to loans with flexible rates for which the rates follow the short-term interest rate.) The practical computation is done as a moving average of those fixed rates that were set in the last 60 months for new loans with 5-year fixed rates, and in the last 24 months for 2-year loans. The former loans are taken to represent loans with a longer fixed period than 5 years, and the latter those with a period shorter than 5 years. For loans with flexible rates, day to day movements are followed directly. With algebraic notation we could write:

$$RS_{01} = \frac{\sum_i w_i^{RS} \overline{R}_i^1}{\sum_i w_i^{RS} \overline{R}_i^0} \tag{20}$$

where $\overline{R}_i^0$ and $\overline{R}_i^1$ denote the average interest rate for loan type i in period 0 and 1, respectively, and w_i^{RS} is the weight for loan type i reflecting its share of the total stock of loans for owner-occupiers.

Depreciation

Depreciation refers to the physical deterioration of houses. The weight for the depreciation subindex is estimated to be 14 per cent of the total market value of all owner-occupied homes, excluding land value (as given by tax records).

From 1999 the depreciation subindex is set equal to the weighted average of the maintenance and repair subindexes (see below).

Maintenance and repair

This product group is divided into a goods and a service part. The goods part, constituting 70 per cent of the total weight, consists of some 20 commodities used in construction, including kitchen and bathroom appliances as well as construction material . Price collectors monitor prices of these in DIY stores.

The services part, constituting the other 30 per cent of the total weight, follows three wage indexes for construction workers, which are computed in SCB's bureau for construction statistics. The subweights used are based on a one-off survey of tax records carried out in 1999 related to tax deductions allowed for construction services. A small part (10 per cent) of this subindex also covers repairs of washing machines, for which normal price collection takes place.

From 2001 this product group will be restricted to materials for minor maintenance and repair. This change will not affect the calculation of the depreciation subindex, however.

Real estate tax and rent for a leasehold site

Real estate tax is a certain percentage of the assessed tax value of real estate in each year but may be changed between years. In principle the assessed values are changed continuously to reflect market values but for policy reasons these changes do not occur regularly.

The real estate tax subindex is designed to reflect changes in final tax payments. At present this is done according to the following formula:

$$I_{01} = \frac{s^1 \sum_{v \in V} k_v^1 \sum_{j \in E_v} t_j^0 \frac{o_j^1}{o_j^0}}{s^0 \sum_{v \in V} k_v^0 \sum_{j \in E_v} t_j^0} \qquad (21)$$

where s^0 and s^1 are the tax rates in period 0 and 1 respectively, k_v^0 and k_v^1 are reduction factors, applicable to newly built houses where v is the year of construction. At present these factors are either 1, ½ or 0.

t_j^0 is the base year assessed value for house j which belongs to tax category ν (which determines whether full, half or no tax shall be paid, depending on the year when the house was built),
o_j^0 and o_j^1 are politically decided annual adjustment factors that, in principle, serve to update the t_j^0 for changes in market prices,

In certain municipalities, leasehold rent is paid for the site on which the house is built. At present no index is computed for this, movements being represented by movements in the All Items KPI.

Electricity

Because of the deregulation of the electricity market in Sweden, the KPI electricity subindex was reviewed and redesigned in 1999. The present description reflects the new index, for 2000.

Following deregulation, electricity service now consists of two parts: Supply (the *power* itself) and Distribution. (Taxes and fees are added on top of both.) By law these have to be provided separately by two different companies, so the consumers purchase electricity from one company and the service of distribution from another one. However, the consumer receives only one electricity bill where the two charges are shown separately.

The consumer has a free choice of power supplier, whereas the local distribution monopoly still remains. Before November 1999, a significant cost was associated with the change of power supplier but this cost was then sharply reduced. The important changes in the market took place only after this reduction.

After deregulation, the consumer choice of power contract has two aspects:

1. *Choice of supplier power company*. So far, very few consumers have used the opportunity to change company. One reason for this is that the competition resulting from deregulation has caused power prices between companies to move quite close together. A second reason is that in some cases a small charge is still levied on consumers who change supplier.

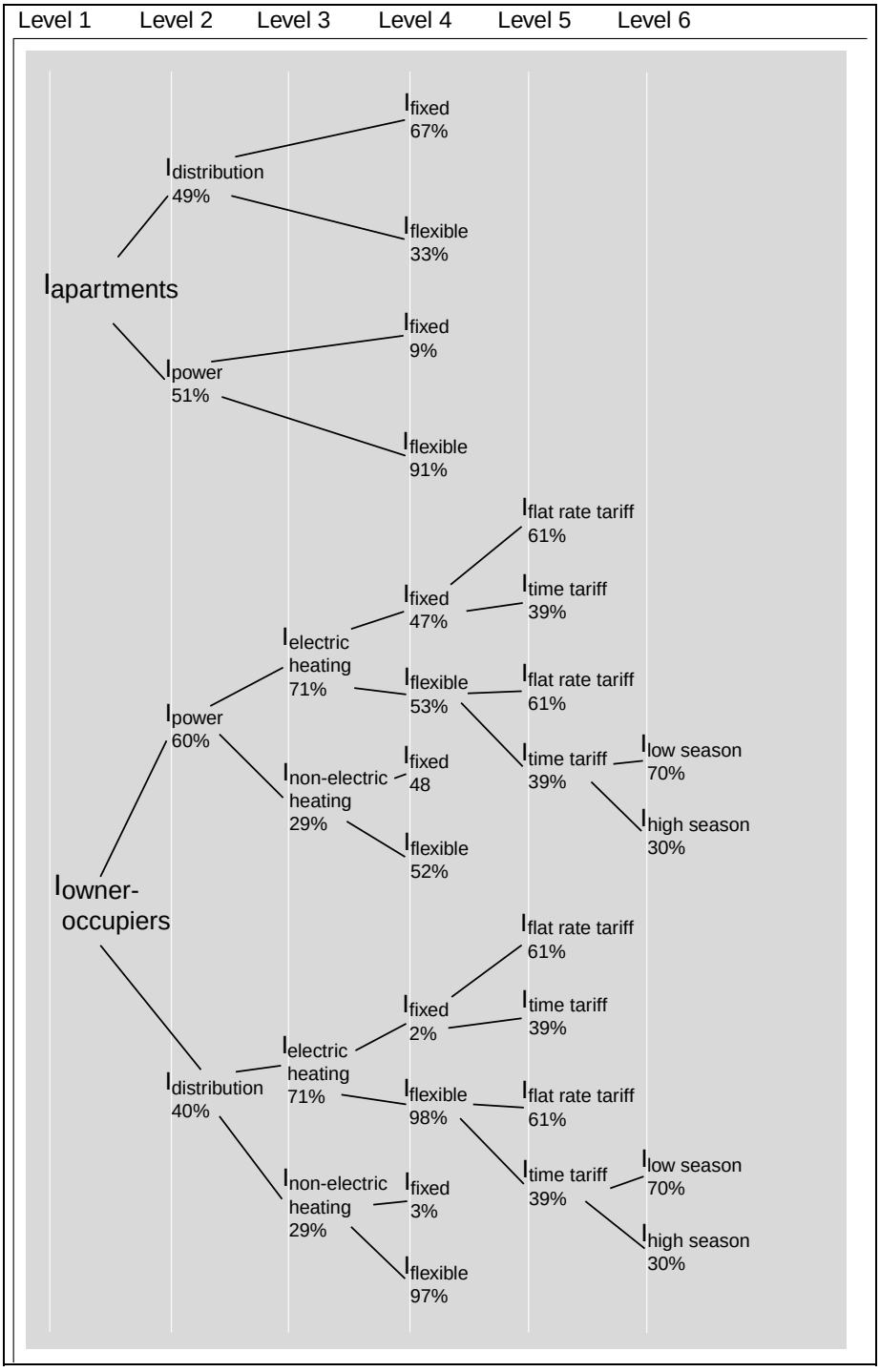
2. *Choice of contract form*. The consumer can choose to accept a power price fixed for a period of 1–3 years, to let the price follow the current price at Nordpool, the Nordic power exchange, or to do nothing, which means that the old contract and price is still in effect. So far, the choice to do nothing has resulted in a higher price.

In some companies, a third type of choice is possible, a choice which also existed before deregulation.

3. *Choice of tariff*. The consumer can choose to have the same price throughout the year and day (*flat rate tariff)* or to have a *time of day and year tariff* with high prices in peak months / hours and low prices in non-peak months / hours.

The hierarchical computation structure employed in computation of the electricity price index, including the internal weights, is shown in Diagram 1, where the successive decomposition of the electricity index into weighted subindices for up to six levels is presented. Weights generally represent the situation two years before the current year. First, apartments are separated from owner-occupied houses into two groups and the owner-occupied houses are divided into those where electricity is the main source for heating (*electric heating)* and those where it is not (*non-electric heating)*. Next the fixed part of the charge is separated from the kWh charge. Both of these are divided into two groups according to whether a *flat rate* or a *time of day and year* tariff is used. The *time of day and year* tariff subindex under the flexible price contract is finally divided into a low and a high season subindex.

Diagram 1 Hierarchichal structure of the electricity subindex



The lowest level subindexes for the flexible prices (such as $I_{high\ season}$, $I_{low\ season}$ or $I_{flat\text{-}rate\ tariff}$ in Diagram 1) involve two more aggregation steps - over companies and over different contract forms. Aggregation over companies is performed using fixed weights reflecting their sales of kWh two years ago. Aggregation over contract forms, however, follow a unit value approach in that information is collected every month on the current distribution of the customers over different contracts and an average price is calculated according to this distribution.

For Distribution, in 2000, an order pps sample, with two strata, of 50 companies from a population of 224 is used. For suppliers there is an order pps sample, with two strata, of 15 out of 153 companies.

Health, general

In Sweden, all health services are, to varying degrees, subsidised by Government. Subsidisation takes many forms. To the extent that the subsidy depends on price only, it is straight-forward to account for it in the index simply by taking account of the final price paid.

Another common form for subsidisation, however, is so called *high cost protection* by which the proportional subsidy for a person or family rises as the amount consumed rises up to some ceiling value, above which further consumption (in some cases) is free. The approach applied in the KPI in such cases to take account of this subsidy is by estimating the average unit price paid by all consumers, after deduction of the average subsidy. In some cases rather crude methods have to be used to do this.

Health, medicines

The first step towards this subindex is calculated by the Government owned *Apoteket AB* which enjoys a national monopoly for pharmaceuticals. It covers all medicines, whether prescribed or not. The calculation uses weights that change every month, in contrast with the rest of the KPI.

The Apoteket index refers to the basic, i.e. unsubsidised, prices so does not take into account any changes in the rate of subsidisation. This is instead done by the SCB, using a micro-simulation model to

analyse the effects of any subsidy changes. This model uses information on purchases of medicine made by a sample of persons in one Swedish county.

Health, dental services

From 1999, dentists have been free to set their own prices. A subsidy, which increases progressively according to a patient's total expenditures on dental services, exists for patients with a high annual total, with no cost ceiling.

The dental price survey at present covers seven different dental treatments, mainly ones with large total invoiced fees for which 37 providers provided data (27 private clinics and 10 county councils). In a first step an unsubsidised price index is calculated for each treatment. In a second step the average subsidy is deducted so that the actual price change, as experienced by consumers is obtained. Finally the total index for dental services is a weighted average over all seven treatments.

Health, care services

This subindex covers all kinds of outpatient health care, provided both by general practitioners and by specialists. Determination of the fees is generally a policy matter - only part of the price is paid by the patient and the other, larger part is paid by Government and financed through compulsory public insurance fees or taxes. The major part of health care is provided by county councils but there are also private practitioners. In principle, all types of fees and the whole country are covered by the index. The weights used are not exact consumption weights – for example, population numbers are used instead of number of consultations for weighting different county councils. In some other cases the weights are quite old.

Subsidisation takes two forms in this area. Firstly, a fixed subsidised price, which may vary between county councils is, charged for a service (e.g. a visit to the doctor), and secondly a high cost ceiling is applied to a patient's annual total cost. The first kind of subsidy is accounted for using the subsidised price itself for index calculation whereas the second subsidy is considered through a crude calculation of average prices net of subsidy.

New cars

Annually, a new sample of new car models is drawn from the Central Car Register by order pps sampling, where the size measure is the number of newly registered cars of each model. Additional specifications for the sampled models are added judgementally in order to obtain a specification that has a unique price. For each car model one outlet is selected where the asking price (i.e. not the actual negotiated price) is collected.

When model changes occur, a quality adjustment is made for changes in specification. In principle, all changes in specification are taken into account. The final decision concerning these adjustments is made by an outside panel of car experts that meets once a year. However, they put a very low value on most specification changes. The major kind of adjustment is for changes in fuel consumption. This is done according to a formula, which estimates the present value of future fuel consumption:

*quality difference= Δ litre/km * 15 000 * 3.09 *price/litre in August last year* (22)

Finally, it should be noted that only specification changes within the same basic car model are subject to this kind of direct quality adjustment. Completely new models and brands are introduced into the index in December with a new index link, effectively using the overlap method.

Used cars

The used car subindex follows the prices of 3 and 5 year old car models, since these are the ages at which used cars are typically sold by companies to households. (Sales between households are, as a matter of principle, not included in the index.) About 4*60 car models[5] are drawn annually from the Central Car Register by order pps sampling. New models are linked in and old models are linked out in the overlap month of December. Average prices of the sampled models, are purchased from Autodata (a Finnish company that keeps records of used car prices in Sweden). Autodata estimates average

[5] 60 models are 3 years old at the beginning of the index year, 60 are 3 years old at the end of the index year, 60 are 5 years old at the beginning of the index year and 60 are 5 years old at the end of the index year.

prices for these models each month according to the following regression model:

$$log(Price) = \alpha + \beta_1 t + \beta_2 m \qquad \text{where } t = \text{vintage and } m = \text{mileage} \tag{23}$$

The regression coefficients β_1 and β_2 are estimated using standard multiple regression techniques. The equation allows estimation of the price of any model for a given vintage and mileage.

Method used in 1997-1998

The prices estimated from (23) are used for index calculation according to the scheme shown in Table 3. For example the average price of a three year old car in March 2000 would be calculated as:

*Av. price = (9/12)*P[Vintage=1997, mileage=6 500] + (3/12)*P[Vintage=1998, mileage=4 500].*

In this way the KPI follows the price of a car of unchanged average age and mileage (in this case 60 000 km) every month. In the case of 5 year old cars, the vintage is two years earlier than in Table X and the mileage is 100 000 km.

Table 3: Weighted average price of a three year old used car

Index month	Pro portion	Vintage	Mileage (km*10)	Pro-portion	Vintage	Mileage (km*10)
Dec -99	12/12	1997	6 000	0/12	1998	4 000
Jan -00	11/12	1997	6 167	1/12	1998	4 167
..						
Dec -00	0/12	1997	8 000	12/12	1998	6 000

This method required matched model comparisons from one year to the next. Model specifications in the Autodata register are not such that it is always easy to see which models should be matched. For this reason this method was abandoned from 1999.

Method used 1999–2000

The prices estimated in (23) are used for index calculation, keeping vintage and mileage constant. In this way a 12-month comparison will be between cars with the same mileage but of an age that differs by about one year.

Public transport services

Intra-regional transport

Transport companies in eight county councils report their fares to the CPI. Except for Stockholm, the calculations are carried out by the counties and only a final county subindex is reported to the SCB. For Stockholm, covering almost 40 per cent of the total national weight, a calculation over all kinds of tickets is made by the SCB. County weights lag one year, dating from the year before the actual reference year.

Special methods are needed when a fare system changes in a county. In those cases it is necessary to obtain at least approximate traffic volumes from the reference year with the level of detail that is relevant under the new fare rules.

Rail fares

Up to 1999 rail transport was almost entirely monopolised by Swedish Rail. Index calculations were based on fares for a large number of distances and special eligibility conditions, detailed weights being obtained from Swedish Rail itself. From 2000 rail transport has been deregulated and more lines will be privatised. The rail fare index remains a sample of tickets for specified distances and eligibility conditions. When a new company takes over the responsibility for a certain line and distance, fares are compared directly between the new and the old company.

Air fare

Air fares are divided into two product groups - domestic and international air fares.

Domestic air fares cover the two important air companies in the market. Only flights from / to Stockholm are covered but all types of tickets and special conditions are included, with fixed weights reflecting estimated household purchases. Detailed weights are obtained from the companies themselves.

International air fares are covered by obtaining fares for a large number of distances, air companies and ticket conditions, The weights refer to private household usage, and business class and

similar fares are not included. Stratification is first by continent and country group, for each of which a number of distances are purposively selected. Detailed weights for distances and ticket types are obtained from one company only and applied to the overall calculation.

Shipping fares

This subindex covers both international and domestic shipping fares. Many international distances and shipping companies are covered. Detailed weights are of the quantity type, reflecting the total number of passengers travelling different distances. Different prices for different types of tickets are taken into account. Assumptions concerning the distribution of the passengers on different days of the week are made.

Taxi fares

A self-weighted sample of 100 taxi journeys from all over Sweden, each specified in detail, is used. The sample has been fixed since it was drawn in 1989 and later as a subsample in 1992, except that a firm which has disappeared is replaced by a new one. The price of the journey is determined by adding tariff components according to taximeter rules, fixed, pre-agreed prices, common for example for airport trips, not being considered.

Removals

An index is imputed from the construction price index. This index reflects the change in production cost for removal services by weighting a number of cost factors together, such as salaries and wages, truck prices, interest rates (reflecting capital cost), petrol and repairs. The prices paid by consumer are expected to vary around these cost changes in the long run.

Telephone services

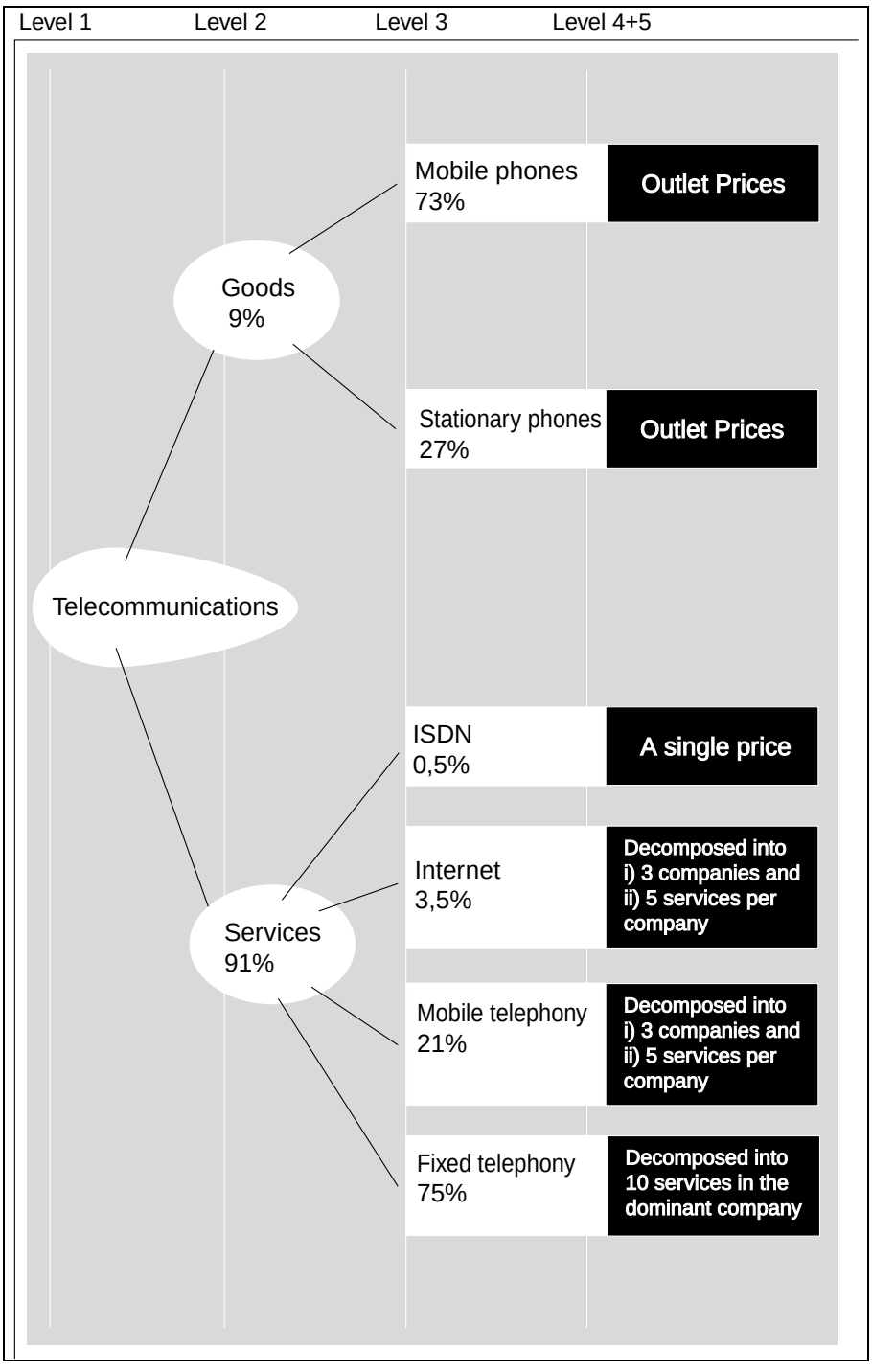
The Swedish telephone market has been deregulated since the mid 1990. This holds for all parts of the market – fixed telephony as well as mobile telephony and Internet access services. The incumbent company retains its dominant position in the fixed telephony market but not in the other markets.

The KPI telephony subindex is decomposed into a hierarchical set of subindexes according to Diagram 2, where the internal weights for 2000 are also shown. A fixed weight aggregation is applied throughout all levels of the index. Weights applied come mainly from a telecommunication survey for 1997 carried out by the SCB. Weights provided by the companies themselves are also used, where available and appropriate.

The first step divides the index into goods and services. Next, fixed line telephony, mobile telephony and Internet access are separated. For fixed telephony, the following 10 services are included:

1. Entrance fees (1 per cent). For the first installation. Two fees are included: for adults and for youth (below 26).
2. Rental (35 per cent). Refers to the fixed fee per year; there is only one price in the market.
3. Call connection (13 per cent). Refers to the start-up cost for a call; only one price in the market.
4. Domestic calls (26 per cent). From 2000, local, regional and national calls have all been charged uniformly. Two rates are weighted together – a low one for nights and weekends and a high week-day day-time rate.
5. International calls (8 per cent). An average price over all countries is monitored. Weights are by call-minutes to the respective countries. Only one company is included.
6. Calls to mobile telephones (13 per cent). Charges for calling different providers are weighted together, using call minutes in the reference period as weights.
7. Miscellaneous (4 per cent). Sub-services are: Removal of subscription to a new address, Transfer of subscription to another person, Directory enquiries, Calls to premium rate services (so called 071- and 072-numbers). In the first three cases, there is only one price in the market but for Calls to premium rate services an average price is computed.

Diagram 2 Hierarchichal structure of the telecommunications subindex



Mobile telephony is decomposed into the three existing providers with weights according to their most recently ascertained revenues. Sub-services in each company are roughly connection rate, monthly charge, weekday call rate, weekend and night rate and call connection rate. These are weighted according to the most recently ascertained revenue breakdown in each company, price updated to the price reference month.

Internet services are also decomposed into three companies, according to three year old revenues. Only the monthly access charge enters the index.

The subindices for goods in Diagram 2 are based on a traditional outlet price measurement carried out by local price collectors. Representative products are an ordinary table telephone and three specified models of mobile phones.

Changes in the above methods will be made for 2001. The details are not yet known at the time of writing.

Data processing equipment

This product group is represented by desktop PCs (65 per cent), laptop PCs (19 per cent), printers (9 per cent) and pocket calculators (7 per cent). Local price collectors are used to obtain prices for all products. A new model is entered as soon as the old one is no longer for sale.

The index calculation (except for pocket calculators) is done by monthly chaining, a calculation method that is unique to this product group. The price change from the preceding to the current month is calculated for those models that existed in both months. These monthly price changes are then chained to an index from December, last year. With the so called RA formula (above) the following index is obtained for year t, month m:

$$I_{t,m/t-1,dec} = \prod_{m'=1}^{m} \frac{\sum_{i \in \Omega_{m'} \cap \Omega_{m'-1}} \frac{p_{i,m'}}{p_{i,m'-1} + p_{i,m'}}}{\sum_{i \in \Omega_{m'} \cap \Omega_{m'-1}} \frac{p_{i,m'-1}}{p_{i,m'-1} + p_{i,m'}}} \qquad (24)$$

where Ω_m is the set of models in month m and $p_{i,m}$ is the price of model i in month m.

Informal and not yet documented experiments indicate that the results from the method applied are broadly comparable with those that would be obtained from a more sophisticated hedonic method.

Games of chance

The price of a bet is viewed as the share of the bet, in real terms, which is retained by the organisers, i.e., not paid back as prizes. The index thus consists of two multiplicative components - one reflecting changes in the percentage take of the organisers **and** one reflecting changes in the cost of maintaining the real value of a bet, measured as the KPI for all other products.

Thus if, for example, the percentage take of the organisers decreases from 51 per cent to 49 per cent and the KPI for all other products is 102, the games of chance subindex would become:

102*(49 / 51)=102*0.961=98

From 2001, all major lotteries and games of chance will be included in the KPI, so that the total coverage of the market is around 90 per cent. Important areas are horse-racing (ATG), football betting ("stryktips" and others) and "bingolotto". Different lotteries are weighted in the subindex, using the consumption value (bets minus prizes) of each as the weight.

Books

The book subindex has two main strata: *book clubs* and *book shops*. In the five book clubs included, sampled by cutoff, newly published books, mostly fiction, are included by averaging their prices in the most recent club catalogue according to detailed inclusion rules, concerning the number of pages etc.

For book shops, ten book titles (or, in some cases, book series) are included, from ten book shops all over the country. A simple geometric mean is computed of the weighted (by the turnover of the bookshops) geometric average price for each title. The book titles (of so called "indicator" books) are selected so as to be of constant quality over long periods of time.

Newspapers and magazines

This subindex is divided into four strata. Value weights are obtained from *AB Tidningsstatistik*. They are not always updated each year.

Newspapers, single copies. Nine newspapers are included in a cutoff sample. Where prices differ between days of the week, the day of the week is also included in the specification.

Newspaper subscriptions. 20 newspapers are included in a cutoff sample. Annual subscription rates are used.

Magazines, single copies. Ten magazine titles are included in a pps sample (the size measure being the total annual revenue for single copies) with a certainty stratum for the largest magazines. Very small magazines are excluded.

Magazine subscriptions. Ten magazines are included in a pps sample (the size measure being the total annual revenue for subscriptions) with a certainty stratum for the largest magazines. Very small magazines are excluded.

Package holidays

45 specified holidays from seven holiday providers, sampled by cutoff, are included in the index with equal weights.

A specified holiday is described in detail, with respect to all price-determining characteristics. These are: i) destination, ii) point of departure, iii) length of the holiday, iv) month and approximate date of departure, v) group of travellers, vi) accommodation and vii) holiday provider. This method implies that, for example, a June holiday in Spain is included in the index in December to May or June according to the listed prices applying to a booking made on the 15th of those months.

Prices are taken from travel catalogues. Last-minute prices are not included. In months where a certain holiday is not available for purchase it is excluded from the index. For December, however, prices are instead carried forward from the last month when the holiday was available for booking.

The sample of specified holidays is a *quota sample*. An example of the procedure is given in Exhibit 7.

Exhibit 7: Quota sampling of package holidays

Example: A sample of 20 package holidays is desired. It is known that 60 per cent of the holidays are to Spain, 30 per cent to Greece and 10 per cent to Portugal. 70 per cent of the travel groups are 2 adults, 20 per cent are 2 adults + 1 child and 10 per cent are 2 adults + 2 children. 20 per cent stay in 2-star hotels, 40 per cent in 3-star, 30 per cent in 4-star and 10 per cent in 5-star hotels. The sample is then selected purposively in such a way that all these proportions are

retained in the sample, which then becomes self-weighted. For example, there will be 6 holidays to Greece, 4 will be in 2-star hotels, 14 of the travel groups will be 2 adults etc.

Education

Almost all education in Sweden is funded by Government and thus free of charge from a private consumer perspective. The small private part that is included in the KPI consists of study circles, organised by special educational associations (*studieförbund*). For these, a sample of specified circles in several regional bransches of four different associations are included.

Other private education, not included in the KPI at the moment, consists only of some private boarding schools and the like which are very small in terms of total consumer expenditure.

Social protection (in the HICP, not in the KPI)

The index for social protection covers only children's day-care. Care of the elderly is included from 2001. Both types of services are subsidised by the public sector.

The index is composed of two multiplicative parts:

- An index for effects of tariff changes, which reflects changes in tariff components in a sample of 20 Swedish municipalities. The sample of municipalities is a two-stage sample where the first stage is a sample drawn for the FASIT micro-simulation model, used for calculating income statistics. The second stage sample is a pps sample from the first stage sample. From each of the 20 municipalities a number of tariff components are collected (e.g. price per child for full-time care for a family of two children in a certain income bracket), and the tariff index is obtained by the FASIT model.
- An index for effects of income changes, estimated from official income statistics with a 3-month lag. This index is also obtained by the FASIT model, based on a weighted average of data for blue-collar and white-collar workers.

The subindex for children's day-care is also used as the index estimate for fees for children's after-school centres (Coicop 10.X).

Insurance

Four categories of insurance are included in the KPI. The weight for each of these categories reflects the service charges applied by the insurance company, defined as premiums received plus premium supplements minus claims paid out. (Up to 1999 KPI weights were calculated according to the gross principle, reflecting gross premiums without subtracting claims.)

For all kinds of insurance, the basic product-offer included and used as a price indicator is an insurance policy of constant coverage and quality in real terms. In each sampled insurance company, policies are specified with respect to all their price-determining characteristics, which are then held constant for month-to-month pricing irrespective of any changes in the policies of the sampled insurance holders.

Home insurance. Home insurance covers the contents of a home and often includes some other things like legal protection and travel insurance. In the case of a single-family house, this insurance is normally combined with a building insurance (see below) but the two parts are separable and separated due to HICP requirements.

A two stage sample of 12 companies in the first stage and 52 insurance policies in the second stage was drawn in 1992 and is still used today. Probability sampling was applied in both stages - in the second stage the sample was drawn by statisticians in the insurance companies themselves. The policies are self-weighted. This is achieved by making the sample size in each company proportional to its premium sum.

The index is calculated according to an unweighted RA formula.

Building insurance. Building insurance covers damages to the building and its site. The same sample and calculation formula as for home insurance is used. In practice, only the part of the sample that refers to home-owners is used, in this way an effective sample of 28 policies is obtained.

Car insurance. Car insurance covers both damages and theft to the insurer's car and damages to third parties, caused by accidents.

Cars are drawn from the central car register as a simple random sample, at present consisting of 48 cars drawn in 1998. For those cars there was information on the company with which the car was insured and the price determining characteristics of each insurance policy was obtained from those companies and kept fixed thereafter. Month to month pricing is accomplished using information from the same companies. The index is calculated as a ratio of arithmetic mean prices (A) which provides an estimate of a Laspeyres type index.

Reliability and uncertainty

The reliability of price change estimates is influenced by a multitude of factors. A special investigation on these problems was undertaken by the KPI Commission (1999). Below we give a summary of some points made in this report.

Conceptual problems

Reliability must be related to the chosen measurement concept, in its most general form, being defined as the difference between the estimate actually obtained and the one that would have been obtained with full and perfect data relating to the ideal concept. However, this way of looking at reliability leads to complications for CPIs like the Swedish KPI, whose ultimate aim it is to estimate changes in the cost of maintaining a constant standard of living according to the theoretical notion of the true cost of living index. This is because it is an abstract concept that can not be exactly expressed as a mathematical function of observable quantities.

In practice, the two most important areas of conceptual difficulties in CPIs are:

i) *Owner occupied housing*. Although there is some (but not full) agreement among economists that a user cost model is relevant for determining the target price index for living in one's own house, there is much disagreement about the precise definition of this user cost model. In fact, no two countries use the same detailed model. Important differences exist as to the treatment of interest rates, taxes and other aspects of the index. In (Dalén, 1999) it was shown that the differences between the results for 12-month changes of the All Items CPI based on several different plausible detailed concepts for owner-occupied housing was between –1.4 and +0.5 index units for individual years and -0.3 to -0.6 on average over a 20-year period. These differences could not properly be labelled "biases", since no general agreement on the correct answer exists, even retrospectively.

ii) *Quality change and new products.* Although economic theory provides some guidance as to the valuation implied by consumer

> choices of changes in the set of consumer products and their characteristics, it is not possible to determine the measurement target with exactness, even in retrospect. The arguments this clearly. In the light of the varying stands taken by persons this aspect can be said to be of a similar order as that of owner-occupied housing, that is at least 0.5 index units.

One way of expressing this is to say that there is a *conceptual grey zone*, of the order of one percentage unit, within which we cannot state with any assurance how well our measurement results reflect what should be understood by, for example, the word *inflation*.

Bias

Despite these conceptual difficulties, a bias assessment for Sweden was put forward in Dalén (1999). This assessment was guided by the concept of the cost of living index. It looked at the bias components focused by the Boskin Commission in the U.S but also took in some other aspects of the index peculiar to Sweden.

Exhibit 8: Bias components according to the Boskin Commission

Essentially all ”economic” bias components are due to the fact that a CPI follows the price of a fixed set of products with fixed weights possibly purchased in a fixed set of outlets. This is in contrast to the real world where products and outlets are continually changing and where consumers are able to take advantage of these changes for increasing their standard of living by substituting old products or outlets for new ones. More specifically the Boskin Commission considered the following components of bias:

- *High level substitution bias.* This refers to consumer substitution between broad product groups such as food for clothing or meat for fish as a result of different price changes for these groups.
- *Low level substitution bias.* This refers to consumer substitution between specific brands of the same product or between different outlets selling the same product as a result of different price changes for these brands / outlets.
- *New products and quality change.* This refers to the effects on the standard of living of completely new consumer products such as the video recorder, mobile phone or a new medicine compared to an earlier situation where they did not exist. It also refers to the effects of quality changes of new models incorporating another set of characteristics compared to the old ones. Examples are new PC models with more internal memory, hard disk space etc.
- *New outlets.* This refers to the effect on the average price level of completely new outlets entering the market such as discount outlets, factory outlets, Internet shopping etc.

The assessment for Sweden was made in two parts – for the non-housing components (2 / 3 of the weight) and the housing components (1 / 3), respectively.

Non-housing components

The first observation is that high-level substitution bias is close to zero in Sweden. This is due to the unique construction of the KPI with a long-term index that uses current period weights (see formula 5 above). Since 1990, an elementary aggregate formula which is comparable to the geometric mean has also been used for most product groups which makes low-level substitution bias very small.

For quality change and new products in general, the Boskin approach was taken as the starting point for assessments by product group, but in a number of cases particular circumstances in the Swedish KPI justified modifying the estimates. For example, one key issue is how the health component, which contributes much to the U.S. bias, should be looked at. In a government funded system such as that in Sweden, much of the quality improvements affects government consumption rather than household consumption. Taking this view, bias becomes smaller, although the bias for medicines is still considerable. The non-housing quality and new product bias was estimated to contribute to an upward KPI bias of 0.33, distributed as shown in Table 4.

Table 4: Assessment of bias resulting from new products and quality change Housing not included.

Product group	Bias	Weight	Contribution to total bias
Food (incl. away from home)	0.31	20	0.062
Household appliances	3	1.3	0.039
Furniture + telephone	0.5	6	0.030
Clothing and shoes	0.5	5.5	0.028
Transport, cars	2.1	3	0.063
Transport, other	0	13	0.000
Health	1.2	2.8	0.032
Entertainment	0.67	8.4	0.056
Other products	0.28	6.5	0.018
Sum		**66.5**	**0.328**

For estimating new outlet bias, data on market shares for low-price outlets from 1986 up to today were used. Since these market shares were increasing, it was concluded that there was an overestimation bias of 0.03-0.05[6].

There are also sources of bias of a different nature than those Boskin considered. Some product groups are not covered by the index. Two

[6] All numeric estimates refer to a 12-month change and are expressed in percentage points. For example an error of 0.03 is to be added/subtracted to an index number like 102.5.

such groups are childcare and care of the elderly with a combined weight of 0.9 to 1.5 during the years 1980-1996. National Accounts consumption data in current and constant prices, using volume data like number of children enrolled etc., show implicit price increases above the KPI average for 1980-1996 which implies an underestimating bias of 0.03 on average and more than that in the 90's. There are other cases of under-coverage both in the product and in the outlet dimensions of the KPI. Examples are trade union fees, objects of art, hospital care, some seasonal goods, mail order, market stalls, and inter-regional bus fares. Although there is clearly a risk of bias resulting from these exclusions, the direction of this bias is unknown.

The failure to measure the price actually paid is also a source of bias. This failure may be caused by the use of list prices (e.g. for cars) and by the presence of discounts, coupons, "club prices", bargaining etc., which are not accounted for. These phenomena have probably become somewhat more common during the last 10-20 years, causing a small estimated upward bias of 0.02 or so.

Sampling bias due to non-probability sampling is usually neutral in the long run. However, there is one case where there is a clear danger of an underestimation bias. This is due to the tendency of price collectors to choose an excessive proportion of regularly (non-sale) priced variants in the price reference period, whereas the proportion of sales prices in the prices collected becomes proportional to their real share later in the year. In the KPI, this error is corrected for with respect to clothing but it may exist for other products. Should these other products belong to the high-tech sector, this factor would somewhat offset the upward bias for such products. This has been allowed for in the estimate of quality change bias presented above.

Housing

Housing consumption in Sweden comprises four distinct parts: i) Rented dwellings, ii) Dwellings in housing cooperative building societies, iii) Owned one-family houses and iv) Owned secondary homes.

For rented dwellings, there are many possible sources of bias, but in both directions. The main underestimation risks are: i) Failure to account for depreciation and ii) New, more expensive apartments are

linked in without comparison backward (note that the Swedish rental market is regulated). The main over-estimation risk is the failure to fully account for many small quality improvements in apartments; for example, cable TV is given a very low value. The net effect of these three sources of bias is judged to be more likely to be an under- than an overestimation, with a most likely effect of –0.04.

Owner-occupied one-family housing accounts for about 15 per cent of the KPI weight and its treatment thus influences the index outcome a great deal. At the same time the measurement target is disputed – there is *conceptual uncertainty*. The method used today is a kind of user-cost method where mortgage interest plays a dominant part. A sensitivity analysis was therefore done to look at other possible methods. These were: i) Rental equivalence, where the rent index was imputed to represent a large part of the weight, ii) Direct consumption where the purchased house is considered as immediately consumed, like other durables in the CPI, iii) Exclusion, where owner-occupied housing is simply left out of the index and iv) The final proposal in the Commission Report, namely to treat the real interest rate as constant and use real estate prices as the price indicator. All these alternatives gave higher KPI outcomes for most years from 1981 to 1997 indicating that there would be an underestimation bias if one of the alternative concepts were accepted as right in principle. The range of the bias was from –1.41 to +0.46. On average, it was from –0.57 (rental equivalence) to –0.26 (exclusion). The final estimate of bias was taken to be –0.30, which reflected the difference compared with the final Commission Proposal.

Consumption of secondary homes (12 per cent of household consumption) is excluded from the KPI today. It is represented in equal amounts by rents and owner-occupied housing. No obvious over-or underestimation bias results from this approach.

Summarised bias assessment

The summarised bias assessment, shown in Table 5, is very close to zero for the long-term index. It is estimated to be 0.13 for the short-term index which determines the 12-month inflation rates that are published. For the Swedish HICP, which excludes owner-occupied housing, the above assessment would result in an estimate of around 0.5. However, here it must be noted that the HICP is not a cost-of-

living index by intent so that the relevance of such an assessment may be questioned.

Tabell 5: Summary of bias assessments

Source of bias	Size of bias
Substitution bias, high level	LTIX: -0.04; STIX +0.08
Substitution bias, low level	0
New outlets	+0.04
New products and quality change	+0.33
Under-coverage, public services	(-0.04)[7]
Price measurement errors	+0.02
Owner occupied housing	-0.30
Rents	-0.04
Sum	**LTIX:+ 0.01 (-0.03)[3] STIX: +0.13 (0.09)[3]**

The estimates refer to averages over several years. For a single year they could be markedly different. The uncertainty of the assessment itself, which is mainly due to the conceptual problems mentioned above, should be emphasised.

Random sampling errors

The KPI uses a primary stratification into item groups which in turn are combined into separate and independent price surveys. The variance of the whole price index is therefore a weighted sum of the variances of the separate surveys:

$$V(I) = \sum_h w_h^2 V(I_h) \qquad (25)$$

A reason for reasonably assuming that all these surveys are independent (but not the product groups themselves) is that there is no common regional sampling scheme used in them.

Altogether, there are about 60 different surveys. Some of them cover only one product group and have simple designs. Those which cover their universes completely may be assumed to have zero variance.

[7] Numbers in brackets refer to historical estimates that are not considered relevant for a bias assessment for the future.

Some price surveys are more complex, however. This is especially the case for that large part of the index where outlets and products are simultaneously sampled. Here, outlets are sampled by probability (pps) from the Central Business Register. Products are sampled by pps for daily necessities but by the representative product method for other local price collection items. The final sample in these cases is considered as drawn from a two-dimensional universe of products and outlets. The final sampling units are thus sampled products sold in sampled outlets – *a cross-classified sample.*

For a cross-classified sample the total variance can be decomposed into three parts:

- *Variance between products* (in the same outlet)
- *Variance between outlets* (for the same product) and
- *Outlet and product interaction variance*

The cross-classified model is fairly close to the actual sampling design for local price collection, although some products are in fact purposively drawn.

Systematic computations according to this model were carried out for a few years in the mid 90's. Total variance was estimated at 0.04, corresponding to a 95 per cent confidence interval of ±0.4. This estimate appeared to be fairly stable for the years (1991-1995) for which the model was tried. This estimate primarily refers to a price change from December one year to December the following year, that is to a 12-month change in the same index link. However, there are reasons to believe that roughly the same estimate applies to 12-month changes involving two consecutive links. The sampling error of index estimates at lower aggregation levels are generally higher, due to smaller underlying sampling sizes. The clothing subindex has a particularly large sampling error (of several percentage points), due to the large variations in price and price change within this group.

Dalén and Ohlsson (1995) provide exact formulae and more details of the procedures used.

Mistakes and the shadow index

In addition to the statistical uncertainty discussed above, mistakes are sometimes committed, which are discovered after publication of the results. However, following each monthly publication, the All Item KPI (index with 1980=100) is legally fixed; it cannot be revised. These fixed numbers are used for legal and contractual index-linking. On the other hand, no subindices are fixed in this way and they therefore can be, and occasionally are, revised when necessary.

Mistakes are of two kinds: i) mistakes within the framework of the chosen method and ii) mistakes affecting the choice of method. When such mistakes are discovered, a "Shadow index" is computed, which uses the corrected indices consistently backwards for the whole series of All Items indices and subindices. This ***shadow index*** is used for computing all rates of change, which are therefore not always consistent with the official fixed index numbers. There have been six such mistakes since 1980.

- January 1983 – mistakes in the price data for electricity and distant heating.
- January 1987 – the abolition of a certain tax on housing was not noticed until the index for February.

- January-March 1990 – An unfortunate choice of elementary aggregate formula.

- January 1992 –July 1993 – A selection bias in the index for clothing, which was not eliminated until the index for April 2000, lowered the total index level approximately 0.1 per cent.

- January 1995–March 1997 – an inappropriate method was used for owner-occupiers mortgage interest costs. The effect on the All Items KPI level was eliminated as from the index for January 1998.

- July–September 1997 – an erroneously entered rent quotation was overlooked, which affected the index for three months.

In Sweden's Statistical Databases as well as on Statistic Sweden's web-site, the revised shadow indices for the All Items KPI and for product categories are presented in a consistent manner. Publication in printed reports are of the same kind.

Satellite indexes

The Harmonised index of consumer prices (HICP)

The HICP is an index intended for making international comparisons of inflation. HICP numbers exist from 1995 onwards, with 1996=100 and are comparable between different countries in Europe.

The HICP exists for all EU countries plus Iceland and Norway and, aggregated, for the EU and the European Monetary Union (EMU) as a whole. Recent HICP indexes are regularly published in press releases from Eurostat, the EU Statistical Office, and are available on the website: europa.eu.int/comm/eurostat/

The HICP serves as a target variable for monetary policy in the EMU and is a fundamental economic indicator for the European Central Bank (ECB). Having an internationally comparable measure of inflation first became an immediate concern for assessing compliance with the so called convergence criteria for membership in the EMU. In 1993, Eurostat initiated a co-operative effort with the EU's National Statistical Institutes, which led to the HICP.

The content and methods of the HICP are agreed on by the Member States of the EU and are laid down in a number of regulations and guidelines shown in Exhibit X. At present these regulations deal with coverage, new products, updating of product samples, adjustment for quality change and index formulae, although some rules are not very far-reaching. These requirements have also led to the inclusion in the KPI of certain goods and services not covered earlier, such as parking fees, bank services and used cars. For the decomposition of the HICP (and from 2000 also the KPI) into product groups, the international standard classification known as COICOP (Classification of Individual Consumption by Purpose) is used.

The coverage of the KPI and the HICP in terms of product groups differ to some extent. In Annex 1, the detailed coverage can be found from the weights column, where a zero weight indicates that the product group in question is excluded.

In the years past new regulations and guidelines have continually been added. The working process for agreeing on these regulations is

1. Discussions in a Working Party (WP) for the HICP. In the WP, the CPI / HICP producers from National Statistical Institutes in all 15 EU Member States are represented as are price statisticians from Candidate Countries from Eastern Europe. Some users, primarily from central banks (including the ECB) also attend. Often, more complicated issues are first examined in special Task Forces, composed of members from volunteer countries. The WP makes recommendations for the adoption of regulations.

2. Decisions by the Statistical Program Committee (SPC), composed of Directors General of NSIs.

3. Council or Commission decisions according to general EU procedures.

Like the KPI, the Swedish HICP is a chain index with annual links but the links are all based on the short-term index so that there is no long-term index involved in the HICP calculation.

The HICP does not replace the KPI. The KPI will remain the official measure of changes in consumer prices in Sweden in the future.

The Net Price Index

The NPI is defined as a CPI in which both reference year and current year consumer prices have been reduced by the sum of all indirect taxes which, at different stages of production, enter into the cost of production of final consumer products. The prices are also adjusted for certain subsidies. The index is intended to follow the movement of sales revenues that part of the private sector producing consumer goods and services would receive after deduction of indirect taxes and after addition of subsidies. These sales are defined as referring to an unchanged amount and composition of goods and services, produced with a constant technology.

In the NPI, indirect taxes are subtracted from the consumer prices as recorded in the KPI. Not only indirect taxes on the direct sale of consumer products are removed but also indirect taxes that are levied on the raw materials, semi-finished goods, input materials and labour

that are estimated to have contributed indirectly to output of the products. Also, certain cost elements of an indirect tax nature that are related to the use of capital goods are removed. Whether an indirect tax is considered as fully or only partly shifted to consumers is not taken into account, the full value of indirect taxes always being removed from the consumer prices. Subsidies applying to consumer products are treated symmetrically, additions to the consumer prices being made.

The NPI is computed every month, but only at the All Products level. For reasons related to the method of calculation used, the NPI cannot be decomposed into subindexes.

The boundary drawn between indirect taxes and other Government revenues and between subsidies and other Government expenditures is based on National Accounting principles, albeit with certain exceptions.

In principle the NPI can be defined as

$$NPI = \frac{\sum_{i \in X} (p_i^1 - s_i^1) q_i}{\sum_{i \in X} (p_i^0 - s_i^0) q_i} \qquad (26)$$

where X is the domain of products defined by KPI and s_i^0 and s_i^1 are the amounts of taxes / subsidies per unit of product i in period 0 and 1 respectively.

Another way of expressing the index, which comes nearer to its actual computation, is given by the following formula

$$NPI = \frac{KPI - \sum_j w_j^T I_j^T}{1 - \sum_j w_j^T} \qquad (27)$$

where w_j^T is the share of a certain type of tax / subsidy j in total private consumption and I_j^T is an index showing the change in tax / subsidy j since the base period.

The following indirect taxes and subsidies were allowed for in the calculation of the NPI for 1999.

Indirect taxes: Value added tax, beverages tax, tobacco tax, energy tax, carbon dioxide tax, sulphur tax, nuclear power tax, different electricity taxes, vehicle tax, vehicle sales tax, custom duties, advertising tax, gambling tax, fertiliser tax, pesticide tax, waste tax, crude pine oil tax. natural gravel tax, real estate tax, labour protection tax, general payroll tax, special pensions tax and road user charge for lorries.

Subsidies: Operational subsidy to public transport companies, Central Government newspaper support, Central Government literature support, support to sparsely-populated areas, regional employment and transport subsidy, agricultural subsidies, mortgage interest subsidies for dwellings and subsidies to study circles.

The Net Price Index (NPI) was first published in May 1965. Between 1959 and 1978 index numbers were published for February, May, August and November. NPI is currently presented on an index reference base of 1980=100.

Indicators of underlying inflation

Starting with the publication of the Swedish Consumer Price Index (KPI) for August 1998, SCB has calculated and published two measures of underlying inflation. These measures, called UND1X and UNDINHX, are defined by the central bank of Sweden and published monthly on its behalf . Their purpose is to provide alternative measures of inflation that exclude temporary effects that enter into the KPI, these normally being unimportant for monetary policy.

In both of the two measures of underlying inflation, the KPI domain is modified so as to exclude mortgage interest on owner-occupied houses. In UNDINHX the domain is further reduced to exclude goods that are mainly imported. Prices are adjusted for changes in indirect taxes and subsidies. Thus alterations in indirect taxes and subsidies, except those related to wages and salaries, are excluded from price movements. (These alterations are increases or reductions in the tax/subsidy rates, which result in a change in the amount of

indirect taxes / subsidies paid, expressed as a proportion of the price of a single good or service.) Taxes and subsidies on final consumption are deducted, while taxes on raw materials and semi-manufactured goods are not.

In principle both indexes could be expressed as

$$UND_X = \frac{\sum_{i \in X} (p_i^1 - \Delta s_i) q_i}{\sum_{i \in X} p_i^0 q_i} \qquad (28)$$

where X is the domain of products defined by UND1X and UNDINHX, respectively and

Δs_i is the change between periods 0 and 1 in the amount of indirect taxes and subsidies per unit of product i which is due to increases or reductions in tax / subsidy rates (the difference between the actual amount of indirect taxes / subsidies paid in period 1 and the amount that would have been paid if the tax / subsidy rates had remained unchanged).

Another way of expressing the index, which comes nearer to the way in which it is actually computed, is:

$$UND_X = KPI_X - \sum_j w_j^T (I_j^T - I_j^{T*}) \qquad (29)$$

where KPI_X is total KPI with exclusions defined by the domain X,

w_j^T is the share of a certain type of tax / subsidy j in total private consumption,

I_j^T is an index showing the actual change in tax / subsidy j since the base period and

I_j^{T*} is an index showing what the change in tax / subsidy j would have been given an unchanged tax / subsidy rate.

Depending on the nature of the particular tax involved (e.g. whether it is specific or ad valorem), the exact computation takes a number of different forms. More details are given in Johansson and Sjögren (1999).

The Price Basic Amount

National basic pensions and certain supplementary pensions, certain life annuities, study assistance and certain maintenance allowances are guaranteed a given real value by means of adjustments made using a Price Basic Amount (previously Basic Amount). As from 1982 this has been calculated for a whole year.

For example, for 1999 the Basic Amount amounts to SEK 36 396 multiplied by a factor amounting to the change in the consumer prices between June 1998 and June 1997. The Basic Amount is rounded to the nearest hundred crowns. A detailed account of the rules for calculating it is provided in the KPI Annual report (PR 15).

The rules governing the changes in the Basic Amount, although defined in relation to the KPI, are ultimately a matter of governmental policy. From time to time they are set by Parliament and Government. For example, inflationary effects of the Swedish depreciation in November 1992 up to October 1994 were excluded from the Basic Amount. By providing a distinctly separate measure for this policy purpose the statistical integrity of the KPI itself is maintained.

References

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

Dalén, J. (1999): *Bedömning av biasrisker i konsumentprisindex (KPI)*. In KPI Commission (1999, below). In Swedish

Dalén, J. and Ohlsson, E. (1995): *Variance Estimation in the Swedish Consumer Price Index*. Journal of Business and Economic Statistics, Vol. 13, No. 3, 347–356.

Diewert, W.E. (1976): *Exact and Superlative Index Numbers*. Journal of Econometrics, Vol.4, 115–145; reprinted in Essays in Index Number Theory, Vol. 1 (eds. W.E. Diewert and A.O. Nakamura), Amsterdam: North-Holland, 1993, pp. 223–252.

Johansson, J. and Sjögren, I. (1999): *UND1X and UNDINHX: Implementation of Two Measures of Underlying Inflation by Statistics Sweden*. In Proceedings of the Measurement of Inflation Conference, edited by M. Silver and D. Fenwick. Cardiff University, pp. 336–344.

Konüs, A.A. (1924): *The Problem of the True Index of the Cost of Living*, translated in Econometrica, Vol. 7, (1939), 10–29.

KPI Commission (1999): *Konsumentprisindex, betänkande från utredningen om översyn av konsumentprisindex*. Justitiedepartementet, Stockholm. SOU 1999:124. In Swedish, summary in English.

Norberg, A. (1999): *Quality Adjustment – the Case of Clothing*. In Proceedings of the Measurement of Inflation Conference, edited by M. Silver and D. Fenwick. Cardiff University, pp. 410–426.

Ohlsson, E. (1995): *Coordination of Samples Using Permanent Random Numbers*. In Cox et al: Business Survey Methods, Wiley, pp. 153–169.

Rosén, B. (1997a): *Asymptotic Theory for Order Sampling*. Journal of Statistical Planning and Inference, Vol. 62, 135–158.

Rosén, B. (1997b): *On Sampling with Probability Proportional to Size*. Journal of Statistical Planning and Inference, Vol. 62, 159–191.

Socialstyrelsen (1961). *Konsumentpriser och indexberäkningar åren 1931-1959*. Sveriges Officiella Statistik, Stockholm. In Swedish

SOU (1943): *Betänkande angående levnadskostnadsindex*, SOU 1943:8, Finansdepartementet, Stockholm. In Swedish.

SOU (1953): Konsumentprisindex, betänkande angående omläggning av levnadskostnadsindex, SOU 1953:23, Civildepartementet, Stockholm. In Swedish.