Distr.

GENERAL

CES/AC.49/1999/10
14 September 1999

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

INTERNATIONAL LABOUR ORGANIZATION (ILO)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/ILO Meeting on Consumer Price Indices

(Geneva, 3-5 November 1999)

SAMPLING AND NON-SAMPLING ERRORS IN A CONSUMER PRICE INDEX

Invited paper submitted by Statistics Netherlands[1]*

Summary

To control the accuracy of a Consumer Price Index (CPI) the magnitude of all possible errors has to be assessed. In this paper the different types of errors that occur during the process of compiling a CPI are described. Further, some methods that are used in practice to measure and minimise the errors in a CPI are discussed.

Keywords: price index, accuracy, measurement, control.

Introduction

1. In many countries a Consumer Price Index (CPI) is used as an indicator of inflation. A CPI plays also an important role in the indexation of wages, social security benefits and taxes. Bias and errors in a CPI can have enormous implications for the whole community of a country. Bias can be defined as any systematic deviation with respect to some ideal population index. In discussing bias most people do so by considering how well a CPI is able to measure changes in the aggregate cost of living. The cost-of-living index, that is a theoretical concept based on unobservable variables, has to be translated into an operational goal or object of estimation. Most statistical offices adopt a fixed-weight price index as the operational goal. The difference between a cost-of-living index and a fixed-weight index is known as substitution bias. This bias shows up at different aggregation

[1] Paper prepared by Mr. Martin Boon, Statistics Netherlands. The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

levels: commodity substitution bias, outlet substitution bias and elementary substitution bias. The fixed-weight index is estimated from prices of a sample of commodities observed in a sample of outlets, and expenditure shares of commodity groups observed in a sample of households. In the estimation process of the operational goal different errors in both price indices and weights can occur, ranging from sampling to non-sampling errors. In section 2 the different kinds of errors in an estimated CPI are briefly described. Section 3 discusses the way errors such as sampling variances can be measured (or estimated). Section 4 deals with some procedures to minimise errors in an estimated CPI. Section 5 contains some final remarks.

1. Description of types of error

2. One of the main objectives of a sample survey is to compute estimates of population characteristics. Such estimates will never be exactly equal to the population characteristics. There will always be some error. This error can have many causes. Figure 1 gives a taxonomy of possible causes. See also Balk and Kersten (1986) and Dalén (1995) for an overview of the various sources of stochastics and errors experienced in calculating CPIs.

Total error:
- Sampling error
  - Selection error
  - Estimation error
- Non-sampling error
  - Observation error
    - Overcoverage
    - Measurement error
    - Processing error
  - Non-observation error
    - Undercoverage
    - Non-response

Figure 1. A taxonomy of errors in a consumer price index

2. Two broad categories can be distinguished: sampling errors and non-sampling errors. *Sampling errors* are due to the fact that an estimated CPI is based on samples and not on a complete enumeration of the populations. Sampling errors vanish if one would observe the complete population. As already mentioned, statistical offices usually adopt a fixed-weight price index as object of estimation. The fixed-weight index can be seen as a weighted average of partial indices of commodity groups with weights being expenditure shares. The estimation procedures most statistical offices apply to a CPI involve different kinds of samples. The most important are:

- for each commodity group a sample of commodities to calculate the partial price index of a commodity group;
- for each commodity a sample of outlets to calculate the elementary price index of a commodity from individual price observations;

- a sample of households needed for the estimation of the average expenditure shares of the commodity groups[2].

3. The sampling error can be split in a selection error and an estimation error. A *selection error* occurs when the wrong selection probabilities are used. The probability of drawing an element out of the sampling frame has to coincide with the selection probability that is specified in the sample design. The *estimation error* denotes the effect caused by using a sample based on a random selection procedure. Every new selection of a sample will result in different elements, and thus in a different value of the estimator.

4. *Non-sampling errors* may even occur if the whole population is investigated. They denote errors made during the process of recording the answers to the questions. Non-sampling errors can be subdivided in observation errors and non-observation errors. *Observation errors* are one form of non-sampling errors. These denote errors made during the process of obtaining and recording the answers.

5. *Overcoverage* means that elements are included in the survey not belonging to the target population. For outlets, statistical offices have usually inadequate sampling frames. At Statistics Netherlands, for instance, the statistical business register (SBR) is used as sampling frame for outlets. In this register outlets are classified according to major activity. A business register usually has extensive overcoverage, because it contains out-of-scope outlets (for instance firms that sell their services to businesses rather than to households). In addition, there is usually no detailed information on the commodities sold by an outlet. So, it is possible that a sampled outlet may turn out not to sell the commodity in question.

6. *Measurement errors* in a household expenditure survey and/or price survey occur when the respondent does not understand the question, or does not want to give the right answer, or when the interviewer and/or price collector makes an error in recording the answer. In household expenditure surveys, households appear to systematically underreport expenditures on commodity groups such as tobacco and alcoholic beverages. Price data are usually collected by field collection, mail questionnaires, telephone and/or electronic data capture (by means of CD-ROM, diskette or via the Internet). Until now, in most countries the main collection method is field collection by price collectors that regularly visit outlets. Price collectors can make mistakes in the outlets by collecting prices of wrong commodities.

7. The collected price data are processed in different stages such as data coding, entry, transfer and editing (i.e. control and correction). At each step mistakes, so-called *processing errors*, may occur. In the Netherlands, price collectors write down prices on paper forms at the outlets. After they have returned home from collecting prices, they use a laptop computer as input and transmission medium for the price information on the collection forms. This way of processing prices is susceptible to errors.

---

[2] Some countries use data from National Accounts instead of a household expenditure survey to obtain the expenditure shares.

8. *Non-observation errors* are errors made because the intended measurements could not be carried out. *Undercoverage* occurs when elements in the target population do not appear in the sampling frame. The sampling frame of outlets can have undercoverage, which means that some outlets where commodities are purchased can not be contacted. Most statistical offices exclude mail order firms and non-food market stalls from outlet sample selection.

9. Another non-observation error is *non-response*. Non-response errors may arise from the failure to obtain the required information timely from all the units selected in the sample. One can distinguish total and partial (or item) non-response. Total non-response occurs when selected outlets cannot be contacted or refuse to participate in the price survey. It is also possible that mail questionnaires and collection forms are returned by respectively the respondent and the collector of the price survey after the deadline for processing. Mail questionnaires and collection forms that are only partially filled in by respectively the respondent and the collector of the price survey can be considered as partial non-response. If the price changes of the non-responding outlets differ from those of the responding outlets, the results of the price survey will be biased.

10. In a household expenditure survey one also encounters total and partial non-response. Households drawn in the sample can refuse co-operation. Partial non-response occurs if certain types of households refuse to give information about their expenditure on certain commodity groups.

2. Measuring errors

2.1 Estimation of variance

11. The variance estimator depends on both the chosen estimator of a CPI and the sampling design. In Boon (1998) an overview has been given of the sampling methods that are applied in the compilation of CPIs by European statistical institutes. It appears that only four EU national statistical institutes use some form of probability techniques for outlet selection and only Statistics Sweden uses probability sampling for item selection. In the absence of probability techniques, so-called judgement and cut-off selection methods are applied.

12. In view of the complexity of the (partially connected) sample designs in compiling a CPI, an integral approach to variance estimating seems to be problematic. That is, it seems to be difficult, to present in the form of but one formula a measure of the variance of a CPI in which all sources of sampling errors are caught. Usually partial (or conditional) measures are developed, in which only the effect of a single source of variability is quantified. For instance, Balk and Kersten (1986) have calculated the variance of a CPI due to the sampling variability of household expenditure surveys. These variances are conditional on the assumption that the partial price indices are known with certainty. The conditional sampling errors have to be put together into an unifying framework. With such a framework we can assess the relative importance of the various sources of sampling error.

Under rather restrictive assumptions Balk (1989) has derived an integrating framework for the sampling error of a CPI.

13. There are various types of procedures for estimating CPI sampling variances. Design-based variance estimators (i.e. variances of Horvitz-Thompson estimators) can be used in combination with Taylor linearisation procedures for sampling errors arising from a probability sampling design. For instance, assuming a cross-classified sampling design, in which samples of commodities and outlets are drawn independently from a two-dimensional population, with probability proportional-to-size sampling (PPS) in both dimensions a design-based variance formula can be derived (Dalén and Ohlsson, 1995). Dalén and Ohlsson (1995) found that the sampling error for a 12-month change of the all-commodity Swedish CPI is in the order of 0.1-0.2%.

14. The main problem with non-probability sampling is that there is no theoretically acceptable way of knowing whether the dispersion in the sample data accurately reflects the dispersion in the entire population. In other words, in this case the estimation of standard errors is very difficult. In practice, for non-probability sampling approximate techniques are generally applied for variance estimation. One is quasi-randomisation, i.e. assuming that a given set of data has been generated by a random procedure although probability sampling was not strictly used (Särndal, Swensson, Wretman, 1992, pp. 574). Thus, assumptions are made about the probability of sampling each commodity and outlet. The problem with this method is that it is difficult to find a probability model that is an adequate approximation to the method actually used for outlet and item selection in practice. Another technique is the replication method, including random groups, balanced half-samples, jackknife and bootstrap. This is a completely non-parametric method of estimating sampling distributions and standard errors. The replication method works by drawing a large number of sub-samples with replacement from the available sample. From each sub-sample the parameter of interest is estimated. The distribution of the resultant estimates approximates the true sampling distribution of the estimator when applied to the original sample. For more details on the replication method see Särndal, Swensson, Wretman (1992, pp. 418-445).

2.2 Qualitative descriptions of non-sampling errors

15. Unlike sampling errors it is difficult to find quantitative measures of non-sampling errors. So in the case of non-sampling errors the use of qualitative descriptions is the only alternative. For instance, the coverage of the sampling frames as a proxy of the target populations can be addressed (including gaps, duplications and definitional problems). The percentage of the target outlet samples from which responses or usable price data were obtained (i.e. response rates) can be provided. Any known difference in the prices of responding outlets and non-responding outlets can also be described as well as a brief indication of the method of imputation or estimation used to compensate for non-response.

3. Procedures to minimise errors

16. We can control the *estimation error* through the sampling design. For example, by increasing the sample size, or by taking selection probabilities proportional to some well chosen auxiliary variable, the error in an estimated CPI can be reduced. The choice of an adequate sampling design for a CPI is an extremely complex matter. The target population is the set of all goods and services that are acquired, used or paid for by households from outlets in a particular time period. A proper probability sampling procedure selects a sample by a random mechanism in which each good or service in the population has received a known probability of selection. In combination with a Horvitz-Thompson estimator a probability sampling design will produce an index that is (approximately) unbiased and precise.

17. The following three probability sampling designs are used extensively in survey practice: simple random (SI) sampling, PPS sampling, and stratified sampling with SI or PPS sampling in each stratum. The advantage of SI sampling is its simplicity; it gives each population element the same probability to be included in the sample. Sampling proportional to size has the advantage that the most important elements have a large chance of being sampled, while at the same time allowing a random element in the selection process. For instance, at Statistics Sweden outlets are selected with probabilities proportional to some proxy for size, namely number of employees. Unequal probability designs can lead to substantial variance reduction in comparison with equal probability designs. In stratified sampling, the population is divided into non-overlapping sub-populations called strata. For instance, at the UK Office for National Statistics (ONS) the population of outlets is split by outlet type (multiple, independent, or specialist, such as baker and butcher) to form different strata. A sample according to a certain design is selected in each stratum. One of the reasons why stratified sampling is so popular is that most of the potential gain in precision of PPS sampling can be captured through stratified selection with SI sampling within well-constructed strata. Stratified sampling is in several aspects simpler than PPS sampling.

18. Due to the fact that appropriate sampling frames are lacking, samples are sometimes obtained by non-probability methods. Judgement (or expert choice) sampling is one form of non-random selection. In this case an expert designates certain ‘typical’ elements where data are to be collected for conclusions about the whole population. With skill on the part of the expert a fairly good sample may result, but there is no way to be sure. A more sophisticated non-probability method is quota sampling. In quota sampling the population is divided in certain strata. For each strata, the number (‘quota’) of elements to be included in the sample is fixed. Next the interviewer in the field simply ‘fills the quotas’, that means in case of outlet sampling that the selection of the outlets is based on the judgement of the price collectors. Another non-probability method is cut-off sampling in which there is a deliberate exclusion of part of the target population from sample selection. In particular, this procedure is used when the distribution of the values of the target variable is highly skewed. For instance, a large part of the population may consist of small outlets whose contribution to the total sales is modest. One may decide in such a case to exclude from the frame the outlets with the lowest sales. Because the

selection is non-random, non-probability methods lead to biased estimates. However, empirical results of research undertaken by Statistics Netherlands show that a non-probability selection method do not need to perform worse, in terms of the mean square error, than a probability sampling technique (De Haan, Opperdoes and Schut, 1997).

19. Provided that the sampling design is given, the sampling variance of an estimated (all-commodities) CPI can in general be lowered by:
- increasing the sample of households needed for estimating the expenditure shares;
- constructing homogeneous commodities, i.e. groups of varieties (of a commodity) having similar price changes, or by enlarging the outlet sample per commodity;
- constructing commodity groups made up of commodities having similar price changes, or by enlarging the commodity sample.

20. It is important to allocate optimally the available resources both between and within the different CPI samples, because badly allocated samples may lead to unnecessary high sampling errors. The Swedish variance estimation results, presented in Dalén and Ohlsson (1995), showed that the error due to commodity sampling is relatively high compared with those due to outlet sampling. Then it is worthwhile to increase the sample size of commodities and reduce the sample size of outlets. This means that systematic analysis of sampling errors offers possibilities for precision improvements and/or cost decreases. The problem of optimum sample allocation is usually formulated as the determination of the sizes of the samples of commodities and outlets and their distribution over the strata that minimise the sampling error of an all-commodities CPI subject to the available budget.

21. As already mentioned, a business register usually is not an adequate sampling frame for outlets, because it has extensive *overcoverage*. It is recommended to set up an appropriate sampling frame by enumeration of the main outlets within each sampled municipality. Outlet enumeration yields a list of all outlets in the municipality together with the commodity groups that belong to their assortments. The UK ONS is the only EU national statistical office that establishes a sampling frame through outlet enumeration in the main shopping areas within each sampled locality. A less expensive way to organise such an outlet survey is to ask the price collectors to make a list of outlets where purchases are made by households; they usually know the local situation very well.

22. The populations of commodities (and varieties) and outlets are continually changing through time. The composition of most commodity groups is not constant over time, due to commodities disappearing from the market and newly appearing commodities. The factor time plays also a disturbing role with outlets: outlets become closed, either temporarily or permanently; new outlets can have started; the importance of some type of outlets can diminish or just increase. The samples of commodities (and varieties) and outlets should be reviewed periodically and they should be updated to maintain their ‘representativity’ with respect to the current buying habits of all households.

23. *Measurement errors* caused by underreporting of expenditures can be adjusted by using other available statistics such as National Accounts consumption data. See Linder (1996) for more details on the Dutch way to link the estimates to National Accounts data. Measurement errors by price collectors can be avoided by using hand-held computers in the outlets. In this way checking of observed prices is adequately done at the point of price collection (i.e. in the outlet) by automatically comparing the currently observed price with the previously observed price for that particular quotation (limit on percentage price change) and to all prices for that commodity (upper and lower limits for that price). Since January 1995, the UK ONS uses hand-held computers for local price collection (Haworth, Fenwick and Beaven, 1997).

24. In addition, a continuous attempt to avoid measurement errors is necessary. This can be done by providing an adequate training program to the price collectors. Such a training program has to include:
- obtaining a background understanding of the nature and uses of a CPI;
- accompanying an experienced price collector;
- attending a course and/or reading a manual on the price collection process.

25. It is useful to appoint data collection supervisors that conduct quality assurance checks on the data collectors. It is also a good idea to organise regularly meetings where price collectors and statisticians of the head office can interchange their experiences. Then the statisticians come into touch with conditions in the field and can use the opportunity to provide more information about frequently made price collection errors and new representative goods.

26. It is important to check the collected price data for *processing errors* and where possible, to correct errors. This activity is called data editing. Editing can be carried out on the data of the individual observations, in other words micro-editing. Selective editing and macro-editing can be interesting alternatives at a time where resources to be spend on data editing must be minimised, while at the same time maintaining a high level of data quality. Selective editing is a form of traditional micro-editing, in which the number of edits is kept to a minimum. Only edits having an impact on the survey results are carried out. Macro-editing offers a top-down approach. Edits are carried out on aggregated data (for instance commodity group price indices) instead of individual records (i.e. price observations). Micro-editing of individual records is only invoked if macro-edits show something is wrong.

27. *Non-response* usually introduces selection bias. In general, we can distinguish three methods for the treatment of missing price observations. First, we can exclude the corresponding price from the data set of previous prices, so that the set of previous prices is ‘matched’ with the set of current prices. Second, matching can be achieved by using an imputed (or artificial) price. The imputed price can be calculated by either carrying forward the previous price observation or extrapolating the previous observation using the change of the other price observations for the same commodity. Finally, it is possible to re-weight the sample. The objective of re-weighting is to inflate the weight given to the prices of the responding

outlets. This compensates for those prices that are lost through non-response.

28. In a household expenditure survey missing data are usually imputed with the help of information on the same household from a previous observation period or other households from the same observation period. To reduce bias in the expenditure patterns caused by selective non-response, in a household expenditure survey the sample of households is generally post-stratified by a number of household characteristics such income, composition and size.

4. Conclusion

29. It is important to regularly monitor and improve the quality of a CPI. The quality of a CPI concerns aspects like accuracy, relevance, comparability and coherence, accessibility and clarity, timeliness and punctuality. To control the accuracy of a CPI the errors in a CPI have to be assessed. The total error in a CPI can be divided into sampling and non-sampling errors. Sampling errors are due to the fact that an estimated CPI is based on samples. Non-sampling errors may even occur if the whole population is investigated. These errors comprise coverage errors, measurement errors, processing errors and non-response. To measure the sampling variance of a CPI several estimation techniques exist, depending on the sampling design that is applied in practice.

30. In order to ensure public confidence in a CPI, a detailed and up-to-date description of the methods and data sources should be published. The document should include, among other things, objectives and scope of the index, details of the weights, and, last but least, a discussion of the accuracy of the index. A description of the sources and magnitude of the sampling and non-sampling errors (coverage, non-response rates, etc.) in a CPI provides users with valuable data on the limitations that might apply to their uses. An example of a good handbook of CPI methods is the one that is published by the US Bureau of Labor Statistics (BLS) which deals in a separate section with the varieties and sources of possible error in the index (see BLS, 1997).

5. References

Balk, B.M., 1989, On calculating the precision of consumer price indices. Contributed Papers 47th Session of the ISI, Paris.

Balk, B.M. and H.M.P. Kersten, 1986, On the precision of consumer price indices caused by the sampling variability of budget surveys. Journal of Economic and Social Measurement 14, pp. 19-35.

Boon, M., 1998, Sampling designs in compiling consumer price indices: current practices at EU statistical institutes. Research in Official Statistics, 1 (2), pp. 39-52.

BLS, 1997, The Consumer Price Index. In: Handbook of Methods, pp. 167-202.

Dalén, J., 1995, Quantifying errors in the Swedish consumer price index. Journal of Official Statistics, 11 (3).

Dalén, J. and E. Ohlsson, 1995, Variance estimation in the Swedish consumer price index. Journal of Business and Economic Statistics, 13 (3), pp. 347-356.

Haan, J. de, E. Opperdoes and C. Schut, 1997, Item sampling in the consumer price index: a case study using scanner data. Research report (Statistics Netherlands).

Haworth, M.F., D. Fenwick and R. Beaven, 1997, Recent developments in the UK Retail Prices Index: quality management. Proceedings of the Third International Conference on Price Indices, ed. by B.M. Balk (Statistics Netherlands, Voorburg).

Linder, F., 1996, Reducing bias in the estimation of consumer price indices by using integrated data. Research report (Statistics Netherlands).

Särndal, C.-E., B. Swensson and J. Wretman, 1992, Model Assisted Survey Sampling (Springer-Verlag, New York).