European Conference on Quality in Official Statistics (Q2016)
Madrid, 31 May-3 June 2016

Uncertainties in the Swedish PPI and SPPI

Kristina Strandberg[1], Anders Norberg[2], Marcus Fridén[3]

[1] *Statistics Sweden, Stockholm, Sweden; kristina.strandberg@scb.se*
[2] *Statistics Sweden, Stockholm, Sweden; anders.norberg@scb.se*
[3] *Statistics Sweden, Stockholm, Sweden; marcus.friden@scb.se*

Abstract
In estimating the uncertainties in a sample survey it is easy to concentrate on the sampling error since it often can be quantified numerically. In the Swedish PPI and SPPI there is an established formula for estimating the sampling error. The formula takes into account the multi-stage sampling design as well as the finite population correction. However, a significant part of the uncertainties in these surveys are non-sampling errors, such as specification error, measurement error and data processing error. An effort has been made to estimate the impact of these errors and for each stratum the error contribution from non-sampling sources. It is not clear how to combine the sampling error with the non-sampling error into an overall measure of total error. We propose a method to estimate the total survey error and to identify strata with the biggest total uncertainty.
Keywords: variance estimation, non-sampling error, total survey error.

1. Introduction

Work of exploring the uncertainties in the Producer and Import Price Index (PPI) and the Services Producer Price Index (SPPI) was prompted by a project at Statistics Sweden where the main task was mapping the uncertainties in the Swedish Gross Domestic Product (GDP). A better understanding of the uncertainties in the primary statistical sources would be a first step towards mitigating errors in the final product. Studies on the topic have been carried out in project form (Isaksson et. al. 2014 and 2015). In the pilot, a number of sensitivity analyses for the Swedish GDP were conducted, investigating the impact on the GDP of different input data sources for product groups and industries, varying price indexes, and of different deflation methods.

The main study further explored the new insights made possible by the sensitivity analyses. One focus was the "SCM method": a method for automatic balancing of the national accounts (Calzaroni et. al. 1998). The method was first described in 1942 and it got its name from the creators; Stone, Champernowne and Meade. A short description of the method is given in chapter 2.

Throughout the studies, variance estimation methods for the PPI and the SPPI were studied. The conclusions are summarized in chapter 3. Different ways to evaluate non-sampling errors were looked into, which is discussed in chapter 4. A necessary effort to evaluate the total uncertainties of these indexes, however not scientifically stringent, is presented in chapter 5.

2. Price Indexes and the National Accounts

2.1. Producer and Import Price Index and Services Producer Price Index

PPI is a monthly survey that aims to show the average change in prices in producer and import stages for different product groups. Prices are measured in the first distribution stage, when products exit the production process from Swedish producers or when products cross the Swedish customs frontier entering the Swedish market.

The quarterly SPPI aims to show the average price change for services produced in Sweden.

2.2. Population, Sample Selection and Estimation

Objects in the PPI and the SPPI are transactions and the target population is formed by all the transactions referring to goods and services sold by a Swedish producer or imported by a Swedish importer. Market weights for product groups are calculated from previous year's sales. Design weights on the other hand, stem from the sampling design. These weights are combined to form the weights used in estimation. With this information, the price change between the current period and a base period is calculated for each of the product groups in the survey.

A sampling frame is constructed by combinations of companies and product groups from other surveys, such as Foreign Trade in Goods and the Structural business statistics. From the frame, a first-stage πps-sample of companies is selected within each stratum/product group. In a second-stage Statistics Sweden and the respondent cooperate in selection of a typical transaction to follow monthly. In total, about 9000 transaction prices are collected for PPI/SPPI during 2015

2.3. National Accounts

One important aspect of the national accounts (NA) is that all values are presented in current prices as well as fixed (free from inflation) prices. Turning a current price value into a fixed price value is called deflation.

Most fixed price calculations in the Swedish NA are done by dividing values in current prices with corresponding price indexes. Indexes used are the Consumer Price Index (CPI), the PPI, the SPPI and the Building Price Index. Uncertainties in these price indexes are thereby carried directly into the NA.

There are three approaches for calculating the GDP in the NA: the production approach, the expenditure approach and the income approach. Theoretically these estimates should be the of the same size, but since they are calculated from different independent sources, there is usually a discrepancy. Therefore a post-adjustment ("balance") of these estimates is necessary in order to comply with the requirements of the accounting system. The balancing of the accounts is a big task, relying almost exclusively on subjective methods.

In short, the SCM is a generalized least square method, using inverted uncertainty measures for industries and product groups as weights. The method is used to minimize the sum of squared discrepancies between the estimates in the accounting system. A fully automatic and reproducible balancing method could be an important tool to create benchmark values for the NA. The method has been used for example by Calzaroni (1998) and Chen (2006 and 2012).

3. Sampling Errors

Results from earlier investigations of the variance estimation methods were studied and confirmed by a simulation study. Two variance estimates were picked out as the best based on their performance in the simulation. For sample allocation we use a formula by Wingren (2009) (**Fig. 1**), the finite population correction is however not necessary. For quality measures Roséns (2010) variance estimation formula for πps-sampling (**Fig. 2**) proved to be the best.

Wingren's method (1)

$$\widehat{V}_1(I) = \frac{\sum_{i=1}^{n} \omega_i^2 \times \hat{V}_1(\bar{I}_i)}{\sum_{i=1}^{n} \omega_i^2}$$

where $$\hat{V}_1(\bar{I}_i) = \frac{\sum_{j=1}^{n_i} \dot{\omega}_{ij} \left(I_{ij} - \bar{I}_i\right)^2}{\sum_{j=1}^{n_i} \dot{\omega}_{ij} \times n_i} \times \frac{n_i}{n_i - 1} \times (1 - f_i)$$

and $$\omega_i = \sum_{j=1}^{n_i} \dot{\omega}_{ij} \qquad \bar{I}_i = \frac{\sum_{j=1}^{n_i} \dot{\omega}_{ij} I_{ij}}{\sum_{j=1}^{n_i} \dot{\omega}_{ij}}$$

i = stratum
j = observation
I_{ij} = price ratio for observation j, stratum i
n_i = number of observations in the sample in stratum i
f_i = finite sample correction (in terms of market shares)
w_{ij} = weights used in estimation

Rosén's method (2)

Variance estimates are valid for population totals. We adapt the method to our purposes according to d).

a) Population total, $\tau(y)$, is estimated by;

$$\hat{\tau}(y) = \sum_{i \in urvalet} \frac{y_j}{\lambda_j}$$, where the λ_j are the selection probabilities

b) The variance of the estimator is given, with good approximation, by:

$$V[\hat{\tau}(y)] \approx \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_j}{\lambda_j} - \sum_{j=1}^{N} y_j \cdot (1-\lambda_j) / \sum_{j=1}^{N} \lambda_j \cdot (1-\lambda_j) \right)^2 \cdot \lambda_i \cdot (1-\lambda_i)$$

c) Consistent estimation of $V[\hat{\tau}(y)]$ is given by:

$$\hat{V}[\hat{\tau}(y)] \approx \frac{n}{n-1} \cdot \sum_{i=1}^{N} \left(\frac{y_j}{\lambda_j} - \sum_{j \in urvalet} \frac{y_j \cdot (1-\lambda_j)}{\lambda_j} \Big/ \sum_{j \notin urvalet} (1-\lambda_j) \right)^2 \cdot (1-\lambda_j)$$

d) We let $y_j = \log(index_j) \times \omega_j$. The population total for y_j divided by the total weight then becomes a weighted mean of the indexes.

4. Non-Sampling Errors

It is a well-known problem in production of statistics that non-sampling errors are rarely estimable. For our purposes, we have asked expert personnel to decide if the non-sampling error contribution is either "Low", "Medium" or "High." Non-sampling errors have been evaluated on product group level, i.e. the three, four or five-digit level of SPIN[1]. The aim has been to take into account all sources of error which are not covered by the sampling variance. The following five sources of error have been identified:

- Specification error[2]
- Frame error
- Non-response error
- Measurement error
- Data processing error

---

[1] The Swedish version of the European product classification CPA

[2] Specification error proposed by Biemer et. al. (2014) is not defined as an error source in the Code of Practice.

Frame error is assumed to be distributed evenly across product groups, and therefore not differ much between product groups. The same reasoning is used for data processing, which is carried out automatically in the software. The only data processing done manually is quality adjustments and they are always judgmental which gives a certain degree of uncertainty. This has been taken into account in the overall assessment, but not given as much impact as the error sources mentioned below.

Specification error is considered to give the highest contribution to the overall error. Specification error means that we do not measure exactly what we want to measure. Such errors in PPI are due to use of list prices and hourly rate methods. The proportion of list prices and hourly rates based methods, in total, has been calculated for each product group. An occurrence of between 0 and 25 percent is deemed “Low.” Over 25 percent but less than 50 percent is deemed “Medium,” and 50 percent of more is deemed “High.” The remaining error sources, non-response and measurement error have also been evaluated and are considered low risk.

These five separate sources of error have been combined into one over all measure. More than an error, this measure could be seen as “risk.” As an example, prices for personal computers are measured by highly experienced staff, using A-methods. Still, the character of the product group such as fast product development and frequent quality adjustments, result in a higher risk of faulty values being collected. PPI consists of a total of 508 strata. Out of these, 101 were deemed to have a non-sampling error contribution that is “Low,” for 260 strata the contribution was deemed “Medium” and for 147 strata the contribution was deemed “High.” The corresponding numbers for the 69 SPPI strata are 19, 27 and 23.

It should be pointed out that evaluating the non-sampling error in this manner is a highly subjective method. Undoubtable will it be affected by the prior knowledge and experience of the person/persons doing the evaluation. For the work presented in this paper, one subject expert handled the entire task. A better method would be to obtain several opinions and use an average.

5. Combining Error Measures

5.1 Random Errors vs. Bias

It is, for example, reasonable to think that list prices (which are used in place of hard to collect transaction prices) are systematically higher than transaction prices including withdrawn discounts etc. Hence, per definition a list price is a source of bias. It is a harder task to determine if a list price will result in an overestimation or an underestimation of the index. List prices are usually fixed for a period of time, probably resulting in periods of systematic underestimation of the true price movements but not necessarily an underestimation of change. When the list price is adjusted according to the market, it is safe to assume that the price movement will be an overestimation of that period's true price movement. Monthly indexes are then averaged to create a yearly index and it is hard to say if the final result will be an over- or an underestimation of the true value.

We argue that non-sampling errors in the PPI and the SPPI mainly can be looked upon as random. We cannot say that a certain error source always result in an overestimation or an underestimation. Random errors will show up in the variance estimate and there is a risk that adding separately estimated non-sampling errors will overestimate the total uncertainty. We have not seen any evidence that a stratum is unreasonably punished with a high total uncertainty measure. It is however fully possible that a stratum that is almost completely covered in the sample and thereby have very little (if any) sampling error, still can have a substantial amount of uncertainties added from non-sampling sources. The combination of the sampling error and this non-sampling error is what we are trying to estimate.

5.2 Creating One Uncertainty Measure

It is not clear how to combine the sampling errors with the non-sampling errors. Literature studies have not resulted in any obvious solutions. We propose a method where the assessed non-sampling errors are given numeric values and then combined with the estimated sampling error.

As a starting point, we assume that the sampling errors and the non-sampling errors accounts for 50% each of the total uncertainties in the entire survey (i.e. all strata combined). This is a strong assumption, but we don't have any evidence suggesting the scale to tip either way. We developed this algorithm:

1. The assessed non-sampling errors are given numerical values according to Low = 1, Medium = 3 and High = 9.

2. Errors are rescaled so that sampling errors and non-sampling error account for 50% each of the total uncertainty in the survey, for all strata combined.

3. For each stratum, the *Total Uncertainty* (TU) is calculated as:

   $$TU = Sample\ Variance + Assessed\ Non\text{-}Sampling\ Error$$

4. An indicator is calculated as:

   $$Indicator = weight^2 \times TU$$

5. Plot standard deviation vs. stratum weight for all strata

6. In the plot, mark strata with high indicator values.

For the Swedish PPI this algorithm results in the plot shown in in **Fig. 3**. For the sake of clarity, strata with extreme weights (for ex. mining, petroleum and auto manufacturing) are excluded from the plot – these strata should always be examined carefully. As can be seen in the plot, it is the combination of variance, non-sampling uncertainty and weight that decides if a stratum is marked as influential or not.

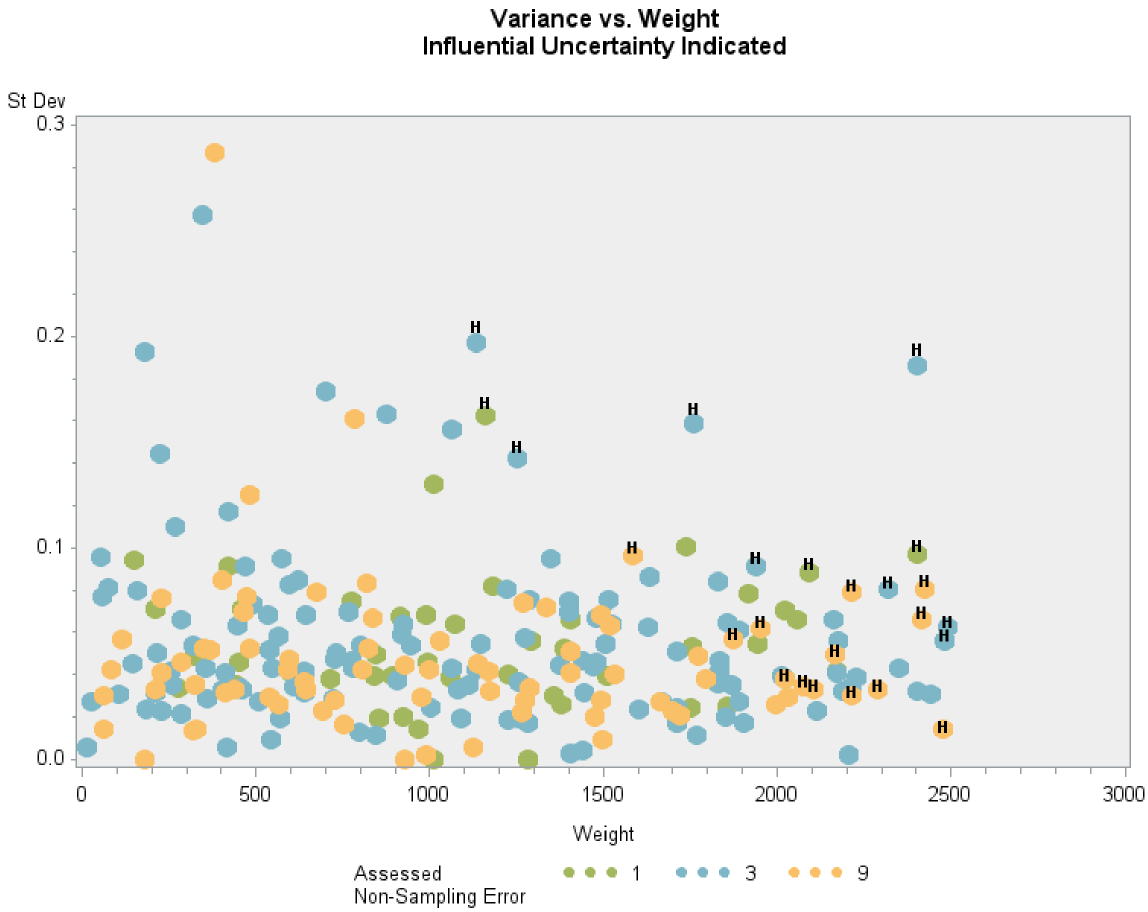

Fig. 3. A plot of standard deviation vs. weight for strata in the PPI. Strata with extreme weights are excluded in the plot. Strata with influential total uncertainty, i.e. a high indicator value are marked with an H.

6. Final Remarks

Under the umbrella of the GDB project, work was carried out with three major aims:

1. Gaining a better understanding of the error profile in the PPI and the SPPI
2. Creating a measure of total uncertainty to use in the SCM method for automatic balancing of the NA
3. Identifying strata where error mitigation efforts are needed the most

A careful study of the error profile for the PPI and the SPPI helped shed light on where quality improvement efforts would be most efficient. After an evaluation of the different error sources we found that mitigating non-sampling errors is the most efficient way of improving the quality of the surveys.

By combining the sampling variance with the non-sampling error according to the algorithm presented, we were able to create numeric estimates of the total uncertainty of each stratum in the PPI and the SPPI. While the estimates are not statistically stringent, they could be used in the SCM method to automatically balance the NA. In addition, estimates are also used to create a plot where strata with big uncertainties are identified. This information can be used for quality reports as well as a tool in the manual work of balancing of the NA.

7. References

Biemer, P., Trewin D-, Bergdahl H. and Japec L. (2014), A System for Managing the Quality of Official Statistics, *Journal of Official Statistics*, Vol. 30, pp. 381-415.

Calzaroni, M. and Puggioni, A. (1998), *Evalutaion and Analysis of the Quality of the National Accounts Aggregates*, Report to the Task Force on Accuracy Assessment of National Accounts, Eurostat.

Chen, B. (2006), *A Balanced System of Industry Accounts and Structural Distribution of Aggregate Statistical Discrepancy*, working paper WP2006-8, Bureau of Economic Analysis, Washington, DC.

Chen, B. (2012), A Balanced System of U.S. Industry Accounts and Distribution of Aggregate Statistical Discrepancy by Industry, *Journal of Business and Economic Statistics*, Vol. 30 (2), 202-211.

Isaksson, A., Xie, Y., Strandberg K., Lindholm, P., Lennmalm, A. and Lennartsson, D. (2014), Känslighetsanalys i BNP Förstudie, Statistics Sweden

Isaksson, A., Fridén, M., Lennartsson, A., Lennmalm, A., Saltveit, M., Strandberg, K. and Xie, Y. (2015), Känslighetsanalys i BNP Huvudstudie, Statistics Sweden

Rosén B. (2010), Teori och praktik för urvalsundersökningar, Kompendium, Mattematisk statistik Stockholms universitet, Sweden.

Stone, R., Champernowne, D.G. and Meade, J.E. (1942), The Precision of National Income Estimates, *The Review of Economic Studies*, Vol. 9, 111 – 135.

Wingren, J-E. (2009), *Variance Analysis for PPI and SPPI*, final technical implementation report, Price Statistics Unit, Statistics Sweden.