United Nations ECE/CES/GE.22/2012/4

Economic and Social Council

Distr.: General
28 March 2012
English
Original: French

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Consumer Price Indices

Eleventh session
Geneva, 30 May–1 June 2012
Item 3 of the provisional agenda
Price collection methods

New experiences with scanner data in the Swiss CPI[1]

Note by the Federal Statistical Office of the Swiss Confederation

Summary

Since 2008, the Federal Statistical Office (FSO) has used scanner data from large retail chains to calculate the Swiss Consumer Price Index (CPI) and the Harmonized Index of Consumer Prices (HICP). Initially, the use of scanner data was restricted to the food and near-food commodity groups, with four of the large retail chains already participating. In the second phase, the project examined the practical aspects of price collection using scanner data for non-food categories. Analysis shows that it is possible to identify non-food groups of commodities that are sufficiently stable to allow collection using scanner data, together with the conventional sampling and calculation methods. However, other commodity groups with products subject to rapid technological change or fashions are not suitable.

In another component of the project, Switzerland has acquired a Statistical Analysis Software (SAS) tool that makes it possible to calculate indices using the Rolling Year Gini Eltetö Köves Szulc (RYGEKS) method. These indices are used as references for purposes of analysis. The main features of the tool are presented separately.

[1] Late submission of this document owing to delays by external partners.

I. Introduction

1. A number of national statistical offices now use scanner data to produce their official price indices. Since 2008, FSO has used scanner data from the large retail chains to calculate the Consumer Price Index (CPI) and the Harmonized Index of Consumer Prices (HICP). The use of scanner data makes it possible to avoid having to collect prices in the field and provides a source of better quality data. It also represents significant cost savings and reduces the administrative burden for the large retail chains.

2. The Swiss approach to the use of scanner data is to continue using the same sampling and calculation methods as previously, while enjoying the benefits of an improved source of data. The price information provided by the scanner data gives better coverage in time and space than does the conventional method of price collection in the field. To keep the same sampling methodology, the scanner data are tailored to provide all the information needed by the collector, who simply works on-screen rather than collecting the information in the field. The sample used to calculate the CPI thus continues to be taken on a monthly basis, using the same methods as with the conventional field system (Becker Vermeulen, 2006; Müller, 2010).

II. Project update for food and near-food categories

3. For the food and near-food categories, the first task was to establish whether the standard sampling and calculation methods used by FSO could also be used with scanner data. In other words, the objective was to assess whether scanner data could be used as an improved source of data with the existing sampling and calculation procedures. Once the feasibility of this had been demonstrated in the context of the Swiss CPI, the next objective was to include scanner data from different retail chains (Müller, 2010).

A. Inclusions

4. A pilot software was used to bring the first retail chain into the project in July 2008. The development of a generic software tool then made it possible to include a second chain (April 2010), a third (April 2011) and finally a fourth (April 2012). The possibility of including a fifth retail chain is currently being looked at.

5. A preparatory and test phase precedes each new inclusion. The main stages are:

- Contact the retail chain for agreement to supply scanner data
- Establish required form and quality of scanner data to be supplied
- Carry out test price collection and analyse results
- Develop an emergency plan in case of non-delivery of scanner data
- Reorganize price collection in the field
- Include scanner data for the retail chain concerned

6. In the food and near-food categories, the four retail chains already included currently account for between 75% and 80% of total market turnover. It would still be possible to include other retail chains in the future.

B. Problems encountered

7. The inclusion of several retail chains for the food and near-food commodity groups has highlighted a number of practical issues to which pragmatic solutions have had to be found. They are related to specific features of the retail chains and the resultant quality of the scanner data. These factors may have a significant influence on the commodity groups for which scanner data can be used.

8. The problems encountered and measures taken to address them are:

- Degree to which the computer system is integrated into the retail chain's management. This may lead to missing scanner data for whole product groups (often fresh products). Where this is the case, price collection in the field has to be maintained alongside collection using scanner data.
- Absence of a clear description of the products, making price collection from scanner data difficult (item selection and replacement). Price collection in the field has to be maintained in this case too.
- Lack of a clear indication of the quantity referred to by price indicated. This also makes it difficult to collect prices using scanner data, particularly when replacing items. Different and not always clearly identifiable practices in reporting prices for a certain quantity have been found, particularly for groups of commodities where prices are set regionally, for which scanner data aggregated by region are required. Collection in the field also has to be maintained in this case, to avoid bias in the indices.
- Marketing effects: when a retail chain is first included, measures must be taken to avoid price collection beginning with especially low prices.
- Cost: price collection using scanner data makes it possible to improve the quality of the indices. However, the savings made on collection costs are sometimes low or even non-existent, particularly for small retail chains. The potential savings depend particularly on the existing conventional data collection. In such a case, the decision must be taken by the management.

III. Non-food category

A. Introduction

9. In the second phase of the project, it had been planned to look at the non-food category, for which alternative processing methods would need to be analysed. The objective was to assess whether it was a suitable time to use alternative sampling methods (rules for price collection, implicit quality adjustment methods) and calculation methods to make greater use of the potential of scanner data (Müller, 2010).

10. After preliminary analysis, the non-food category was divided into stable non-food and dynamic non-food. Stable non-food includes the groups of non-food commodities that show similar behaviour to the food and near-food groups. Dynamic non-food corresponds to commodity groups that have rapidly changing selections or an important technological component, or are affected by fashions.

11. It was finally decided to concentrate solely on stable non-food in order to make greater use of the data while retaining the same sampling and calculation methods. Thus, the main objective was to identify groups of non-food commodities with sufficiently stable behaviour and for which quality characteristics play only a minor role.

12. Despite the definite improvement in quality, the introduction of price collection using scanner data for just one part of stable non-food is clearly less effective than for food and near-food, given the market structure and available data sources.

13. The work on dynamic non-food has not been continued. This was decided because of the current lack of any simple methods for dealing with the problem of price skimming, as well as the frequent changes of selection which, together with the need for quality adjustments, risk leading to bias in the indices. Given the resources available and the methods that could be used in the project, it was considered too ambitious to take on the dynamic group.

B. Analysis method for non-food category

14. The first task was to identify the groups of commodities that could be considered stable and those that were dynamic, that is, to identify non-food groups with enough stability characteristics to be processed using scanner data, like the food and near-food groups.

15. A test price collection was then carried out using scanner data to compare the outcome to the results of price collection in the field. RYGEKS reference indices were also produced to judge the comparability of the results.

1. Data source

16. Agreement was reached with a retail chain already involved in price collection using scanner data for food and near-food on supplying test data for non-food. The data covered a 39-month period from January 2008 to March 2011. Several non-food commodity groups were not available and thus had to be excluded from the analysis.

17. It was not possible to purchase data allocating the items covered by the study to the Classification of Individual Consumption According to Purpose (COICOP) from a marketing institute, as is done for food and near-food. For the purposes of the analysis, therefore, the items were allocated on an aggregated level, with the problems that entails (Müller, 2010). It is not always possible to do this where the structures are different, particularly if the retail chain structures items according to brand. In that case, allocation has to be on item level. Here too, it was only possible to draw clear conclusions for groups of commodities where allocations were made without any major problems.

2. Definition of stable non-food and dynamic non-food

18. The analysis showed that items with stable behaviour and items with dynamic behaviour are not always restricted to specific survey positions or commodity groups. The approach used to define stable non-food in contrast to dynamic non-food was to consider that the survey positions can have more or less stable or dynamic behaviour depending on the characteristics of their respective selections.

19. We used criteria based on the lifespan of the items (no marked rotation in the selection), their mean ranking in terms of turnover and the suspected presence of price skimming to help us decide on the stability of the selection and thus to put each position in the basket of goods on a scale ranging from pure stable non-food to pure dynamic non-food.

20. This classification made it possible to conduct an initial screening, which identified those positions with a selection judged stable enough to be included in the definition of stable non-food, and those judged not stable and thus considered as dynamic non-food.

21. The positions identified as part of stable non-food make up groups 1 to 3 in the typology shown below. The positions considered as part of dynamic non-food, and thus not suitable for collection using scanner data, make up group 4 in the typology.

3. Test collection

22. Once the stable non-food groups had been identified, a test collection was carried out. The aim was to compare the indices produced using the conventional price collection method, a collection using scanner data and indices produced using the RYGEKS method (see below for details). This was to ensure that collection using the current methods produced close results (considered comparable), or that any differences found could be explained.

23. To speed up the test collection, a manual collection was simulated using scanner data. This was done in SAS and included the main characteristics of a manual price collection using scanner data.[2]

C. Analysis results and interpretation

24. The results of the analyses show that non-food can be broken down into four distinct groups, the first three covering stable non-food categories and the fourth representing dynamic non-food.

Group 1

25. The first group contains those positions identified as clearly suitable for collection using scanner data, together with the conventional sampling and calculation methods. It covers the following:

- Materials for the maintenance and repair of the dwelling (COICOP 04.3.1)
- Household utensils (COICOP 05.4)
- Tools and equipment for house and garden (COICOP 05.5.2)
- Household goods (bin liners, glues, string, etc.) (COICOP 05.6.1)
- Stationery and drawing materials (COICOP 09.5.3 and 09.5.4)
- Travel goods and accessories (COICOP 12.3.2)

26. The analyses showed that these positions have stable selections, little or no price skimming and that the RYGEKS indices are very close to the results produced by the conventional collection method (part of the results are given in annex 1, figures 1 to 4).

27. The differences noted in comparison with the conventional collection method arise primarily from differences in samples. For instance, the items selected in the conventional collection in the field are practically never the ones that have generated the largest turnover. A conventional price collection is thus based on items in the selection that are stable but generally have not generated the largest turnover. Price collection using scanner data makes it possible to focus better on those items with the highest turnover, which is particularly desirable for positions that include homogeneous items.

[2] As the simulation tool is still under development, the results are not presented in this paper.

28. There are also some seasonal effects taken into account in the scanner data indices that do not appear in the indices based on the conventional collection methods, which demonstrates the better coverage in both space and time with the scanner data.

29. For some positions, special collection rules — which already exist for collection using scanner data for food and near-food — have been introduced to ensure that collection using scanner data produces plausible results. In particular, they make it possible to avoid some of the dangers linked to the influence of low prices during the period preceding item replacement.

Group 2

30. Group 2 consists of survey positions that have strong potential for the introduction of price collection using scanner data together with conventional methods, but where the analyses have been only partly conclusive. They are:

- Household textiles (COICOP 05.2)
- Household appliances (COICOP 05.3)
- Motorized tools for house and garden (COICOP 05.5.1)
- Electric appliances for personal care (COICOP 12.1.2)

31. In the end, these positions had to be left out for the moment. Problems were encountered in allocating items on the aggregated level to the CPI basket of commodities. The partial lack of relevant test data or the absence of a comparison with the conventional collection method (which was sometimes the case with the large retail chain in question) were also reasons for these product groups being left out.

32. Further analyses (additional manual test collection, full data source, more specific allocations) need to be carried out before a final decision can be taken on the possible inclusion of these positions.

Group 3

33. The third group contains products that could potentially be used in the future for collection using scanner data but that have not yet been analysed. Two specific areas are concerned: conventional clothing (COICOP 03 in part), which would require an extended analysis framework, and furniture (COICOP 05.1), for which data are not currently available.

Group 4

34. Lastly, the fourth group contains products which are clearly not suitable for collection using scanner data with the methods currently available. The products concerned are other clothing, bicycles, televisions (see figure 5, annex 1) and computer hardware. These are clearly product groups that consist mainly of dynamic non-food items.

35. In conclusion, these analyses have enabled us to draw up a restricted list of positions of the stable non-food category, for which collection using scanner data could be introduced in the short term (group 1). The second list of positions (group 2) is a strong contender for a future introduction, but still requires further analyses. The third list of positions (group 3) also has some potential, but the analyses are yet to be carried out. The positions mostly linked to dynamic non-food (group 4) are considered to be clearly unsuitable for collection using scanner data.

IV. The RYGEKS algorithm

36. The objective of another component of the project was to develop an appropriate tool for analysis and comparison in the form of an SAS module for the production of indices based on the RYGEKS method.

A. The RYGEKS context

1. Basis

37. The use of scanner data in constructing price indices has been discussed in the international literature for several years. However, an automated collection method that makes best use of the information available still represents a challenge, and the first valid solutions have only just been developed.

38. Ivancic et al. (2009) presented some key work in this area, with scanner data from several Australian retail chains and solutions for price index calculations. The methods were taken up very quickly by de Haan and van der Grient (2009) and then Nygaard (2010), who confirmed the Ivancic et al. (2009) results. The method developed by Ivancic et al. (2009) is still a long way from being directly usable in calculating an official index, but its properties have convinced a community of specialists who use it for analysis and comparison.

39. Ivancic et al. (2009) used extended scanner data collected in Australia over 15 months giving weekly sales transactions by store for several retail chains. They showed that the use of superlative indices (such as the Fisher and Törnqvist indices) applied to raw scanner data can lead to unstable or biased indices (chain drift). The problems are a result of:

- Chaining: it is particularly common in food categories for new items to appear and others to disappear from sales. Periodic chaining has to be used to take account of this. It has been demonstrated that frequent (weekly and, in some cases, monthly) chaining can lead to unstable indices. The alternative is to use a fixed base index, which does not need to use periodic chaining but also has serious disadvantages.
- Price bouncing: special offers can have a significant effect on sales, with an increase in turnover. Subsequently, when prices return to their normal level, turnover may fall temporarily to almost zero. Where this happens, an index within which prices are weighted according to sales and which does not return to its previous level is liable to systematic bias.
- Seasonal items and missing prices: these are found by Nygaard (2010) and van der Grient and de Haan (2011).
- Price aggregation: the concept of price aggregation can have an effect on the results of index calculations. Ivancic et al. (2009) discuss the effects of price aggregation on different chains and sales points. However, both Nygaard (2010) and van der Grient and de Haan (2011) maintain aggregation concepts in existing index calculations. Some questions do nevertheless arise where scanner data are used in a price index.

40. In order to solve these problems, Ivancic et al. (2009) propose a multilateral method using GEKS (named for its authors, Gini, Eltetö, Köves and Szulc), an index commonly used to compare purchasing power between countries (ILO, 2004, Annex 4; PPP Manual Eurostat-OECD, 2006).

41. The GEKS methodology itself cannot be applied to a published price index, given the need to review previous periods with each new collection period included in the

calculation. Ivancic et al. (2009) thus developed a method called RYGEKS (Rolling Year GEKS) to correct that shortcoming. By using a sliding 13-month window of data for the index calculation, the results can be "fixed".

2. Advantages of the RYGEKS method

42. The RYGEKS method has the following advantages:

- RYGEKS represents a complete price collection that makes use of all the available data. The prices are weighted according to turnover or to their share of sales in the relevant food index.
- New and disappearing items in the selection are included in the index calculations when available.
- Chaining is not an issue in the context of RYGEKS.
- The effects of price bouncing and missing or seasonal items are kept under control.
- RYGEKS indices do not suffer from chain drift. This is a key property, which shows the GEKS/RYGEKS method to be an effective analytical tool (Ivancic et al., 2009). The result is confirmed for food but must be put into perspective in the context of non-food.

3. Limitations of the RYGEKS method

43. The GEKS/RYGEKS method cannot be considered infallible, for a number of reasons:

(a) In some circumstances, the raw data may need to be corrected (de Haan and van der Grient, 2011) to avoid index errors. Inadequate data may produce implausible results;

(b) While there have been a fair number of experiments for the food group, where RYGEKS produces generally reliable results, little has been published on other areas, which use specific rules for price collection and experience frequent changes in quality (particularly computers and cars);

(c) One condition for the use of RYGEKS is that a retail chain's items have to be allocated to the standard CPI basket, and this may be complicated or laborious for certain groups, as in the case of clothing, where retail chains often structure items by brand. In such cases, it would not be efficient to use RYGEKS.

4. Developments in the context of published indices

44. It seems unlikely at this time that national or international agencies would approve the use of a RYGEKS index to calculate a published index, as there has certainly not yet been enough practical experience. Furthermore, a multilateral method does not seem very practical in a working context – the methods for checking the plausibility of reported prices are mainly based on a bilateral index approach.

45. Nevertheless, the national statistical offices of both the Netherlands (van der Grient and de Haan, 2011) and Norway (Johansen and Nygaard, 2011) do take established index formulae (Jevons and Törnqvist, respectively) and apply them to scanner data using partly automated methods, working with a larger number of prices than in the case of conventional price collections in the field. This can be done by processing raw data at the elementary index sample level. It is thus possible to avoid the problems noted by Ivancic et al. (2009). In that context, the RYGEKS method can only be used as a reference tool to check the quality of the results.

B. Use of RYGEKS in the Swiss Federal Statistical Office

46. We took as our basis a SAS code from the RYGEKS algorithm kindly made available to us by Statistics Norway. Adaptations were made to take account of the particularities of the Swiss CPI and various extensions that are explained below. The mathematical formalization of the RYGEKS method used is given in annex 2.

1. Structure of the Swiss CPI

47. The structure of the Swiss CPI is specific in a number of respects. Initially, Jevons-type elementary indices are calculated for each position in cells defined by region and by distribution channel. The distribution channels may include several sales points.

48. The Jevons indices are then aggregated to partial indices using an arithmetical weighting system and then to the CPI total using Lowe's formula. The weightings are updated periodically.

49. The scanner data provided by the retail chains are aggregated to represent their own pricing policies. The "centralized" products, which are sold at the same price throughout the country, are thus distinguished from the "decentralized" products, the prices of which may vary by region. This is particularly the case for some fresh products.

50. The RYGEKS method is generally proposed for the calculation of elementary indices. In the context of the Swiss CPI, the original code was adapted to make it possible to create a single elementary index for the "centralized" products and separate elementary indices for each region for the "decentralized" products.

51. The original code was furthermore reworked so that RYGEKS indices could be created at two different levels of the basket of goods, that is, the more detailed level (survey positions) and the first weighted level.

2. Use of aggregated scanner data

52. The scanner data are made up of transactions over a certain period of time and in a large number of sales points within a retail chain. To calculate the indices, the transactions are aggregated by item and sometimes by sales point to take account of the different pricing policies (see above). A unit price is created by dividing the cumulative turnover by the quantity sold. This takes account of the price variations over the period in question and may be considered as a mean value.

53. Ivancic et al. (2009) discuss the effects of different types of aggregation of raw scanner data. In our work, we maintain fixed aggregation rules for the whole CPI basket of goods. It would be possible to aggregate items over different retail chains, but we consider each chain as a separate distribution channel and avoid aggregating them.

3. Extensions

54. The following extensions were added to the basic code:

- Interface programming to make it possible to use our own source files with the scanner data; these also contain the necessary allocation of items to COICOP.
- The original code was intended for use with a Törnqvist index rather than the Fisher index used by Ivancic et al. (2009). We therefore also wanted to use the Fisher index to be able to compare the results of the different formulae.
- The way the results are presented was modified. The objective was to produce an interface that would make it possible to read the RYGEKS results and be able to aggregate them to the CPI total for comparison.

4. First results using the RYGEKS tool

55. As stated in chapter 4, the RYGEKS tool was used, in particular, to produce reference indices for analysis of the stable non-food category (see annex 1, figures 1 to 5). It was also used to make a comparison for the food group (COICOP 01) for a particular retail chain between an index produced with prices from scanner data but based on a price collection using conventional methods and the RYGEKS method (see annex 1, figure 6). In the latter case, the results were very close.

V. Conclusions

56. Since the project began in 2005, the objective has been to make the most pragmatic use possible of scanner data, firstly because it can ease the burden on the companies involved, and secondly because these data can significantly improve the quality of results. Scanner data are currently being used for food and near-food items from four retail chains, and a fifth chain is preparing to join the system. Prices for part of the non-food items will also be collected using scanner data during this year.

57. Overall, we draw very positive conclusions from what has been done and what remains to be done under this project. The scanner data are a very rich source of information for price statistics, but their use also raises many practical issues and represents a significant investment in terms of information technology development. Before embarking on such a project, it is important to establish the cost-benefit ratio, according to the specific features of the national market, and the availability, and particularly the quality, of scanner data, as these affect the quality of the published CPI.

VI. References

Becker Vermeulen, C. (2006), Recent developments in the Swiss CPI – Scanner data, telecommunications and health price collection, Paper presented at the ninth meeting of the Ottawa Group, 14–16 May 2006, London, United Kingdom;

Eurostat-OECD (2006), Methodological Manual on Purchasing Power Parities;

van der Grient, H.A. and J. de Haan, Scanner Data Price Indexes: The “Dutch Method” versus Rolling Year GEKS, Paper presented at the twelfth meeting of the Ottawa Group, 4–6 May 2011, Wellington, New Zealand;

de Haan, J. and H.A. van der Grient, Eliminating Chain Drift in Price Indexes Based on Scanner Data, Paper presented at the eleventh meeting of the Ottawa Group, 27–29 May 2009, Neuchâtel, Switzerland;

ILO, IMF, OECD, Eurostat, United Nations, World Bank (2004), Consumer Price Index Manual: Theory and Practice, Geneva: ILO Publications;

Ivancic, L., W.E. Diewert and K.J. Fox (2009), Scanner Data, Time Aggregation and the Construction of Price Indexes, Discussion Paper 09-09, Department of Economics, University of British Columbia, Vancouver, Canada;

Johansen I., Nygaard R., Dealing with bias in the Norwegian superlative price index of food and non-alcoholic beverages, Paper written for the twelfth meeting of the Ottawa Group, 4–6 May 2011, Wellington, New Zealand;

Müller, R., Scanner data in the Swiss CPI: An alternative to price collection in the field, Room document for the Joint UNECE/ILO meeting of experts on Consumer Price Indices, 10–12 May 2010, Geneva, Switzerland;

Nygaard R. (2010), Chain Drift in a Monthly Chained Superlative Price Index, Paper presented at the Joint UNECE/ILO Workshop on Scanner Data, 10 May 2010, Geneva, Switzerland.

Annexe 1

Résultats

Les courbes présentées montrent la comparaison entre l'indice calculé avec les prix provenant du relevé traditionnel sur le terrain et les indices RYGEKS produits avec les données scannées.

Figure 1
Relevé traditionnel vs RYGEKS : COICOP 04.03.1 (produits d'entretien du logement)

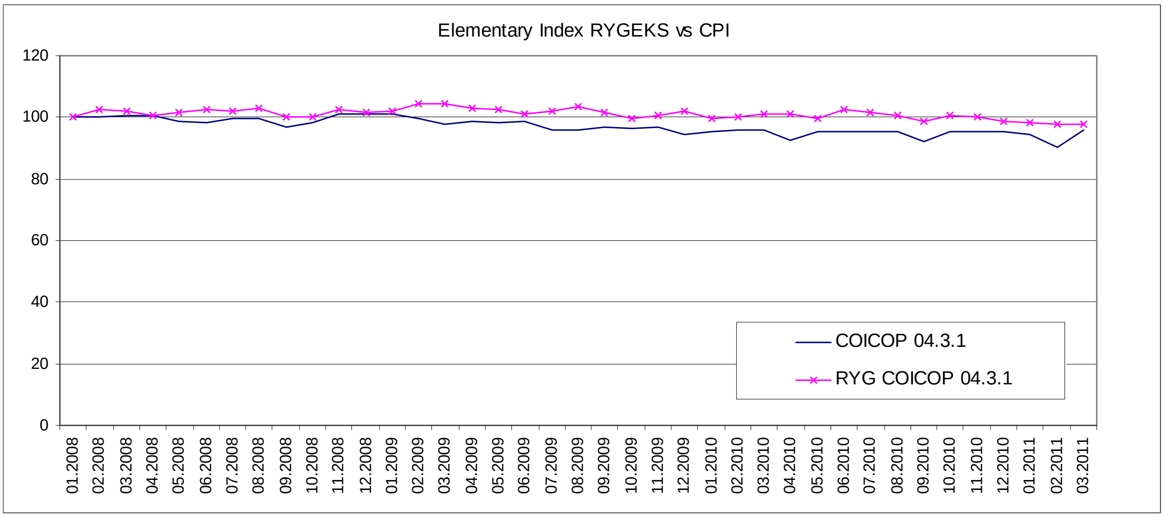


Figure 2
Relevé traditionnel vs RYGEKS : COICOP 05.5.2 (petits outils)

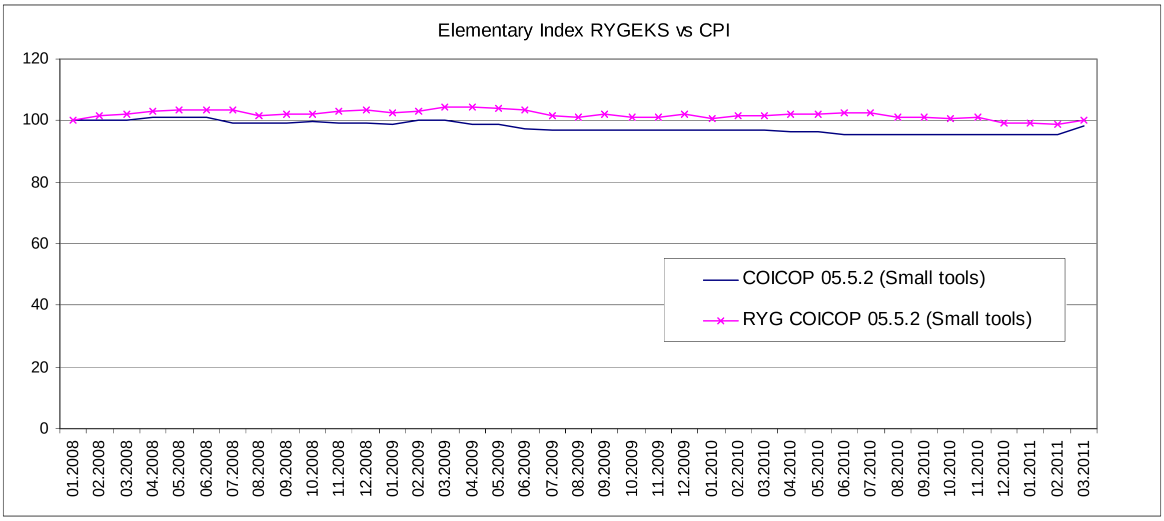


Figure 3

Relevé traditionnel vs RYGEKS : COICOP 05.5.2 (accessoires divers)

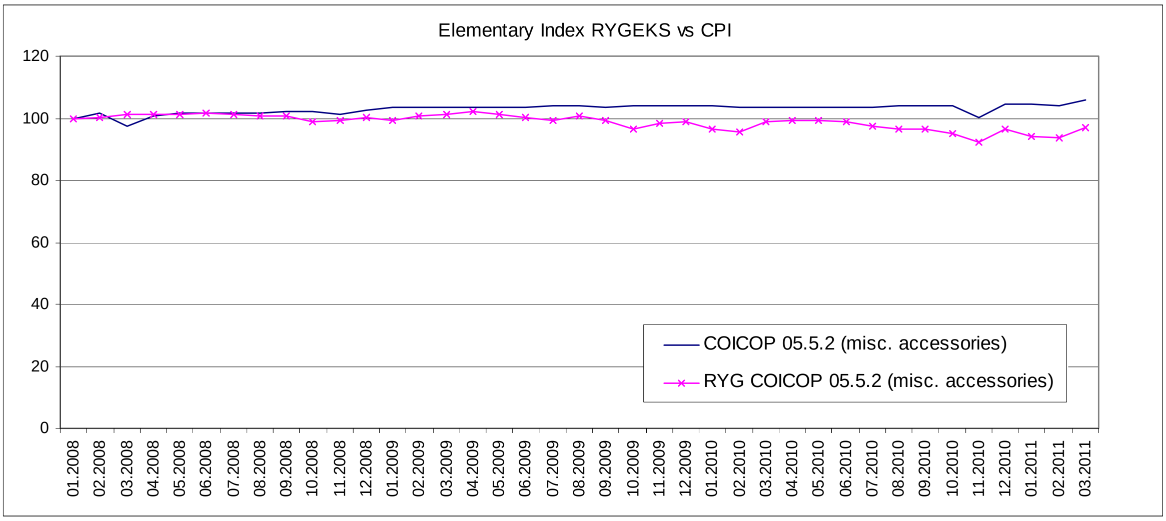
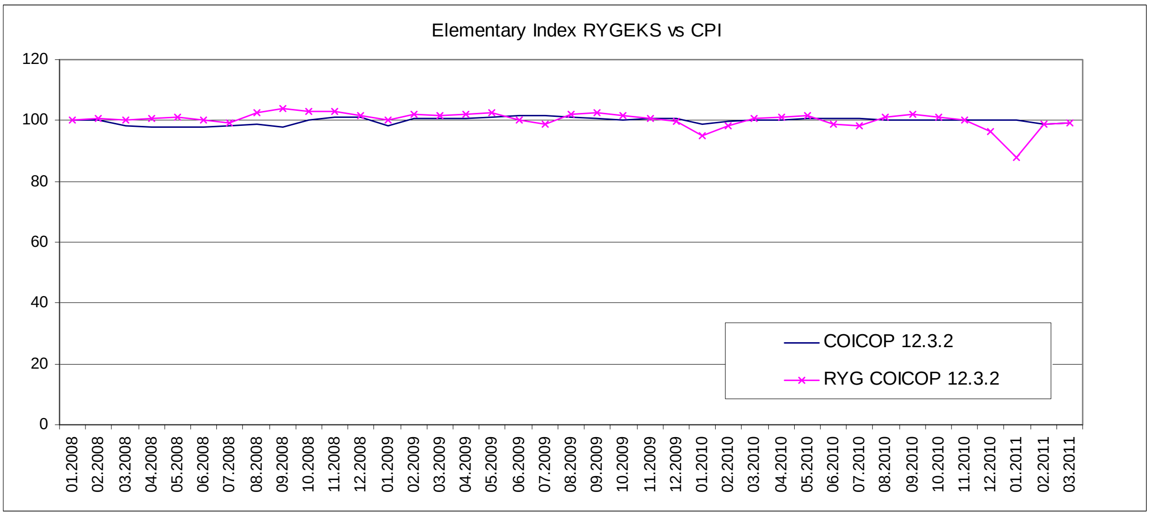


Figure 4

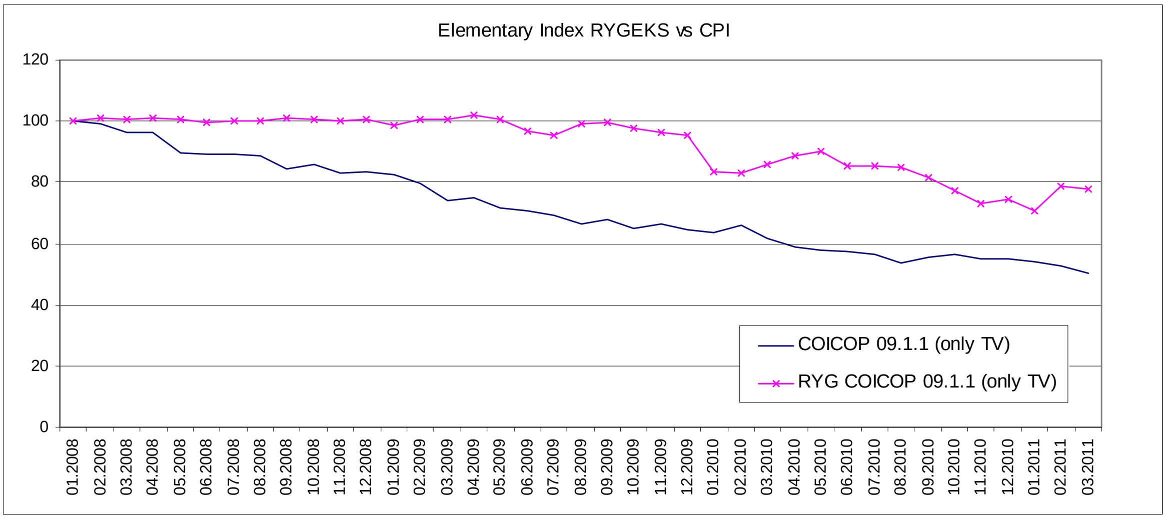
Relevé traditionnel vs RYGEKS : COICOP 12.3.2 (articles de voyage)



Le graphique suivant présente un contre-exemple où l'indice RYGEKS s'éloigne significativement de l'indice produit avec les prix provenant du relevé traditionnel sur le terrain. Il est à noter que les prix relevés de manière traditionnelle pour ce type de biens font l'objet de traitements spécifiques d'ajustement de qualité tandis que la méthode RYGEKS ignore ce genre de procédures.

Figure 5

Relevé traditionnel vs RYGEKS : COICOP 09.1.1 (TV)



L'exemple suivant présente la comparaison entre l'indice produit avec un relevé traditionnel basé sur les données scannées et un indice RYGEKS pour la chaîne de distribution X pour le groupe de l'alimentation et des boissons non alcoolisées (COICOP 01).

Figure 6

Relevé traditionnel vs RYGEKS : chaîne de distribution X, COICOP 01

Annexe 2

Formalisation de l'algorithme RYGEKS implémenté à l'OFS

Pour les calculs, l'algorithme décrit par Ivancic et al. (2009) a été suivi :

Matrice d'indices bilatéraux

Indice de Fisher [F(i/j)][3]

où

$$F(i/j) = \sqrt{L(i/j) \cdot P(i/j)}$$

Il s'agit de la variation de prix de la période i par rapport à la période de base j, où L représente un indice Laspeyres et P un indice Paasche. L'indice Laspeyres tient compte des parts de vente de la période de base, l'indice de Paasche des parts de vente la période finale.

$$L(i/j) = \sum_k w_{ki} \left(\frac{p_{ki}}{p_{kj}} \right) \qquad P(i/j) = \left(\sum_k w_{kj} \left(\frac{p_{kj}}{p_{ki}} \right)^{-1} \right)$$

où p_{ki} est le prix de l'article k à la période i et w_{ki} la part des ventes de l'article k lors de la période i (pour l'ensemble des articles d'une position).

Transformation RYGEKS

La transformation selon GEKS produit un indice transitif :

Calcul GEKS pour les premières 13 périodes :

$$\begin{pmatrix} F(1/1) & F(1/2) & ... & F(1/13) \\ F(2/1) & & & \\ .. & & & \\ F(13/1) & & & \end{pmatrix} \begin{matrix} \rightarrow ps(1)(1) = MoyGéom\{F(1/1), F(1/2), \ldots, F(1/13)\} \\ \vdots \\ \rightarrow ps(1)(13) = MoyGéom\{F(13/1), \ldots, F(13/13)\} \end{matrix}$$

ainsi,

$$ps(1)(t) = MoyGéom\{F(t/1), F(t/2), \ldots, (F(t/13)\}$$

$$\text{RYGEKS(t)} = \text{ps(1)(t)/ps(1)(1)} \quad (\text{ t} = 1, \ldots, 13)$$

Pour t > 13 :

$$\text{RYGEKS(t)} = \text{RYGEKS(t - 1)} \cdot \text{ps(t - 12)(t)/ps(t - 12)(t - 1)}$$

où

[3] La formalisation est présentée ici avec l'indice Fisher. Le code implémenté permet également l'utilisation d'un indice Törnqvist.

$$ps(t-12)(t) \quad = MoyGéom\{F(t/t-12), F(t/t-11), \ldots, F(t/t)\}$$
$$ps(t-12)(t-1) = MoyGéom\{F(t-1/t-12), F(t-1/t-11), \ldots, F(t-1/t)\}$$