

Multipurpose Price Statistics

Roberto Barcellan*

**European Commission, Eurostat*

Ottawa Group

Copenhagen -1-3 May 2013

Abstract

Price statistics is a fundamental and critical statistical area that provides key information to policy makers and citizens. The development of price statistics in recent years has been characterised by an increasing demand for information on different dimensions of prices: the focus, while continuing to be on the evolution of prices over time (covered by Consumer Price Indices) and on their spatial comparisons (covered by Purchasing Power Parities), has progressively moved to a combined view of both temporal and spatial dimensions of price statistics. References to detailed average price levels become more and more usual among price statisticians. At the same time, the evolution of the production process for the compilation of price statistics has been profoundly influenced by technological innovations, emerging of new sources for data collection leading to the availability of huge amounts of basic price information.

Hence, it follows that in the coming years the development of price statistics needs to converge to the integration of temporal and spatial price information to generate synergies and to fully profit from this vast amount of already available and potentially available basic price information. The modernisation of price statistics will then rely on the concept of "multipurpose price statistics".

The evolution towards multipurpose price statistics generates several methodological and practical challenges. On the basis of the experience cumulated at European level and of the related analysis, this paper defines the concept of multipurpose price statistics and explores the associated issues, their role and influence in the production of official price statistics and presents the strategic approach initiated at European level to set up an integrated system of price statistics.

Key words: multipurpose price statistics, CPI, PPP, statistical production process.

Acknowledgements: this paper is largely based on the Eurostat Consumer Prices Research studies "Experimental analysis into the measurement of indicative price levels for consumer products", 1st-4th pilots, published between February 2009 and December 2012 by Eurostat (author Jan Walschots). It also has its roots in the proceedings of the Eurostat Conference "Reviewing the Business Architecture of Consumer Price Statistics", held on 15-16 October 2009 in Luxembourg. Special thanks to the colleagues in Eurostat working on multipurpose statistics for their input to the paper.

1. Introduction

It is often argued that the recent shift towards evidence-based policy and the sector and product market monitoring exercise call for a new approach to price statistics that can be described as a "structural approach to prices". In this sense, In Europe, Eurostat and the National Statistical Offices (NSOs) are investigating possibilities to extend the system of price statistics to modernise current data collection practices. Attention is concentrated in combining temporal and spatial dimensions of consumer price collection and index compilation. This would be an extension to existing work on price statistics that includes the Harmonised Index of Consumer Prices (HICP) and Purchasing Power Parities (PPP) by emphasising the synergies between these two statistical areas. Both HICPs and PPPs are compiled as index numbers. Those two statistics show, respectively, aggregate data on the development of price levels over time and comparisons of price levels across countries. The new approach, referred to as "multipurpose price statistics", aims to present indicative price levels for specific groups of products as an additional set of price data. It is by no means a replacement for the measurement of inflation (HICP) or international price level comparisons (PPPs).

The approach to multipurpose price statistics has been developed on the basis of pilot projects for the collection of detailed prices aiming at producing detailed average prices for a well identified list of products. The pilot projects run since 2009 provided experimental results and tried to explain the strengths and limitations of the results produced. It must be stressed that the results provided in the reports regularly published by Eurostat in relation to the detailed price level pilot projects do not necessarily represent national average prices – they are indicative price levels for those product groups that consumers usually or typically buy nationally.

Multipurpose consumer price statistics aim to develop further existing price index statistics by increasing the level of detail of the collected and released information. They also aim to produce statistics on detailed average price levels to supplement the existing price index statistics. Furthermore they promote research on the modernisation of price collection processes with the aim to make best use of collected data for various statistics and to promote further harmonisation of the various price statistics.

More detailed price statistics and detailed price level data are needed for several purposes, including for product market and sector monitoring in order to contribute to the improved monitoring of the euro area Single Market policies, the understanding of how markets work and their potential value for developing HICP communications. More detailed price statistics are also useful for developments such as the HICP at constant tax rates and HICP administered prices. In addition, detailed average prices are of interest on their own and are already released to the general public in the 'Consumer Markets Scoreboard' published by the European Commission (DG SANCO).

The development of a new system of price statistics and a coherent set of publications on prices is a complex project. Before reaching a final and stable production process that provides price levels data at the detailed product level, research projects are needed to investigate experimentally the qualities, advantages and disadvantages, of alternative approaches. This approach requires also to define the appropriate methodological and practical framework for the regular compilation of this "new" price statistic figures.

Defining and computing the price level of a product group involves a number of methodological requirements that are different from the methodological requirements of HICP and PPP calculations.

Therefore HICP and PPP data collection are not immediately and fully usable for making a publication of detailed average price levels. The novelty in the pilot projects and analysis run by Eurostat and the European Statistical System (ESS) is, however, the tentative to identify and consider appropriate re-use of selected existing price data in order to limit costs and response burden.

The basic idea is to set up the best strategy for collecting detailed price data starting from the traditional HICP collection at different points of time and comparing them with PPP-data. Any adjustments to the price observation process should however be subject to the condition that the continuing compilation of the HICPs and PPPs are not negatively affected.

The research in this area is only at the beginning, nevertheless results are encouraging as well as the interest they generate. Eurostat and the European Union NSOs are currently running additional research studies for the period 2011-2014 on the use of HICP, PPP and other alternative data sources to develop multipurpose consumer price statistics. This shall lead to decisions for the medium-term on how to make available detailed data on price indices and indicative average price levels.

This paper explores some methodological aspects related to multipurpose price statistics, takes into account the results of the pilot projects underlying this experiment and aims to test whether and to what extent the proposed approach is a feasible way to reach comparable and usable results, to find out what are the difficulties in practice and eventually to consider possible adjustments to the process to improve results in the future.

Section 2 is devoted to the brief presentation of the HICP and PPP frameworks. Section 3 introduces the basic terminology associated to multipurpose price statistics; section 4 highlights some methodological and practical challenges; section 5 introduces the concept of price statistics data warehouse; section 6 explores the results of the pilots projects and section 7 points to the future work in this area.

2. Methods

The main idea of multipurpose price statistics is to enhance the traditional collection of basic price data for the compilation of (harmonised) consumer price indices by combining it with key characteristics of the PPP basic price collection. In practice, a set of consumer product groups is identified on the basis of a reduced version of the PPP standard definitions and basic prices, complemented by related information, are collected following a HICP oriented approach.

2.1. HICP and PPP data

The **HICP** is a monthly price (inflation) measure covering the whole of each European Union (EU) countries; Croatia; Iceland; Norway; Switzerland; and Turkey. Using HICP data has the advantage that there already exist a great number of monthly price observations. The HICP do not need to measure prices for each and every product in each and every outlet. It suffices if a sample of product offers is taken in each elementary product group that is representative of the price development in that product one or more consumption segments. The product descriptions used for the HICP allow for some flexibility in order that the HICPs can best represent price developments in national markets and to allow price collectors to find in each selected outlet a product that is relevant in that outlet and according to the given product description. Price collectors will, in general, select the same product-offer each month in order to optimize price comparisons over time.

At the very detailed level, HICP data comparability across countries may be rather limited. Firstly, because markets are different, there will be many differences between the product-offers that are selected from within each elementary product group mainly due to differences in the sampling approaches adopted by the NSIs. Secondly, product definitions in the Member States may be a mix of tight and loose descriptions which can mean that the NSIs may collect different mixes of products. On this point however it should be noted that comparability across countries is not only an issue for HICPs – it can also be so for PPPs. Thirdly, the products for which prices are observed, may not necessarily be market leaders or most representative products in the market (the HICP should after all follow prices not only for market leaders).

Moreover the comparability over time of price levels will be more limited than that of price indices. Resampling and replacements will result in the observation of differing products over time and there may be differences in actual outlet distribution in the successive samples.

On the other hand, **PPPs** are based on a selection of products that are rather narrowly defined. Using PPP data would in principle give more comparability of product specifications across countries, and that data is based on some knowledge or judgment concerning the representativeness of the products in each Member State. Products according to the same strict product description are observed and their prices compared across countries. However, in many cases these products are available and observed only in part of the countries and only contribute to the price level comparisons among these countries. Within the PPP survey, at EU level, prices for one specific product are only collected once per three years. Supplying PPP data more frequently would involve considerable additional costs. Moreover PPP data at present is limited to capital cities. Finally the number of price quotes collected per product for the PPP exercise is low in practice and this means that, at the most detailed level, the existing PPP data set has not been considered sufficiently reliable for publication.

2.2. HICP-PPP mix: the ideal approach

In October 2009, the Eurostat Conference on "Reviewing the Business Architecture of Consumer Price Statistics" concluded that the solution to meet the increasing requirements of detailed average price information had to be found within the existing constraints of the price statistics production process. As such, the only way forward was modernising, improving the efficiency and trying to get something in common between PPP and HICP. The proposal of the Conference was to modernise the production system by bypassing the strict separation between the existing PPP and HICP collection methods and looking at integrating the two approaches in the best and modern way of a data warehouse approach.

The data warehouse approach combines the "detailed" and harmonised product information of the PPP collection with the "temporal richness" of price observation of HICP/CPI collection. Hence, the basic idea of multipurpose price statistics is to collect price observations and the necessary complementary information (product characteristics) and stock them in a data warehouse structure. The corresponding repository of basic prices and related information should then be used for the compilation of the relevant multipurpose price statistics (above all detailed average price levels). The system qualifies as a sort of extended and higher frequency PPP collection or a HICP well-identified-specific-list based collection.

Such an approach raises methodological and practical concerns.

3. Terminology

Before exploring methodological and practical considerations related to multipurpose price statistics, it is important to define the key terms pertinent to this area.

Multipurpose price statistics

Basic price information (basic prices and related information) collected for different price statistics purposes.

Detailed Average Price Levels

Indicative price levels are:

- averages of prices measured in the outlets selected for price observations,
- for the product offers as chosen by the price collectors,
- selected following the procedures of the statistical institute in each country,
- in many cases, but not always, for the most popular product types,
- often, but not always, prices at which the products are sold in the most popular outlets.

The prices collected may be affected by special offers and seasonal sales, and to a degree that may differ across countries.

Therefore, these prices do not necessarily represent national average prices. They show indicative price levels within the scope of the definitions used, for specific products that consumers usually or typically buy nationally and that are therefore used to calculate the national HICPs.

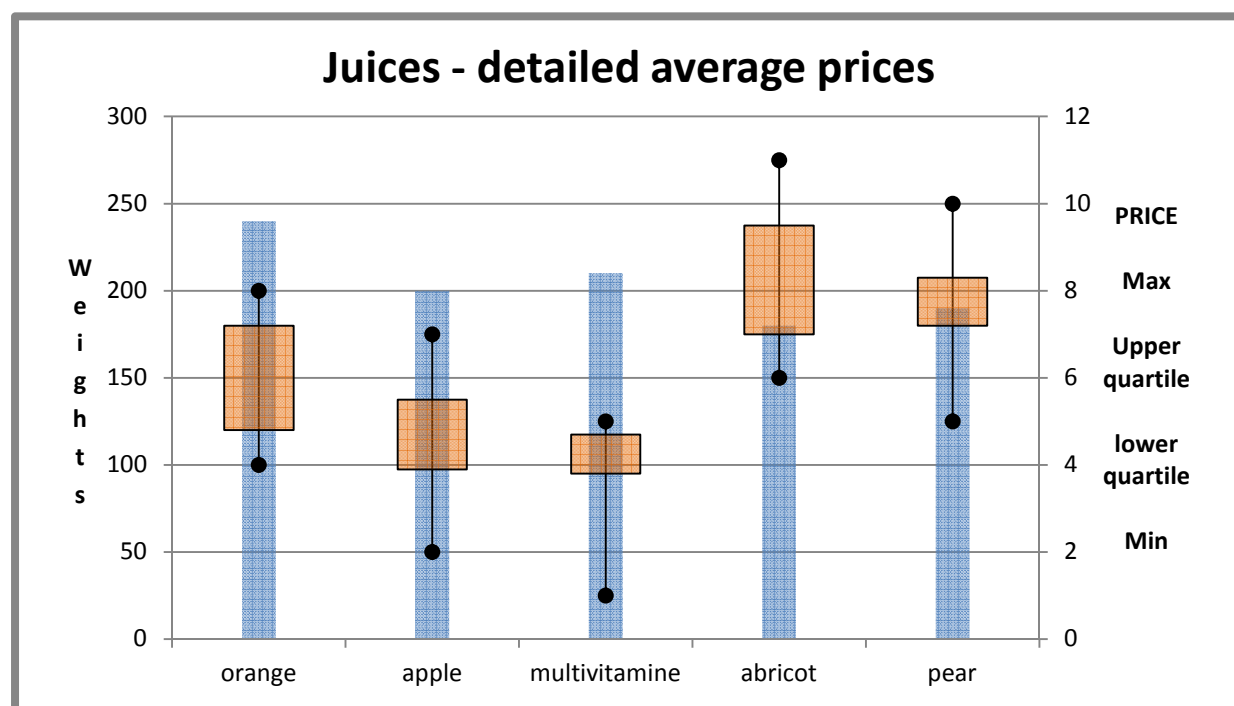
The prices observed are those of 'product-offers'. This means that once a product and outlet sample was drawn in line with HICP legal framework and sampling rules in the EU Member State the price observer collects the price at which a product is offered in an outlet. Short term consumer responses to changing market circumstances do not influence the resulting average price. If a special offer in one outlet induces a large number of consumers to buy the product there at a low price this will affect the average price paid by consumers but not the average of the price offers collected.

Detailed price levels are in many cases computed using an unweighted arithmetic average of the price observations. In some countries regional price observations were weighed using e.g. regional population figures.

Price level dispersion

In assessing indicative price levels, it is important to associate them with a measure of dispersion. Following well established statistical practices, the average price levels can be supplemented by other statistical measures such as the median price level as well as by quartile dispersion measures (combined with an absolute dispersion measures based on minimum and maximum prices) - See Fig. 1.

Fig 1: Detailed average prices – an example: juices



4. Methodological and practical challenges

The main methodological challenge in multipurpose price statistics is to combine the spatial (PPP) and temporal (HICP) dimensions.

The design of price collection in the HICP relies on the purpose of the HICP itself: a measure of inflation. Therefore it does require the entire methodological framework proper to consumer price indices compilation going from sampling aspects to representativeness, quality adjustments, etc. On the other side, the spatial elements is caught by the PPP oriented approach that focuses on the comparability of products across countries, usually enacted by a very precise description of the products for which basic price information has to be collected.

Multipurpose price statistics require then to design the collection of basic information to match the PPP approach in a less detailed way (selected list of group of products – no forced adherence to a very restrictive description of the product) and to combine it with the HICP methodological and practical collection instruments (sampling, geographical coverage, quality adjustment, frequency).

The outcome is a set of information on a well identified group of products, collected at high frequency, over the entire geographical territory and representative of the consumer expenditure habits. This information is ideally mapped in a data warehouse to be used to derive the relevant price statistics.

The main aim of multipurpose price statistics is to provide the information for the derivation of different price statistics, mainly detailed average price levels, but potentially also PPP more frequent information and qualified HICP-product movements (dedicated product indices).

In this respect, comparability aspects across countries and over time are quite relevant.

4.1. Comparability across countries

There are several factors that may influence price differences and limit the comparability of the results. These factors may be differences between the market circumstances in the various countries and differences resulting from the statistical process.

- The aim of the product selection is to find products which are available in possibly all EU Member States, are of a comparable quality, and that may be representative for certain consumption segments. However, one product cannot represent the full market in all countries for the consumption segments involved. This means that the average price of e.g. a loaf of white bread may not be indicative of the price level of bread in general in a given country.
- Furthermore, even if products are identical across countries, they may not have the same relevance for the consumers in the various countries. For example, a product may be a market-leader in one country, and its price very representative, while at the same time being a niche product in another country, where its price could be much less representative.
- The positions of the selected products on national markets may differ across countries even within the limits of the general product description used. For example there may be differences in market situations of 'table wines' between wine producing countries and countries that do not produce wine themselves.
- Product descriptions used in national HICPs may vary, representing different parts of the range of products as allowed by the general product description. If the product description in a country is tighter than the general product description, the prices collected will be a selective part of all the products, and therefore the average price may reflect this. Likewise if in a country a very wide product description is used the average price may refer to a larger range of products than in another country.
- For some products the focus on price observations may be on most sold varieties, whereas for other products the focus on price observation may be on the higher end or lower end of the market.
- The outlet structure of the product market in a country may be different. Prices charged in supermarkets may be different from prices in specialist stores or open markets. Price differences between countries for ice-cream may relate to differences in market structure (locally produced or sold in supermarkets).
- The outlet distribution in the samples of the product may be different. If prices for a product were observed mainly in supermarkets average prices may be different from prices observed in specialist stores or open markets. The outlet distribution in the samples at the individual product level need not be representative for the distribution of outlets where the product is actually sold.
- For some types of products, HICP samples may cover only price for some types of brands and exclude some other brand levels. If the selection of brand levels for the same products differs across Member States, this will influence comparability.

Quality differences

Even within the boundaries of the product descriptions, quality differences may exist to a varying degree. These may reflect differences between the national markets, when consumers in one country may tend to buy higher quality products and in another country they are more interested in

lower prices for lower quality products. They may also occur in the price collecting process when product descriptions differ across countries.

An example of a product where these differences seem to be important and further research may be needed is clothing and footwear, for which it is very difficult to ensure that prices for products of comparable quality are measured across countries or even are available on the market. But there are other product groups for which samples would need to be further harmonised to further improve comparability, for example orange juice which may be based on concentrate, sterilized, etc.

4.2. Comparability over time

When comparing the indicative price levels results across different time periods they should be interpreted with great care. The differences may come from various reasons other than price changes as measured by HICP-inflation:

- Inflation is always measured by comparing prices in national currencies. The indicative price levels used for the pilot projects on detailed average price levels have all been translated into euro price levels to make possible cross-country comparisons. Changes in currency exchange rates will directly affect comparability over time for the countries that are not members of the euro area.
- Products may be different in various periods. HICP samples are regularly updated in order to follow market developments and to keep the samples representative. HICP production processes aim to eliminate any unwanted index changes caused by product replacements from the inflation figures. These differences in products due to replacements will not always be clear from the product descriptions used in the pilot studies on indicative price levels. HICP price developments are corrected for quality changes. It is not possible to assess what the impact of quality change may have been on the development of indicative price levels.
- Differences across time in the distribution of outlets where prices have been collected may influence the differences in the indicative price levels. Changes in the outlet sample are in many cases dealt with so as to have no direct effect on HICP inflation results.
- Price indices generally represent the average price developments for several representative products within the same COICOP-group. Therefore even the most detailed price indices are generally based on more price observations per month and may be expected to be more reliable than indicative price levels.

5. Towards a multipurpose price statistics data warehouse

From a practical point of view, the multipurpose price statistics approach leads towards the creation of a data warehouse for price statistics, provided that the collection of basic price information is structured and harmonises and mapped to a common data model for price statistics. The combined collection of basic prices and related information (product characteristics) put the data warehouse approach in an intermediate situation with respect to HICP and PPP collections. Such an approach is potentially amplified by the technological advancement in price collection, such as:

- Use of scanner data, with the possibility to collect a huge amount of detailed information on product prices, product characteristics, product volume of sales. Such information should be mapped at broader product level to a structured data model to feed a price statistics data warehouse.

- Use of electronic devices for the traditional HICP collection combined with the contemporaneous collection of additional information on product characteristics in a structure (harmonised) way and in connection with a structured data model to feed a price statistics data warehouse.
- Use of internet price collection for prices and related information according to a data model to be targeted to feed a data warehouse.
- Combination of detailed microdata price information (as collected according to the three previous approaches) with PPP methodology ensuring an increased geographical and temporal coverage.

6. Results of the pilot projects

The report on the Eurostat Consumer Prices Research studies, fourth pilot, as published in December 2012, measures indicative price levels from HICP price collection in June 2011. Prices for 156 products have been published (see Table 1). A large part of the prices for the identified food products and for alcohol and tobacco could be provided by Member States. In other COICOP groups a larger part of the results was rejected.

Some initial conclusions following this research are as follows:

- Given that the regular PPP data set has not been considered sufficiently reliable for publication, the pilot project investigated only one possible way to compile price level data. Further research on the use of HICP, PPP and other alternative data sources is in progress and will be needed to be able to make a well-founded choice for the medium term on the most effective way to make available detailed data on price levels.
- If it is possible to compile price level data from the HICP price collection this has the advantage of high frequency combined with a large number of prices and national coverage.

Table 1: Fourth pilot study on detailed average price levels – Eurostat, December 2012

Number of products entered in the project, and results, by COICOP division			
	Requested	In publication	Rejected
1 Food and non-alcoholic beverages	52	52	
2 Alcoholic beverages and tobacco	5	5	
3 Clothing and footwear	29	28	1
4 Housing, water, electricity, gas and other fuels	3	3	
5 Furnishing, household equipment and routine maintenance of the house	23	23	
6 Health	0	0	
7 Transports	2	2	
8 Communications	0	0	
9 Recreation and culture	23	23	
10 Education	0	0	
11 Restaurants and hotels	5	5	
12 Miscellaneous goods and services	15	15	
Total	157	156	1

- There seem to be large differences between the samples of goods and services for which prices are collected in the HICPs of Member States. Despite the fact that for this research the product descriptions used were looser than those used in the PPP exercise, many countries could not provide indicative prices for some of the requested goods and services.
- For some products, it should be assessed whether differences in product descriptions between countries could be reduced or removed in the longer term by harmonizing product descriptions without reducing the representativeness of the samples. Further research on this would be needed. Differences between national HICP samples can be explained to a large extent by the fact that national sample must be representative of the national market situation. The variety of products available for consumers is so large that there are many alternatives that NSOs could choose from for price collection while still remaining representative.
- The process of price collection as developed by the various NSOs differs substantially across countries. National statistical offices may ask their price collectors to collect prices using rather tight product descriptions, or provide very loose product descriptions and ask the price collector to choose a representative product in the outlet. More harmonised procedures and, above all, product descriptions to be used in HICP price collection are a necessary condition for achieving usable results from the multipurpose price statistics point of view.
- Further research should be done on the optimum level of specification in the product descriptions. More detail in the product definition would increase the comparability of the products across countries as well the precision of the average prices, but on the other hand would lead to a smaller part of the markets being covered and less prices (because more products would be unavailable in more countries).
- In price statistics it is generally accepted that the coefficient of variation (CV) for the observed prices of a product should not be larger than 20 per cent, otherwise we can suspect that the product specification is too broad. For the price level data in the fourth pilot, only 40 per cent of the average prices show a CV below 20 per cent. This indicates that efforts are needed to further narrow the product specifications, without reducing the coverage of the markets.
- If we look at confidence intervals, around 80 per cent of the average prices show a low and hence acceptable width of the confidence interval of less than 10 per cent of the estimated average. Further investigations should look at the price distributions for certain products. If, for example, the distribution of prices included in one average price shows multiple peaks it may indicate that different, possibly incomparable, products were included.
- Tentative results show that the confidence intervals are approximately the same, whether the calculations are based on the limited set of prices for so called well-known brands or based on a full set for all brands. Further research into the reasons for these findings would be needed.
- Furthermore in the period 2013-2014 research will be carried out into the modernisation of price collection processes with the aim to make best use of collected data for various statistics and into further harmonization of the various price statistics. As part of this, Eurostat and some NSOs are undertaking research studies on the exploitation of point-of-sale or scanner data for use in the HICP and/or PPPs and which can benefit the further development of detailed average price level statistics.

7. Future work

The multipurpose price statistics approach is an attempt to find the matching point between PPP and HICP collection methods and practices to take strengths from both systems for the production of price statistics in broader sense, without replacing the compilation of HICP and PPP themselves.

The use of detailed price indices might be a step forward in this direction:

- A concrete starting point would be to develop a common classification for PPP and HICP (e.g. a common list of a certain number of products – 300 – defined at, for example, at COICOP six digit level). This approach has been followed in setting up the pilot projects coordinated by Eurostat at ESS level; in addition, Eurostat recently defined a common COICOP 5 digit classification at European level.
- A second step would be the development of a common approach for comparability and representativeness for both PPP and HICP to increase the reliability and consistency of both collection systems. In practice, this approach could correspond to the set up of PPP measures and HICP indices based on the restricted selected number of detailed products. Methodological and practical considerations on this subject have to be further analysed.
- Furthermore, there could be a common production process in part for both PPP and HICP, for a limited set of products. NSOs could move towards a better described and harmonised production process. Full information could be made available centrally by metadata. The option of using scanner data, internet price collection and technologically advanced price collection in the shops should be further explored; in particular, it should be identified whether scanner data can be used to improve the productivity and efficiency of the process.
- The data needs of today and tomorrow should be discussed with users of detailed average price data.
- The potential success of multipurpose price statistics relies on the creation of a macro data warehouse on aggregate level which could include sectoral data and on the creation of a microdata warehouse in order to enable users to improve their potential uses. The data warehouse approach would offer a structured approach to the detailed price information that would combine the traditional HICP price observation with the product details typical for PPP collection. The flexibility of the data warehouse would allow re-structuring the basic information according to the incoming needs and fine tune the detailed average price level compilation to match users' needs. The project PRIX (PRIce statistics neXt generation) recently launched by Eurostat at ESS level moves towards this targets.

References

European Commission, DG SANCO, 2012, Consumer Markets Scoreboard, December 2012.

Eurostat, 2009-2012, Consumer Prices Research studies "Experimental analysis into the measurement of indicative price levels for consumer products", 1st-4th pilots

http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/prices_data_for_market_monitoring

Eurostat, 2009, Proceedings of the Conference "Reviewing the Business Architecture of Consumer Price Statistics", Luxembourg, October 2009.