

**Economic and Social Council**

Distr.: General

7 July 2016

Original: English

---

**Economic Commission for Europe**

Conference of European Statisticians

**Group of Experts on Consumer Price Indices****Thirteenth session**

Geneva, 2 – 4 May 2016

**Report****Note by the secretariat***Summary*

The present document is the report of the meeting of the Group of Experts on Consumer Price Indices, 2-4 May 2016, and is provided to inform the Conference of European Statisticians of the organization and outcomes of the meeting.

The Conference of European Statisticians at its plenary session in April 2016 approved the UNECE Statistical Programme for 2016 (Report of the Conference of European Statisticians, Sixty-fourth plenary session, Paris, 27-29 April; document ECE/CES/91). The Statistical Programme for 2016 includes the list of meetings planned to be organized from January 2016 onwards. The list includes a meeting of the Group of Experts on Consumer Price Indices to be organized jointly with the International Labour Organization and held in Geneva in 2016 (document ECE/CES/2016/12, Annex I).

## I. Introduction

1. The Group of Experts on Consumer Price Indices (CPI) meeting was held in Geneva on 2–4 May 2014. The meeting was organised jointly by UNECE and the International Labour Organisation (ILO). It was attended by Australia, Azerbaijan, Belarus, Belgium, Botswana, Burundi, Canada, China, Colombia, Croatia, Denmark, Estonia, Finland, France, Georgia, Germany, Hungary, Iceland, Ireland, Italy, Japan, Kazakhstan, Kuwait, Kyrgyzstan, Latvia, Luxembourg, Mexico, Mongolia, Morocco, Netherlands, Norway, Philippines, Poland, Republic of Moldova, Russian Federation, Saudi Arabia, Senegal, Singapore, Slovenia, South Africa, State of Palestine, Switzerland, Thailand, Togo, Tunisia, Uganda, Ukraine, United Kingdom, United States of America, Viet Nam. Representatives of the European Central Bank (ECB), European Commission and Eurostat also attended the meeting. The following specialised agencies and international organisations attended: Economic and Social Commission for Western Asia (UNESCWA), Economic Commission for Latin America and the Caribbean (UNECLAC), Economic Commission for Africa (UNECA), ILO, The World Bank, Eurasian Economic Commission (EEC), International Monetary Fund (IMF), Interstate Statistical Committee of the Commonwealth of Independent States (CIS-STAT), Organisation for Economic Co-operation and Development (OECD). Prof. Jess Diamond, Hitotsubashi University, Japan, Dr. Alberto Cavallo, Massachusetts Institute of Technology, Mr. Satoshi Imai, Nowcast Inc., Mr. Walter Diewert, University of British Columbia, Prof. Tsutomu Watanabe, University of Tokyo and Mr. David Fenwick attended the meeting as invited experts.

2. The meeting was chaired by Marcel van Kints, Australian Bureau of Statistics. Jan de Haan, Statistics Netherlands; Antonino Virgillitto, Italian National Institute of Statistics; Patrick Kelly, Statistics South Africa; Mathieu Lequain, Statistics Canada; James Tucker, Office for National Statistics, United Kingdom; and Nada Hamadeh, World Bank, acted as session chairs.

## II. Organization of the meeting

3. The following topics were discussed at the meeting on the basis of the provided papers and presentations:

- (a) Scanner data
- (b) Big data
- (c) CPIs for different population groups, income groups and geographic areas.
- (d) Poster session
- (e) Methodological issues
- (f) House price indices
- (g) Synergies between CPIs and PPPs and integration of survey activities
- (h) Update of the 2004 CPI Manual
- (i) Future work

### **III. Summary of discussions and main conclusions reached at the meeting**

#### **A. Session 1: Scanner data**

Session chair: Jan de Haan, Statistics Netherlands

3. The session on scanner data was based on papers by Denmark, Island, Italy and Japan. Room documents were provided by the Netherlands and Eurostat.

4. A growing number of countries apply scanner data for the production of the CPI. Scanner data can complement and in some cases replace traditional methods for collection of price data through surveys or price collectors, and be used for the calculation of CPI for particular products or groups of products. Scanner data can also be used for sampling purposes and expenditure weights estimation. The use of scanner data raises a number of practical, methodological and conceptual issues that need further research. The following main points and conclusions were raised during the session:

(a) The information contained in scanner data sets is usually found “correct”. Nevertheless, the data should be carefully tested as various countries did come across examples of prices (unit values) with implausible values. Testing of a new processing system, if necessary, is also needed and may take quite some time.

(b) The choice of product identifier (Global Trade Item Number (GTIN), Stock Keeping Unit (SKU), or model number) can significantly affect the results and needs consideration.

(c) When estimating price indices on the whole population of items in scanner data rather than on relatively small samples, a multilateral method seems preferable in order to avoid chain drift. However, more research is required on the choice of index number method and the method to circumvent revisions of previously estimated figures.

(d) Where national statistical offices do not wish to rely on scanner data for the construction of their CPI, scanner data can still inform them on issues such as item sampling, the extent of product substitution, weighting and the nature of promotional sales.

(e) The use of scanner data has the potential of further improving the quality of CPI statistics and making the production more cost efficient (replacing traditional price surveys) as well as reducing response burden. It will be useful to further exchange good practices on how to implement scanner data in the regular production of the CPI.

#### **B. Session 2: Big data**

Session chair: Antonino Virgillitto, Italian National Institute of Statistics

5. The session was based on papers by Italy, Japan, Netherlands, Norway, United Kingdom and Massachusetts Institute of Technology, USA.

6. The use of Big Data has attracted significant interests from national statistical offices (NSOs), while the CPI is one of the few areas of official statistics where practical experiences have been gained. Use of Big Data has large potentials for CPI compilation in terms of real time access to a large and growing amount of information, but also raises methodological and technical challenges as well as legal and institutional issues (e.g. ensuring timely access to the data). The following main points and conclusions were raised during the session:

(a) Web scraping techniques for supporting and complementing traditional CPI compilation have been investigated in NSOs in recent years with the objective of exploiting Internet as a data source to reduce collection costs and increase the frequency and number of products that are represented in the basket. These experiments are starting to produce results in terms of compiling price indices based exclusively on scraped data.

(b) Experiments with different techniques for web scraping demonstrate two main approaches: i) Using tools (robots) that reproduce and automate manual steps by collecting data from Internet. ii) Implementing a specific *parser* (a software programme) for each retailer web site that extracts structured price data from an unstructured web page. The key common challenge reported is the high maintenance costs in the training of robots and the development of parsers, which makes scaling up to full “Big Data” size difficult. Therefore, the experiences in NSOs are currently restricted to a manageable number of web sites.

(c) An extensive comparison between online and offline prices, based on a large-scale, multinational collection of online price data, showed that they are mostly similar, yet significant differences are present across countries, sectors and retailers. The conclusion is that currently the web can be safely used as an effective data-collection tool for some groups of products, such as apparel and electronics.

(d) Recording and classifying price observations correctly is a key common challenge identified by NSOs when using web scraping. Supervised machine learning methods are being experimented for identifying the correct COICOP category for each scraped item from its generic text description found on the web site.

(e) The use of Internet as a data source also gives insights over other aspects of price behaviour, such as the pricing pattern over the product life-cycle and quality changes associated with product replacements. Data seem to support the so-called 50% rule by which half of a price increase is attributed to an increase in quality.

(f) Web scraping and scanner data are two aspects of Big Data that can be combined within the regular production process of compiling CPI. Advanced IT architectures and Big Data technologies can enhance the capabilities of NSOs of collecting, storing and analysing massive quantities of price data.

### **C. Session 3: CPI’s for different population groups, income groups and geographic areas.**

Session chair: Patrick Kelly, Statistics South Africa

7. The session was based on papers by Canada, Germany, Japan, Uganda and USA. Room documents were provided by Colombia and New Zealand.

8. The session comprised papers on both the practical design of CPIs for different groups of the population, as well as their use in understanding the specific inflation experience of different groups.

9. The following main points and conclusions were raised during the session:

(a) Price indices for different socio-economic groups can be used to illustrate variations in changes in the cost of living. Extensive user engagement is helpful to identify the most relevant group indices.

(b) If indices for different groups are to be published, these should be carefully designed. Many countries publish indices for specific geographic areas. Possible sources of

weighting data for area specific CPIs would include population censuses, household budget surveys as well as administrative sources.

(c) Specific indices can be published for different income groups. The household expenditure survey, on which CPI weights are usually based, can be used for deriving CPI weights for different income groups. The accuracy of these types of indices can be improved by identifying and pricing the products consumed by different income groups, rather than only adjusting the weightings for product groups.

(d) CPIs for different groups shed light on the variation in the experience and perceptions of inflation. Analysis of these relationships can be an aid to the management of monetary policy.

(e) CPIs for different groups of the population need to be communicated and explained well to users to ensure correct interpretation and avoiding misunderstanding and misuse of the CPI.

#### **D. Poster session**

10. The session was based on the following presentations:

- (a) The Billion Prices Project. Using Online Data for Measurement and Research. MIT and NBER, USA
- (b) Commercial Property Price Indicators Sources, methods and Issues. Eurostat
- (c) CPI Major Review. INEGI, Mexico
- (d) Modernising official statistics. UNECE
- (e) Selection of CPI calculation formula in three steps. UNECE

#### **E. Session 4: Methodological issues**

Session Chair: Mathieu Lequain, Statistics Canada

11. The session was based on papers by Japan, USA and IMF. Room documents were provided by Belarus, Burundi and Mexico.

12. The following main points and conclusions were raised during the session:

(a) The most common way for calculation of higher-level price indices is to apply a fixed basket approach where indices are calculated as the expenditure weighted arithmetic average of the elementary indices. Higher-level price indices may be calculated by use of direct (long-term) or chained (short-term) elementary price indices. The two methods, if applied correctly, give identical results. There is a choice between Lowe and Young indices, i.e. whether to price-update the expenditure weights from the weight reference period to the price reference period. The choice should be driven by the target of the CPI and considerations of whether the relative quantities or the expenditure shares tend to remain constant. There is no general consensus on which approach to follow.

(b) The Jevons index is generally recommended for the calculation of elementary price indices because of its better statistical properties. The use of chained short-term elementary price indices seems to include practical advantages in dealing with product replacements. However, in all cases, the chained Carli index should be abandoned.

(c) While elementary indices may be calculated geometrically and higher-level indices arithmetically it can be argued that consistency in aggregation may not be an

important criterion. First, the amount of information available on quantities (expenditures) is not the same at the different levels of aggregation. Secondly, there may be different degrees of substitution within elementary aggregates as compared to substitution between elementary aggregates.

(d) Development and implementation of new statistical systems and software may provide increased flexibility and operational enhancements, including methodological improvements such as e.g. improved imputation methods for missing observations. Some functions which were earlier done manually may be automated. The process can be streamlined and better documented.

(e) Cost-of-living indices may be targeted by calculating the CPI using the constant elasticity of substitution (CES) formula for the interim estimates of the chained CPI. Once weighting data referring to the price reference period becomes available the CPI can be revised accordingly.

(f) From time to time it is useful to evaluate methods applied for compilation of price indices for difficult-to-measure products and services, such as insurance, health, telecommunications and digital equipment. This may include alternative ways to derive weights and considering market changes in terms of price behaviour and changes in the composition of consumption. Proper treatment of services poses particular challenges. It was therefore recommended that further work on the treatment of services be carried out to enable the CPI community to develop and share best practices.

## **F. Session 5: House price indices**

Session chair: James Tucker, Office for National Statistics, United Kingdom

13. The session was based on papers by Canada, Poland and Switzerland. A room document was provided by Turkey.

14. Since the financial crisis interest in both residential and commercial house price indices has increased. This has raised a number of methodological issues as well as issues regarding suitable data sources, data quality and timeliness.

15. The following main points and conclusions were raised during the session:

(a) While a number of countries have produced an official house price index for many years, it is a more recent development for others. Important factors driving development are availability of suitable data sources and political impetus.

(b) The use of administrative data sources is a common starting point for house price index development. While the presentations highlighted the advantages of these sources, they also revealed some substantial challenges, specifically the logistics of obtaining and compiling the data, and ensuring that the data are of appropriate quality. This indicates the importance of developing a strong relationship with data suppliers.

(c) Hedonic models can be put to good use in developing and understanding of price-determining characteristics of property types as well as for the decomposition of the overall price into its land and structure components.

(d) While it is clear that there are often other sources of house price information (including from the private sector), there are advantages in having an official index in line with international standards and sharing of good practice between countries.

(e) Potential topics for further work include securing reliable access to high quality administrative data and comparisons of official indices with those produced by other organisations.

## **G. Session 6: Synergies between CPIs and PPPs and integration of survey activities**

Session Chair: Nada Hamadeh, World Bank

16. The session was based on papers by Morocco, South Africa, CIS-STAT, UNECLAC and UNESCWA with Palestine.

17. The temporal and spatial focus of CPIs and PPPs, respectively, has commonly resulted in different production infrastructures for the two indices. While serving different purposes and exhibiting differences in methodologies, index numbers, item lists and validation methods, there is room for increased synergies as both programs collect price data, and the ICP uses the countries' CPI infrastructure and helps improve it in many cases.

18. The following main points and conclusions were raised during the session:

(a) The session examined possible synergies between CPIs and PPPs and integration of their survey activities. Areas where synergies can be achieved include:

- Overlap in survey frames and outlet types/identification;
- Adoption of the Structured Product Descriptions in identifying items;
- Increased overlap in item lists;
- Harmonized classifications used by CPIs and PPPs, as they are both COICOP based;
- Increased alignment in weights;
- Validation methods and techniques;
- Production of sub-national price indices, which are useful for policy-makers;
- Increased production of detailed CPI indices that support the production of PPPs on a rolling basis.

(b) The increased harmonization would facilitate the production of more frequent PPPs with less additional burden to countries, and would help improve CPI infrastructure and processes.

(c) Closer integration of CPI and ICP should be looked at in future in light of the agreement of the 2016 UN Statistical Commission to adopt an ICP rolling benchmark approach of surveys over a three year cycle beginning in 2017, and the call of the Statistical Commission for exploring closer alignment of ICP and CPI price surveys.

## **H. Session 7: Update of the 2004 CPI Manual**

Session chair: Brian Graf, International Monetary Fund

19. During the 2014 CPI Expert Group meeting countries supported an update of the 2004 CPI Manual. The Intersecretariat Working Group on Price Statistics (IWGPS) subsequently agreed to initiate an update of the manual, with the IMF as lead agency. The manual will be updated by a CPI Technical Expert Group (CPI-TEG) chaired by the IMF. The work will be organised under the auspices of the IWGPS. The main purpose of the update is to reflect methodological and practical developments over the previous decade, evolving user needs, and to provide clear and more prescriptive recommendations wherever possible. Draft updates will be made available to countries for consultation.

#### **IV. Recommended future work**

20. The participants recommended that a meeting of the Group of Expert on Consumer Price Indices should be organised in 2018 and included in the programme of work of the Conference of European Statisticians (CES). The following topics were suggested for possible inclusion in the agenda:

- (a) Services in the CPI
- (b) New Data sources
- (c) Use and communication in the CPI
- (d) Methodological issues
- (e) The updating of the CPI Manual

21. A special session for countries with economies in transition should be organised.

22. UNECE should solicit countries volunteering to organise sessions to ensure in-depth discussions of topics. Organisation of workshops and poster sessions should be considered.

---