Partnership, processes and possibilities: the South African experience of integrating PPP and CPI work

Patrick Kelly and Lekau Ranoto

Statistics South Africa

Paper presented at the ILO/UNECE meeting of CPI experts

Geneva

2-4 May 2016

Abstract

Collection of prices for the compilation of PPPs is often perceived to be time consuming and of little immediate value to a national statistics office. Countries may find the integration of survey activities for both the CPI and PPPs to be a challenge. In developing countries this is compounded by the fact that data collection for PPPs is funded by international organisations. South Africa has formed the base currency for Africa for two rounds of the ICP. This paper compares the similarities and differences in the definition, collection and quality assurance processes involved in the compilation of the CPI and PPPs. The paper describes how Statistics South Africa manages to collect data for the calculation of PPP as part of the monthly CPI process without using extra resources. Indeed, rather than being a burden, this has presented opportunities for improving the CPI and providing additional information, such as average prices, to users.

"The similarity of the basic data collection process for the ICP and CPI should be used to optimize the use of resources. To that end, the ICP should use the statistical infrastructure already in place in a country to the maximum extent possible, and helps improve on it wherever possible. This translates in the ICP using the same outlets, the same price collectors and processing procedures, as long as they comply with ICP requirements." (African Development Bank 2015a)

1. Introduction

The significant improvement in the coverage and availability of purchasing power parity (PPP) ratios since 2005 has enhanced the comparative analysis of economic variables between countries. While it is the World Bank that compiles and publishes PPPs, it is national statistics offices that make this possible by collecting and assembling the necessary price and other economic data as part of the International Comparisons Programme (ICP).

The collection of pricing information for ICP is typically the responsibility of CPI units in national statistics offices. This can be seen as an inconvenience of additional work or as an opportunity for integration and learning. Although the differing purposes of PPPs and CPIs require variation in methodology, their commonality of classification and price collection makes their harmonisation possible and useful.

In South Africa, the statistics office used its own staff, processing systems and resources to collect ICP pricing data. In doing so, it identified elements of the ICP process that have been implemented as improvements for its own CPI. These improvements will further streamline collection of ICP price data in future rounds.

2. Background

The methodology employed in collecting data should always be tailored to the purpose of the intended output. While the inputs to a CPI and PPP are ostensibly the same – prices – their intended use sets up differentiated requirements for their collections. For this reason it is useful to compare the definitions and purposes of a CPI and PPP.

CPIs are temporal price indices, mainly concerned with measuring price changes over time. The focus of the index is on products and services purchased by consumers. CPIs may be primarily intended to measure inflation in an economy or to track changes in the cost of living. The relative importance of these two objectives guides price statisticians in their choice of methodology. In practice many national CPIs, including that of South Africa, aim to serve both of these purposes, being a measure of inflation for monetary policy as well as an escalator of contracts, wages and benefits (Statistics South Africa, 2013).

PPPs on the other hand are a form of spatial price indices. They show the ratio of the prices in national currencies of the same good or service in different economies. PPPs are primarily derived for analytical purposes involving international comparisons of national economic size and growth. PPPs allow cross-country measurement of real per capita income and gross domestic product (Word Bank 2013).

The first ICP was conducted in 1970, covering 10 countries. Recently, ICP rounds have taken place in 2005, 2009 (interim round), 2011 and 2015 (interim round). The ICP has been coordinated at a global level by an office in the World Bank.

The surveys are organised by region and are coordinated by a multinational agency located in that region. The main reason for conducting the surveys on a regional basis is that the products to be priced tend to be more homogeneous within a geographical area, the expenditure patterns are likely to be similar, and the language differences are reduced. There are also operational advantages in having the ICP carried out by agencies that are in close proximity to the economies they are coordinating. Africa has participated in all ICP rounds since 2005. ICP Africa is coordinated by the African Development Bank (AfDB). During the 2011 Africa round 50 countries participated, including South Africa. These countries were divided between four sub-regional organisations (ECOWAS, COMESA, SADC and AFRISTAT).

South Africa has participated in the ICP-Africa programme from 2005. Participation in this round was limited due to a substantial overhaul of the national CPI taking place at

that time. However, certain ICP methods, such as the Structured Product Description, were incorporated into the CPI collection method.

During the 2011 round, South Africa was the base country for the African region. This was because the South African currency is widely used in other African countries and is easily traded on currency markets. Furthermore, South Africa is the most industrialised country on the continent with a wide variety of consumer goods available. These features enhance its comparability with other regional base countries.

In most African countries the ICP was funded by the World Bank and the AfDB as a separate programme, enabling extensive country participation. However, this may have led to the ICP being seen as an add-on with limited incentive to integrate it into existing operations.

South Africa, on the other hand, used its own permanent staff for ICP operations and did not request or receive any AfDB funding. Improved integration of the ICP work with the regular CPI work was therefore necessary to ensure achievement of both areas of work.

The AfDB provided extensive support to national statistics offices, including the compilation of manuals for price collection, collection survey forms, data entry sheets, validation tools and training.

3. Product selection

The ICP has to manage two competing challenges with regard to product selection. Firstly, there must be a sufficient number of products that are comparable across each continent and the globe for the purposes of price level comparison. Secondly, the particular consumption practices of each area must be properly reflected to ensure relevance to actual expenditure. The South African ICP basket was determined by first checking the common products in the CPI and ICP baskets. The data are extracted from the CPI database and specific restrictions (e.g. outlet type, units of measurement,

brand, etc.) are filtered per product. This process ensures that the exact required ICP product is matched with the correct CPI product.

The other products might be available in the country but not in the CPI basket. These products are collected by price collectors during the CPI collection as an extra collection. It was possible to accommodate this additional work into the schedules of the permanent CPI price collectors.

During the ICP 2011 round, there were 1 032 household products to be collected. South Africa collected a total of 585 products, of which 369 were common ICP-CPI products and the remaining 216 were collected in addition.

Because of the differing purposes of the project, the basket selection criteria for the two are different. What is common is that both use Classification of Individual Consumption According to Purpose (COICOP) for classification.

Because the CPI monitors price *changes* of the items over time it is agnostic as to the brand, size, etc. of each particular product being priced. In other words there might be different items at different price levels in the CPI basket, but that will not pose a bias in the calculation.

For example, there are different varieties of rice represented by different brands, different sizes and different price levels. Each month the CPI will monitor the price change of each item, and the price changes are aggregated by means of the index formula to achieve an aggregate price change for a particular level of the classification.

Alternatively, the ICP is interested in average prices. Any variation from a strict specification will therefore be seen as an outlier affecting the average price. It is for this reason that the products within the ICP basket are defined in more detail. The strict specification is all the more important to ensure proper comparability across all participating countries.

As a result of the different approach to specification, it happens that a number of ICP products may be priced under a single CPI product description. For example Basmati, white and Jasmine rice are separate ICP products but are all eligible for the CPI

category of rice. Consequently, a manual matching process is undertaken in the first month of the ICP round to identify relevant items priced for the CPI.

Due to the greater flexibility in product choice for the CPI, product specifications such as brand and product names are captured in a free text description field. This gives rise to numerous spelling errors, and reduces our ability to conduct deeper analysis of price behaviour of identical products. However, the stricter ICP approach has illustrated the benefits of ensuring standardised capturing of names and product sizes. Revisions are currently being made to the CPI forms and capturing system to provide for drop down lists to provide for this. It is anticipated that this improvement will assist ICP validation in future rounds.

4. Outlet selection

Together with the price and the product specifications, the outlet type is one of the fundamental elements of ICP collection. To maintain comparability across countries, prices should be collected from the same outlet type.

For CPI, type of outlet is not specified and purposive sampling is used for selection of outlets. That is, any outlet that carries the required product may be selected.

This restriction on the type of outlet in ICP limits the number of products that can be used from the CPI. Even if the products are available in the CPI basket or in the country, they might not be included in the country's ICP basket because of a mismatch of type of outlet. For example, during the 2011 round, some products were required to be priced from informal markets but the CPI basket did not have informal markets as an outlet type. Another example is the requirement for products from 'Butchery without refrigerator', which do not exist in the country.

It was for this reason that South Africa decided to include a field for type of outlet in the household expenditure survey and the CPI collection forms following the 2012 reweighting exercise. Categories of chain store, independent stores and informal markets were introduced. Despite the fact that these outlet types are not identical to the ICP (supermarket, specialised stores, department stores, wholesale stores, discount

stores etc.), they do provide a step towards better alignment and allow for analysis of pricing behaviour across different outlet types. There is still opportunity for expanding the outlet types for CPI collection since this will ensure that the prices are not biased to specific types of outlets. We have noted that most prices in the CPI are from chain outlets since these are pervasive and easy to collect from.

5. Price collection

Prices are collected by both field-based agents (mainly for goods) as well as head office staff (mainly services). The surveying of goods prices entails fieldworkers (price collectors) who visit sampled outlets in order to record actual prices. The collection is carried out on a monthly basis.

The collection of data in the CPI utilises the following forms:

a) Outlet Cover Page, used to capture data on the outlet status, address, the responsible person for collection and quality control.
b) Structured product description form (SPD), used to initiate a new product selected for pricing according to specifications.
c) Pricing form, used to price products that were previously initiated (selected for pricing).

These forms were designed during a substantial re-engineering of the CPI in 2005. The SPD form was adopted from the ICP 2005 round and it has been extremely valuable for product initiation. A variation of this is now also used in the producer price index. Even though the ICP uses spreadsheets for collection, the fields in the spreadsheet correspond to the survey forms designed for the CPI.

For the ICP products which do not form part of the CPI, SPD forms are created and sent to the teams in the field to collect from the same outlets where they are collecting CPI prices. If these products were not available in the existing CPI outlets, new outlets were used. The same CPI publicity methodology is used by fieldworkers when approaching those outlets.

Integration between CPI and ICP is cost effective since there is no need to employ extra collectors and build new systems. However, it is important that price collectors are trained thoroughly for ICP collection so that they understand the different product specifications. Poor understanding on the part of fieldworkers causes delays in the field and results in uncertainties, leading to the wrong item being priced or no price being collected at all.

6. Data processing

After collection, both CPI and ICP forms from the different regions are sent to head office to be captured. Forms are captured on the CPI database, where all editing processes are done following the CPI methodology. At the end of editing, a clean database is ready for any type of analysis. Although the ICP and CPI use different product codes, a correspondence table allows specific CPI items to be extracted using a custom built SAS program. ICP designed a data-entry sheet to be used for capturing all collected prices. It aims to ensure a homogeneous structure of the datasets across all countries. Following capturing, data are merged by an additional tool which gives as output a country data file.

This data entry sheet is not used in South Africa for capturing purposes to avoid creating different databases for CPI and ICP. Rather, everything is captured in the CPI database and entries required for the ICP are extracted to the ICP data-entry sheet. This means that only data that have passed the CPI quality-validation processes are imported into the ICP data-entry sheet. This process results in speedier validation of ICP data.

The data entry sheet provided lessons for the CPI. Since the sheet contains protected cells to avoid mistakes and guides the user during the entry of data, these functions have now been built into the reviewed CPI capturing system where many of the fields will now have drop down lists to choose from. In this way, errors are avoided at an early stage.

7. Data validation / quality assurance

Quality of the data is a vital prerequisite for statistical analysis and output. CPI and ICP projects use different methods to validate the data collected.

Quality assurance for CPI data is done by editing data in the databases. Data editing is an activity aimed at detecting and correcting errors in data. This consists of four phases:

a) *Validation checks* to ensure that the correct item status codes were allocated during capturing and to allocate missing unit codes;
b) *Logical edits* to ensure that the current and previous month's data do not have contradictory values;
c) *Range edits* to identify whether the data item value falls inside a determined acceptable range; and
d) *Final checks* of the correctness of all the allocated item status codes.

After the editing process is complete, the data are exported to a SAS application for compilation.

Validation and calculation of average prices for the ICP are completed in the ICP-developed SEMPER software. SEMPER analyses the observed prices and provides descriptive statistics (central tendency and dispersion) in a few seconds for checking the consistency of data. A junior statistician was allocated to carry out all processing and validation work, which occupied less than three days of her time each month.

SEMPER performs the following actions:

a) *Conformity checking*: coding of product name, quantity, unit of measurement and others characteristics related to the description of the products. Any product which is not part of the list of products will be rejected by the software.
b) *Factor conversion*: To convert the unit of measurement (e.g. grams, millilitres) from that collected to that required for ICP comparison.
c) *Recalculation of prices (range edits)* to identify outliers.

If there are errors detected from the SEMPER results, these are attended to and corrected on the country data file. SEMPER is re-run until all errors are solved. Validated Country Data files are sent to AfDB monthly for consistency analysis per region. If all is validated, then PPP calculations are carried out.

Following the introduction of ICP-inspired enhancements discussed above, South Africa started calculating averages prices from the CPI data, initially at a national level (in 2000), and provincially from 2010. These are not published but made available on request. There has been substantial demand from government departments wanting price benchmarks for procurement activities.

8. Possibilities for integration at a regional level

Representatives of participating countries are frequently gathered together by the AfDB to perform validation checks. The benefit of these is not clear and certainly constitutes an additional burden on staff from statistics offices. These validation exercises may be necessary only because a number of countries do not see ICP as part of their core operations and do not therefore prioritise it. This reduced commitment results in late submission and poor quality of data.

Despite being intimately involved in the collection of ICP data, national statistics offices play no role in the dissemination of the results. Deeper involvement in presenting and explaining results may require additional training, but should lead to enhanced integration and commitment of countries.

9. Conclusion

Resistance to change and integration is common because introducing something new is usually perceived as extra work. Integration between the CPI and ICP in South Africa has proven to be a valuable exercise. Integration has allowed the ICP collection to be conducted with a minimum of additional effort and no additional resources. It has demonstrated areas for improvement of CPI systems which enhance the quality and

range of pricing data. ICP has further provided opportunities for research work such as calculating sub-national PPPs (Kgantsi, 2012).

10. References

African Development Bank Group. 2014. *Comparing the Real Size of African Economies: Results of the 2011 International Comparison Program 2011 for Africa*. Abidjan.

Kgantsi, EM. 2012. *A comparative study of purchasing power parities for the food component using the consumer price index data in the South African provinces.* Dissertation submitted in fulfilment of the requirements for the degree of Master of Science. University of the Witwatersrand, Johannesburg.

Statistics South Africa. 2013. *The South African CPI Sources and Methods Manual*. Pretoria.

World Bank. 2013. *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program—ICP*. Washington, DC.