Can't weight anymore!

South Africa's use of national accounts and point of sale data to update the CPI basket and weights

Patrick Kelly and Matlhatsi Mogalanyane

Statistics South Africa

Meeting of the Group of Experts on Consumer Price Indices
7-9 June 2023, Geneva, Switzerland

Abstract

Funding constraints and COVID-19 lockdowns prevented Statistics South Africa (Stats SA) from surveying household expenditure data in time for a five yearly update of the consumer price index (CPI) basket and weights. This despite the importance of a relevant CPI for economic policy and the urgency of updated poverty information. In 2022, Stats SA instead used recently benchmarked national accounts estimates together with detailed sales data provided by retailers to update the weights and basket. The paper explains the technical methods used to determine changes to the basket and weights and highlights the importance of user consultations during the project. An assessment of the results lays the methodological groundwork for future weight updates. Key features of the update included using a pre-COVID year as the weights reference period, abandoning price updating and an uneven focus on retail goods at the expense of services.

Contact

patrickke@statssa.gov.za
matlhatsir@statssa.gov.za

1. Introduction

According to international standards, the weights and basket of the consumer price index (CPI) should be updated at least every five years. This is to ensure that changes in consumer expenditure are reflected in the weighted aggregates of the measure of inflation. South Africa subscribes to the International Monetary Fund's (IMF) Special Data Dissemination System (SDDS) and considers its CPI to comply with key international methods.

What happens when the country's fiscal framework does not fund a household expenditure survey (HES)? Faced with the possibility of failing on a key feature of international best practice, as well as missing changes in consumer behaviour, Stats SA adopted a new approach to update the CPI basket and weights.

The paper explains how the CPI benefitted from a timely revamp and benchmarking of the national accounts, and the methods used to adjust the weights. Detailed point of sale (POS) data from retail chains provided the basis for changes to selected categories of the basket.

2. Background

2.1 Household expenditure survey

The South African CPI basket and weights were previously updated in January 2017 based on a household expenditure survey – the Living Conditions Survey (LCS) – conducted over 12 months during 2014 and 2015. No additional funding was allocated by National Treasury for this survey, and it only took place due to skimming of budgets from other areas of Stats SA.

Obtaining full funding for the next expenditure survey was seen as essential because the survey is designed to be representative of the entire population with the ability to provide geographic and population group breakdowns. Importantly, the survey captures expenditure on all products in a particular period, providing for calculation of detailed expenditure proportions. The revised CPI manual (IMF 2020, p108) affirms that a household expenditure survey is "*the primary data source for deriving expenditure shares for the goods and services covered by the CPI.*"

Following the 2014/15 LCS, Stats SA made annual requests for a permanent budget for an expenditure survey. Different options, including an annual, multi-module household survey, to a stand-alone expenditure survey were offered. Due to the requirement that an expenditure survey cover 12 months and the detail of the data required from households, the survey is more expensive than most household surveys. The approximate budget was R300 million spread over two financial years. This was more than 10% of the entire annual Stats SA budget of about R2,5 billion. Unfortunately, no additional funds were allocated, and it became clear that no fresh expenditure data would be collected before 2019/20 to allow an update by January 2022 – five years after the previous basket and weights were updated.

Approximately 1/3 of the required funds were provided in the 2020/21 financial year. Although this was inadequate to run the survey, the COVID pandemic set in and scuppered any possible alternative uses of the money.

2.2 The structure of the South African CPI

The structure of the product and geographic classification of the CPI is important to understanding the limitations of not having a comprehensive expenditure survey to update the weights.

The South African CPI closely follows the international Classification of individual consumption by purpose (COICOP) structure as illustrated in Table 1. This structure forms the basis of what data sources are most appropriate for weights at the different levels of the classification.

Table 1. Classification hierarchy of the CPI

COICOP level	Name	Example
2-digit	Division	Food and non-alcoholic beverages
3-digit	Group	Food
4-digit	Class	Bread and cereals
5-digit	Sub-class	Bread
8-digit	Indicator product	White bread
12-digit	Sampled product	Albany 700g loaf of white bread in store xxx

The CPI divides the country into three categories: large or metropolitan urban areas (metros), other (non-metro) urban areas and rural areas. The headline CPI covers all urban areas. In most provinces there is one weight for each category, however, Gauteng province has separate weights for three metros (Johannesburg, Tshwane and Ekurhuleni); Eastern Cape province has separate weights for two metros (Nelson Mandela Bay and Buffalo City) and Kwa-Zulu Natal province has weights for two municipalities (eThekwini and Msunduzi). Weights are allocated to each indicator product in each of these areas which then comprise the elementary indices of the CPI at an 8-digit level.

Indices are published in aggregated format at national and provincial levels. The statistical release contains information at the 3 and 4-digit level while more detailed indices are available in an Excel format on the Stats SA website.

In addition to the classification-based indices, Stats SA computes a number of analytical indices, aimed at shedding light on the inflation experienced by specific groups. These include geographic-based indices for rural areas and the total country, indices for expenditure deciles, a CPI for pensioners as well as core measures of inflation.

3. The solution

3.1 National accounts – international perspectives

National accounts data are used by a number of statistical offices to update CPI weights. Specifically, the household final consumption expenditure (HFCE) component of the accounts measures the expenditure by households using the same concepts and classification as the CPI. The CPI manual emphasises that national accounts should be seen as *"an alternative source for deriving CPI expenditure weights"* (IMF 2020, p110) and should not replace an expenditure survey. However, "*national accounts data may be used when the household expenditure survey is conducted too infrequently to ensure the reliability*

of the CPI" (p33)*.* The manual also acknowledges that the value of national accounts data as an alternative to a HES is that it is updated regularly.

Advantages of using national accounts data are that they are updated frequently, are nationally comprehensive and that the expenditure on certain products (e.g. insurance and gambling) is treated in the same way as required for the CPI.

However, there are some drawbacks to using the national accounts data including the limited level of product detail, some data sources may include business and non profi expenditure which is difficult to remove, and estimates are frequently revised as new data becomes available.

Increasingly, more countries are using HFCE data to update the CPI weights. The European Harmonised index of consumer prices (HICP) requires that countries update their weights on an annual basis and are derived from national accounts. The HICP is deeply rooted in the national accounts with all "*concepts, definitions and conventions adopted in the HICP ...consistent with those used in the European System of Accounts*" (Eurostat 2018). A similar practice is applied in the United Kingdom's CPIH (ONS 2017).

Since 2018, the Australian CPI weights have been updated annually using HFCE data (ABS 2017). Here the growth rates in the matched elementary aggregates are used to update the proportions of the existing CPI weights. For lower-level geographic estimates and indices for special populations, it is assumed that their change in expenditure is the same as the total aggregate.

The Australian initiative is pertinent to South Africa due to the long time period (six years) in between household expenditure surveys, especially when compared to some of the European countries and the UK which have annual or biannual surveys.

In Brazil, the Institute of Geography and Statistics (IBGE) has done substantial research work to identify possibilities and constraints of using national accounts data. They have not yet implemented this approach and there is no timeline to do so. Similar to the ABS, they intend to maintain the existing lower-level proportions of the CPI and to apply the national growth rates reported in the GDP (Ventura 2019).

3.2 National accounts – local developments

Prior to 2016, the HFCE and other GDP by expenditure accounts, were compiled by the South African Reserve Bank (SARB). Following a lengthy capacity building and transition process, Stats SA began publishing these accounts from the first quarter of 2016. The changeover in compilation responsibility was accompanied by a major review and update of the source data and methodologies used for compiling the accounts (Stats SA 2016).

In 2021, Stats SA published benchmarked and rebased national accounts. This led to a 16% upward revision in the size of household consumption in the 2015 base year compared to 9,2% for the economy as a whole. This change resulted from significant improvements in coverage, source data, and classification (Stats SA 2021). Importantly, HFCE was based on the 2014/15 LCS, meaning CPI and HFCE structures shared a common data framework. Improvements were made in the measurement of owner occupied housing, gambling and insurance services. In some of these cases, methods previously used in developing CPI weights were employed.

There is no perfect fit between the data available in the national accounts and that required for the CPI weights. While a common classification is used, the CPI weights are determined with finer granularity than the national accounts. The 2016 CPI basket comprised 404 products (at a national level) while the HFCE reports expenditure for 79 categories at the 4-digit level and 15 where the 3-digit is the lowest level of detail. The national accounts report for the whole territory of South Africa whereas the CPI publishes provincial indices and is weighted at lower levels as explained in Section 2.2.

Table 2. Shares (percentage) of consumer expenditure – 2016

COICOP	HFCE	CPI total country
Food and non-alcoholic beverages	14,6	19,1
Alcoholic beverages and tobacco	4,8	6,0
Clothing and footwear	5,2	4,1
Housing and utilities	15,1	22,5
Household contents and equipment	5,8	4,3
Health	6,9	1,3
Transport	15,7	14,7
Communication	4,2	2,6
Recreation and culture	7,2	4,8
Education	3,2	2,3
Restaurants and hotels	5,1	3,4
Miscellaneous goods and services	12,2	14,9

Table 2 shows the difference in expenditure shares between the 2016 CPI weights and the equivalent HFCE. Despite the methodological improvements in the national accounts, there are large differences with the CPI weights. Significant differences also exist at lower levels of detail. Simply adopting the HFCE structure for the CPI was not feasible as it would would have disrupted the time series.

The benchmarked HFCE was rooted in the 2015 reference year, and a nominal time-series is available for subsequent years. The 2019 nominal values are therefore updated with a more limited set of data than the reference period.

3.3 Other data sources

Stats SA has previously made use of alternative data sources in order to supplement the household expenditure survey (Stats SA 2017).

Extensive use of business survey data assisted with previous basket updates including large sample surveys (LSS) of the retail, accommodation and food and beverages sectors. However, without the detailed HES data, it proved difficult to link the more aggregated LSS results and the CPI basket. The LSS surveys the formal sector businesses and so excludes sales in the informal sector – however this is believed to be a small percentage of overall sales. The different classifications used – LSS uses Central Product Classification (CPC) – is a further obstacle.

For selected product groups, Stats SA has also previously made use of administrative records such as tax revenues on alcohol and tobacco, and motor vehicle registration data for

purchase of vehicles. These data have proven most useful in obtaining a proper level of expenditure as these categories tend to be under reported in the expenditure survey. However, these are already used to compile the HFCE estimates. Some other administrative data and industry reports were sourced to verify the national accounts information.

Summarised point of sale data was provided, in the past and for this exercise, by a number of retail chains covering all product categories where expenditure is primarily through retail outlets. Millions of rows of data contain annual sales values for each unique stock keeping unit (SKU). This data required classification into COICOP, which was achieved by coding the product descriptions. This information was previously used for detailed level basket selection and calculation of weights.

Importantly, all these complementary data sources have previously been used in conjunction with HES data.

3.4 Weights reference period

COVID-19 had a devastating economic impact, starting in 2020 and continuing into 2021. Overall, the South African economy only reached pre-COVID levels in the first quarter of 2022 (Stats SA, 2022). COVID-19 forced changes to consumer spending patterns with big shifts in restaurants and hotels, clothing and footwear, and alcohol and tobacco (see Table 3).

Table 3: Change in HFCE 2019 and 2020

2 Digit COICOP	% change 2019-2020
Food and non-Alcoholic beverages	4.0
Alcoholic beverages, tobacco and narcotics	-13.8
Clothing and footwear	-20.3
Housing, water, electricity, gas and other fuels	5.7
Furnishings, household equipment, maintenance	-3.0
Health	-1.4
Transport	-9.9
Communication	5.1
Recreation and culture	-13.3
Education	7.5
Restaurants and hotels	-40.0
Miscellaneous goods and services	5.8
Total Expenditure	-3.9

Source: Stats SA 2023

When planning the weights update during 2020 and 2021, it was not known how long the pandemic would continue. As the weights remain fixed for a number of years into the future, the CPI manual guides that the weights should reflect a 'normal' consumption period and that weights that are not likely to change much in the future should be adopted.

In the past, Stats SA has used growth factors sourced from specific short-term industry surveys as well as price updating to adjust the weights to refer to the same period as the index reference month – that being the December before implementation. The revised CPI

manual advises that price updating may exaggerate the upward substitution bias inherent in a Laspeyres-type index and cautions against its use.

The year of 2019 was selected as the weights reference period in order to satisfy the considerations outlined in the preceding paragraphs. However, this meant that there was actually only a two year period between the current weights reference period (December 2016) and the new one (2019).

3.5 Calculating the weights

As a result of the differences in the expenditure shares between the CPI basket and the national accounts data, the changes to the weights were derived from the HFCE growth rates between annual estimates for 2016 and 2017 (providing a proxy for December 2016) and 2019.

This method assumes the structural validity of the original CPI weights. It also retains the existing proportions for elementary aggregates, regional baskets and special aggregations for population groups.

Table 4: Changes in HFCE and CPI weights 2016/17-2019

COICOP	HFCE growth	% change to CPI weights values
TOTAL	15,2	15,0
Food and non-alcohol beverages	14,5	14,6
Alcohol beverages, tobacco and narcotics	21,8	22,1
Clothing and footwear	9,1	8,6
Housing, water, electricity, gas and other fuels	14,7	15,0
Furnishings, household equipment, maintenance	15,2	16,3
Health	18,2	17,8
Transport	13,2	15,1
Communication	6,1	6,1
Recreation and culture	13,2	16,7
Education	19,8	19,8
Restaurants and hotels	17,6	17,7
Miscellaneous goods and services	19,3	13,4

As seen in Table 4, HFCE growth and high-level changes in the CPI weights are not a direct match. HFCE growth rates were applied to the lowest detail available (3 or 4-digit) from HFCE to the 8-digit CPI values in a bottom up approach. The new 8-digit estimates were summed up to obtain aggregates at next highest level. Below (Table 5) is an example of the bottom up approach, using transport services as an example. Here the growth rate obtained from the national accounts at the 4-digit level (passenger transport by road) is 18,6%. This factor is applied to the expenditure values for the lower level detail. The total of R68,129 billion is the sum of the 8-digit expenditure values.

Table 5. Applying 4 digit national accounts growth rates to 8-digit CPI values (R'000)

COCIOP	2016 CPI expenditure	HFCE growth rate	2019 CPI expenditure
2 digit Transport			
3 digit Public transport			
4 digit Passenger transport by road	57 422 923	18,6%	68 129 673
8 digit Local bus fares	48 806 307		57 906 453
8 digit Car rental	6 191 030		7 345 374
8 digit Minibus taxi fares	1 553 499		1 843 155
8 digit Long distance bus fares	872 087		1 034 691

Table 6 shows the changes to the headline CPI weights at a 2-digit COICOP level. The changes are all less than one percentage point. The category with the largest increase is alcoholic beverages and tobacco and with the biggest decline is miscellaneous goods and services.

Table 6. CPI headline weights 2016 and 2019

Category	2016	2019	Difference
Food and non-alcoholic beverages	17,2	17,1	-0,1
Alcoholic beverages and tobacco	5,8	6,3	0,5
Clothing and footwear	3,8	3,7	-0,1
Housing and utilities	24,6	24,5	-0,1
Household contents	4,4	4,4	0,0
Health	1,4	1,4	0,0
Transport	14,3	14,4	0,1
Communication	2,6	2,4	-0,2
Recreation and culture	5,2	5,2	0,0
Education	2,5	2,6	0,1
Restaurants and hotels	3,1	3,3	0,2
Miscellaneous goods and services	15,1	14,8	-0,3
Total	100	100	

3.6 Updating the basket

The inherent disadvantage of a Laspeyeres-type price index is that it does not account for changing consumer behaviour, and cannot accommodate new products. As a result a key component of updating the weights must be to review and update the basket of goods and services.

Point of sale data sourced from retail chains provided the basis for making changes to the CPI basket. After classifying and aggregating the sales information into discrete elementary level product groups (indicator products), the top 80% of products are considered for selection. Any new products within the cut-off threshold that were not part of the existing basket are considered for inclusion.

In cases where a new 8-digit product was added to an existing 4 or 5-digit COICOP group, the expenditure share of the higher level aggregate was kept constant and new products were incorporated by distributing expenditure among the selected items.

In some cases a new 4 or 5-digit COICOP group was created. In these cases, the 3-digit COICOP aggregate expenditure was fixed and new expenditure was added.

Based on the availability of products experienced during data collection, some products were grouped together or split to improve homogeneity, sample size and collection rates.

The number of products in the basket increased to 415 from 404 in the 2016 basket. A total of 14 products entered the basket and two were removed. One product was split into two; in two cases, two products were combined into one; and one product was renamed.

A major limitation was that this update was only possible for goods sold through retail channels. No low-level data on services was available. However, industry reports did not indicate significant increases in expenditure on new services products.

3.7 Communicating with stakeholders

Stats SA had widely publicised the lack of funding for the HES and the associated risk that South Africa's CPI may become non-compliant with international best practice. The methods used to update the basket and weights without the HES represented a significant change in methodology for the CPI. Good practice and Stats SA standards require timely and transparent communication to users of the changes.

A key feature of the communication strategy was the establishment of an advisory committee consisting of representatives of the SARB, National Treasury, economists from several commercial banks and economic research institutions. The aim of the committee was twofold. Firstly to obtain inputs and advice from expert users. Secondly and perhaps more importantly, it was to obtain the buy in and understanding of key industry role players. The committee met twice, first to discuss the proposed methods, and then to share the preliminary results.

Separate meetings were held with a broader group of SARB and National Treasury officials, as well as a subcommittee of the Statistics Council prior to the release of the results.

Documents were placed on the Stats SA website together with a note in the monthly CPI release to advise users of the proposed changes. At the time of releasing the new basket and weights, Stats SA held a public briefing and conducted numerous media interviews.

4. Looking forward

Stats SA eventually received funding for an Income and Expenditure Survey which kicked off field work in November 2022. This is a permanent budget allocation which should yield expenditure estimates every three years. Stats SA plans to introduce an update to the CPI basket and weights in January 2025.

The experience of using the national accounts and retailer POS data has pointed to lessons to be considered in the run up to this next update.

- Working with the national accounts data has improved the understanding amongst the price statistics team of how the national accounts are compiled. Much of the adjustment work that was done in previous reweighting exercises (e.g. to raise the

share of alcohol and tobacco) is now done by the national accountants. Using this work will improve the alignment of national accounts and CPI expenditure shares.

- Stats SA will investigate using the national accounts data as the basis for higher-level weights and updating these on an annual basis. This is already the practice in the producer price index and the experienced gained on the CPI illustrated its feasibility.

- Classifying and aggregating the retail POS data is time consuming and technically demanding. Obtaining and processing the data each year, even if the results are not applied to the CPI annually, will enable the development of capability and experience within the price statistics CPI team. This experience can be applied to other projects such as research into the use of scanner data for CPI compilation. The development of a time series will allow proactive analysis of changes in consumer behaviour to guide basket updates. It is possible that this information may assist in reducing respondent burden for large sample business surveys.

- This experience reinforced that there is no proper replacement for a household expenditure survey. Stats SA did not obtain data to update the services components of the basket. Consumer preferences for services may change more rapidly than goods and this basket update missed any of these shifts. The fact that the HES provides detailed information on all COICOP groups is essential for creating a balanced set of weights and a basket that reflects all the shifts in consumer spending.

References

ABS, 2017. An implementation plan to annually reweight the Australian CPI. https://www.abs.gov.au/statistics/research/implementation-plan-annually-re-weight-australian-cpi

Eurostat, 2018. Harmonised Index of Consumer Prices (HICP) Methodological manual. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-17-015

IMF, 2020. Consumer Price Index Manual: Concepts and Methods. https://www.imf.org/en/Data/Statistics/cpi-manual

ONS, 2017. Investigating the impact of different weighting methods on CPIH. https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/investigatingtheimpactofdifferentweightingmethodsoncpih

Stats SA, 2016. Expenditure on GDP: Information Note. https://www.statssa.gov.za/publications/P0441/Information_note_on_Expenditure_on_GDP.pdf

Stats SA, 2017. Introduction of new weights and basket for the Consumer Price Index. https://www.statssa.gov.za/cpi/documents/Introduction_of_2016_CPI_weights_and_basket.pdf

Stats SA, 2021. National Accounts: Source and methods. Report 04.04.04. https://www.statssa.gov.za/publications/Report%2004-04-04/Report%2004-04-042021.pdf

Stats SA, 2022. Economic recovery from COVID-19: Not all countries are equal. https://www.statssa.gov.za/?p=15690

Stats SA, 2023. GDP time series. https://www.statssa.gov.za/?page_id=1854&PPN=P0441

Ventura, 2019. Challenges of Using National Accounts for a More Frequent Update of the Brazilian CPI Weights. Paper presented at the 16th Meeting of the Ottawa Group on Price Indices. https://www.ottawagroup.org/ottawa/ottawagroup.nsf/home/Meeting+16+-+Rio+de+Janeiro,+2019