REGIONAL SEMINAR ON DATA COLLECTION FOR COMPILATION OF CONSUMER PRICE INDEX

Summary report of the seminar

June 2020
REGIONAL SEMINAR ON DATA COLLECTION FOR COMPILATION OF CONSUMER PRICE INDEX

Addis Ababa (online)
18-28 May 2020

Summary report of the seminar

June 2020
To order copies of *Regional Seminar on Data Collection for Compilation of Consumer Price Index*, please contact:

Publications Section  
Economic Commission for Africa  
P.O. Box 3001  
Addis Ababa, Ethiopia  
Tel: +251 11 544-9900  
Fax: +251 11 551-4416  
E-mail: eca-info@un.org  
Web: www.uneca.org

© 2021 Economic Commission for Africa  
Addis Ababa, Ethiopia

All rights reserved  
First printing July 2021

Material in this publication may be freely quoted or reprinted.  
Acknowledgement is requested, together with a copy of the publication.

The designations employed in this report and the material presented in it do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations Economic Commission for Africa concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Designed and printed in Addis Ababa, Ethiopia by the ECA Printing and Publishing Unit. ISO 14001:2015 certified.  
Printed on chlorine free paper

Cover photos: Shutterstock.com
Background and introduction

The outbreak of coronavirus disease (COVID-19) has brought about an increasing demand for statistics and data, as policymakers and the public want to know how it will affect various aspects of the economy. The ongoing social-distancing and lockdown measures introduced to curb the spread of the virus, have also constrained statistical data collection activities, as most of the censuses and surveys are conducted through face-to-face interviewing. This has presented immense challenges to national statistical systems and their operations. While the challenges caused by COVID-19 may differ from country to country (owing to different national circumstances and patterns), there are some commonalities that require joint efforts to find technical solutions suitable for the continent. Once adopted, the modern technology and techniques can be applied and continued after COVID-19.

At a teleconference with the directors-general of national statistical offices in Africa, in May 2020, countries requested from pan-African institutions and development partners an organized setting in which they could share and exchange good practice and experiences. In addition, they requested advice and guidance on how to continue consumer price index (CPI) data collection through new and alternative approaches and techniques. To this end, the Economic Commission for Africa (ECA), along with partner institutions, organized a regional seminar on data collection for the compilation of CPI in the context of COVID-19, via teleconference from 18 to 28 May 2020. The seminar provided a forum for the exchange and sharing of practice and experience among countries, international and regional organizations, academia and development partners.

In his opening remarks, Oliver Chinganya, Director of the African Centre for Statistics, ECA, indicated that the results of a survey conducted with regard to the impact that COVID-19 had had on the activities of African national statistical offices, had been sent to the directors general of national statistical offices recently. He further indicated that exchanges of experiences were required on how to tackle the challenges faced by African countries because of the pandemic. Stressing the need to look for alternative ways of data collection, he informed participants that the price data watch would be launched during the following month, and the importance of having frank discussions so as to share the expertise of participants. To conclude, he informed participants that the report of the regional seminar would be shared with all of them.

The purpose of the seminar was to provide a forum for sharing and exchanging experiences and best practices on alternative methods of data collection for the compilation of CPI during COVID-19 and beyond. The specific objectives of the seminar were to: better understand the impacts and challenges of COVID-19 on the activities and operations of the data collection and compilation of CPI; exchange countries’ experiences of mitigation strategies and techniques; identify and evaluate innovative data collection approaches and methods to ensure consistency throughout this period, and comparability across countries.

The online regional seminar spanned for eight days, usually with two sessions for each day. There were 16 sessions in total, including one opening, one closing, and 14 substantive sessions covering such issues as international standards and recommendations; data harvesting based on website and data science; and data collection through telephone surveys. The following institutions and development partners joined ECA in giving presentations:

- Economic and Statistical Observatory for Sub-Saharan Africa (AFRISTAT)
in all, 432 participants from 51 member States and 29 agencies and institutions registered for the online regional seminar (see annex I).
Chapter 1: International standards and recommendations

SECTION 1.1 BUSINESS CONTINUITY NOTES ON PRODUCING CONSUMER PRICE INDEX DURING COVID-19 BY THE INTER-SECRETARIAT WORKING GROUP ON PRICE STATISTICS

Carsten Boldsen
Chief, Economic Statistics Section
Economic Commission for Europe

SUMMARY OF PRESENTATION
The presenter gave an overview of the guidance notes for producing CPI prepared by the Inter-Secretariat Working Group on Price Statistics (IWGPS). The guidance notes provide information on data collection, imputation methods and publication.

For price collection from outlets that remain open, alternative data sources include the outlet’s website, newspapers and advertisements, telephone, email, in-person price collection by national statistical office staff and scanner data. For restaurants and cafes – take-away or delivery menus may be priced if comparable to the sit-down menus usually followed. For airlines, hotels and package holidays – it may often be possible to collect prices from the web or brochures. When it is not possible to obtain suitable prices, they should be treated as temporarily missing and imputed.

Missing prices for products that are still available may be imputed by the price change of comparable products, the change of the elementary index or the nearest higher-level price index. For products with strong seasonal price variation, monthly or annual price development of similar products may be used to impute missing observations. Carry forward in general is not recommended but, if justified, the last observed price may be carried forward for a limited period. Methods should be "self-correcting", that is to say, the index must show the correct change from the last period when prices were observed to the period when prices again can be collected.

If there are no transactions the recommendation is to impute with a comparable elementary index or nearest higher-level price index. Imputation with the all-items CPI is also a suitable method and can be justified if there are no transactions. This corresponds to leaving the elementary index out of the CPI calculation. For seasonal products, prices of out-of-season products should be imputed in line with methods in the CPI manual (chapter 11). For regional indices, missing prices should be imputed by available prices and indices within the region. If this is not possible, imputations may be made by CPI data of a neighbouring or similar region or the national CPI. Expenditure weights should be kept fixed, adhering to the regular schedule for updating of weights. Changing weights are not consistent with the fixed basket approach and will create breaks in time series.
When deciding on imputation methods, the main uses of CPI must be considered. Different methods have different effects on the monthly and annual rate of changes of CPI. Methods must be documented to support production and for information of users.

The principles for publication of official statistics should be followed. It is important to be transparent to maintain public trust in CPI. Documentation and explanations of methods should be made available to users. National statistical offices should continue to publish CPI sub-indices, and if these are imputed. Information about imputations because of COVID-19 should be provided with the release of CPI. Indices with full or significant imputations due to COVID-19 should be flagged. If possible, address the impact it continues to have on the overall quality of CPI.

QUESTIONS AND ANSWERS

(a). What needs to be considered when conducting the imputation of CPI?

- Impute by prices of similar products, if not available, apply a bottom-up approach imputing with the nearest available higher-level price index.

- Carry forward is generally not recommended and should only be used for products with stable price development and for a limited period.

- Document methods and procedures to ensure that these can be followed and for information of users.

- Expenditure weights should be kept fixed, elementary indices for which prices cannot be collected should be imputed.

(b). How do you make use of data from online website?

- Online prices may be used to replace physical outlet prices for comparable products, quality differences should in principle be adjusted for.

(c). How do you collect data from informal markets?

- In practice, it is a challenge to collect prices from informal markets. Different countries follow different methods to tackle this problem. Alternative data sources may be used.

(d). How do you handle the prices of banned or illegal products (goods or services)?

- Markets may not behave in a normal way in cases of bans imposed, such as on sale of alcohols and on transport services, and sales may take place in black markets perhaps with inflated prices. Illegal products are in the scope of CPI and should in principle be covered. However, for practical measurement reasons, this may not always be possible.

(e). How do you ensure the quality of data during the lockdown?

- It is difficult to evaluate the quality of CPI compiled during COVID-19 lockdown. It is, however, possible to compare it with the normal CPI. Furthermore, it is important to be transparent and provide the users with information on how CPI is computed and, if possible, flag sub-indices that are imputed or with a very high share of imputed prices.

(Compiled by Isidore Kahoui and Negussie Gorfe).
SECTION 1.2 UPDATING AND MODERNIZING THE 2004 CONSUMER PRICE INDEX MANUAL

Brian Graf

Senior Economist, Real Sector Division

International Monetary Fund

SUMMARY OF PRESENTATION

The presentation provided an overview of why the manual had been updated, background, progress to date and next steps, key changes and the way forward.

The primary objectives of updating the manual included: providing clearer, more prescriptive recommendations and guidelines wherever possible; incorporating feedback and experiences with the 2004 manual; incorporating development in methods and practices, and theory and research since 2004; updating material on data sources, data collection methods and related calculation methods; reflecting development since 2004, such as e-commerce and digital economy, and emerging data source such as scanner data and web scraping; and reflecting evolving user needs.

The need to update the manual had been agreed on during the 2014 meeting of the Group of Experts on Consumer Price Indices, organized jointly by ECE and ILO. In 2014, IWGPS endorsed the updating of the CPI manual, with IMF as the lead agency. As the manual would be presented for endorsement by the Statistical Commission as an international standard, the 2004 manual was split into two publications – Consumer Price Index Manual: Concept and Methods and Consumer Price Index Theory. It had benefited from multiple layers of review before it had been endorsed as an international statistical standard by the Statistical Commission in March 2020.

The key updates of the manual include: being more prescriptive; reflecting advances in technology that have given rise to e-commerce and the digital economy; reflecting improvement made to the concepts and methods used to compile CPIs; reflecting evolving data user needs; ensuring broader consistency with the 2008 System of National Accounts; eliminating repetition and ensuring consistency across chapters; incorporating advice provided in the Practical Guide to Compiling Consumer Price Indices into the manual; concluding each chapter with a summary of key points and key recommendations; and standardizing terminology to create more uniformity and consistency across countries. The manual includes three new chapters, some of which required more updating than others. In addition to updating examples to illustrate the concepts and methods described in the manual, new examples have been added to further reinforce key recommendations.

The presenter concluded with a discussion on the way forward and planned outreach to promote the manual. He clarified that the Theory publication would be finalized by the end of 2020; however, because the Statistical Commission did not endorse theory, it would not be presented to the Statistical Commission for endorsement.

QUESTIONS AND ANSWERS

(a). Are we going to get a hard copy of the manual and in different languages?

- Countries hope that there will be more than one copy sent to each national statistical office. It takes time to translate and review, and IMF is making the translated version available as
soon as possible, with the goal of releasing Arabic, French, Russian and Spanish versions by June 2021.

(b). What are the new subject matters that have been addressed in the new manual?

- Chapter 2 contains a discussion on consumption in the 2008 System of National Accounts and CPI. The System of National Accounts provides a conceptual framework for CPI, and guides us when we have issues.

- Chapter 5 provides a detailed discussion on all methods, including telephone collection.

- Chapter 10 on scanner data provides a detailed overview of how to implement scanner data into CPI, including a discussion on appropriate formulas to compile an elementary index using scanner data. The topic of new sources of information remains dynamic as the situation develops, so it is an area that will continue to develop and be the subject of further discussion and research.

- Chapter 11 is about special cases such as e-commerce issues, owner occupied housing, seasonal items.

(c). How can national statistical offices prepare to adopt the Classification of Individual Consumption According to Purpose (COICOP) 2018 from COICOP 1999?

- In chapter 2, the differences between the two are discussed. A few countries have adopted COICOP 2018. Issues are raised based on experience. The change affects CPI, so there is need for descriptive guidance.

- Two countries in the past months have implemented COICOP 2018. They didn’t update the series backwards. However, in accordance with the manual, countries can compile internally back one year – IMF will have more definitive guidance in the coming months. The manual focuses more on COICOP classification. Match between Central Product Classification and COICOP can be discussed as part of the research agenda.

(d). What is the best method for owner occupied housing?

- There is no consensus on methods for owner occupied housing. The manual presents the various approaches (such as rent equivalent, net acquisitions approach, user cost) and the advantages and disadvantages of each method. Countries need to consider their own circumstances and decide. This topic is included in the research agenda.

(e). What kind of capacity-building and technical assistance will IMF provide?

- IMF provides capacity-building to improve price statistics and will conduct seminars on the new manual. Initially, given the current pandemic, outreach activities will be virtual ones with the plan to conduct in-person training when safe to do so. Countries can request technical assistance from IMF (see annex II).

(f). What happens to the Practical Guide with the publishing of the manual?

- The Practical Guide is discontinued. It has been incorporated into the new manual which, going forward, will serve as the single international standard on compiling CPIs. Significant effort has been made to ensure that the manual meets the needs of all countries, including issues specific to developing countries (for example, how to collect prices that are subject to bargaining).

(g). Has the manual included one or more chapters on the mechanisms of integration of the International Comparison Program into consumer price indices?
• There is an International Comparison Program working group that is dealing with the synergy between the Program and CPI. The working group will issue its recommendations to IWGPS, which will review them. An appendix on ICP is included in the manual.

(h). Is there guidance on data collection and compilation for CPI in a crisis such as COVID-19?

• Continuity notes are issued. CPI continuity plans in general shall be in place not only for COVID-19, but for other potential disruptions such as natural disasters. COVID-19 has brought the continuity issue to the forefront.

(Compiled by Ali Yedan and Sheng Zhao).
Chapter 2: Country experience and practice

SECTION 2.1 HOW TO PRODUCE A CONSUMER PRICE INDEX IN A COVID-19 CONTEXT: THE FRENCH EXPERIENCE

Marie Leclair

Head of Consumer Prices Division

National Institute of Statistics and Economic Studies, France (INSEE France)

SUMMARY OF PRESENTATION

The presenter discussed the consequences of the lockdown due to the COVID-19 pandemic. This crisis had resulted in the disappearance of many consumption segments and the adjournment of the price collection in the field, thus, no information could be obtained on prices. To confront the problem, very useful guidelines had been released by international organizations (Eurostat and ECE) on new ways to collect data during the crisis, on imputation methods in case of missing prices or in the disappearance of consumption segments.

Among the new processes to collect prices during the COVID-19 crisis, INSEE France had used scanner data, online price collection and price collection by phone more than usual. The main idea was to keep the basket of products used before the lockdown and to find the prices of these products using one of the new collection processes. This was done carefully because price levels could vary from one data source to another (especially with scanner data, owing to the different treatment of special offers or online prices) and the quality of the products sold might differ.

The presenter pointed out the difficulties anticipated when returning to the field to collect prices after the lockdown, which included not being welcomed in the outlets, the impossibility to touch the products and therefore to collect crucial information such as quantity, many queues, closed shops and important changes in the pattern of consumption.

With regard to imputations, the presenter highlighted four methods that were followed in accordance with Eurostat guidelines: estimation based on available prices for the same product; nearest aggregate estimation (for example, fast food and takeaway food services were used to impute restaurants); estimation based on the all-item index (when there was no similar consumption – tourism, cars, craftsmen); and the carry-forward method (mainly for products purchased on an annual basis). When using these methods, certain special issues could arise, for example, with highly seasonal products (plants, tourism), prices collected in advance for services (airfare, train, among others), and issues to maintain the representability by type of outlets and products.

The presenter pointed out some results to show the shock on the consumption structure and some lessons learned. One important result was caused by the closure of some outlets (restaurants and museums, among other things), which changed the consumption habit of households – more lunching
at home and less transport use. In order to assess the impact caused by the crisis, INSEE France proposed a new casting exercise twice a month, by using new available data sources that were compared with more traditional data sources in a macroeconomic model.

The main lesson learned was that both Laspeyres-type and Paasche-type indices may be useful during the pandemic. The importance of communicating to the users about the quality of data was also stressed.

**QUESTIONS AND ANSWERS**

(a). How do the new sources of data affect the basket?
- We can try to keep the basket fixed even if we use new sources. But we have to double check our data in order to ensure that the methodological change doesn't interfere with the measure of inflation.

(b). How do you use month-on-month change from the previous year for seasonal items?
- It is an option proposed by EUROSTAT to use the month-on-month change from the previous year to adjust the month-on-month change of the current year, and it is good to use it to capture the seasonal change of this year (2020). But we have to take into consideration the availability of long time series to apply that method.

(c). How do we communicate with users about the publication of two indices?
- We produced two indices, Laspeyres and Paasche, but both needed clear notes about methodologies and warning notes if any. We computed the Paasche index only for illustration purposes and not to replace the official Laspeyres index. Changing the weights of the index too frequently and chain-linking them, have their own drawbacks and we may need to be careful.

(d). How do you link results from scanner data and traditional CPI?
- In the context of the COVID-19 crisis, we used scanner data in order to find the prices of specific products whose price was collected previously in the field. We were able to do that because for each product in our basket we also collected the barcodes. But once you have matched the price collected in the field and in scanner data, you have to be really careful and check manually that there is no issue about a difference in the measure of the level of the price.

(e). What data quality assurance protocols have you put in place to build the confidence of data users?
- To give confidence to the users, we always give maximum information about the methodology and the assumptions that enters in the production of the indices. Show the consequence of the assumptions and give as much information as possible about the quality of data and the different issues we face in the production of CPI. We can also try to produce different indices with a variety of assumptions and make clear the consequence of those assumptions to our users.

(f). How do you deal with the missing prices from minimarkets?
- In the French CPI, we tried to be representative for each consumption segment cross-type of outlets, but because of COVID-19 we have less information than usual. Therefore, we were no longer representative at that level but still representative for the whole products sold in minimarkets or for a whole consumption segment (without taking into account the type of outlet). Furthermore, our estimation for each product in supermarkets and hypermarkets was very precise as we usually used extensive scanner data for these segments.
Consequently, we performed a double imputation for minimarket. First, we calculated the index for minimarkets. Second, we imputed the index for a consumption in minimarkets by the index for this consumption segment in supermarkets and hypermarkets. Lastly, we adjusted all these indices (consumption segment cross minimarket) homothetically, such that their aggregation was consistent with the minimarket index that we computed at the beginning.

(g). How can we do quality adjustment when the price is observed from online but on a pure player website (businesses that only sell products online)?

- We can make the usual quality adjustment following the bridge overlap method, meaning that all the prices we collected in April from a pure player website were not used in the inflation computation of April, but we used them for the following month index. For the prices collected in this way, we anticipated that price collection online would continue for more than one month.

(h). How can national statistical offices make decisions for a crisis such as COVID-19?

- COVID-19 is an unexpected crisis. National statistical offices have to make quick decisions and will regularly perform risk analysis and prepare action plans in case we fail to get certain data. COVID-19 has brought up the importance of new products, such as masks and gloves, which were not in our basket as their weights were below a certain threshold. As a result, we may have to monitor how the situation develops and consider the inclusion at some point.

(Compiled by Emmanuel Ngok and Tesfaye Belay).
SECTION 2.2 UNITED KINGDOM CONTINGENCY PLAN FOR PRODUCING CONSUMER PRICE STATISTICS DURING THE COVID-19 PANDEMIC

Michael Hardie

Deputy Director, Prices Division
Office for National Statistics, United Kingdom

SUMMARY OF PRESENTATION

The presentation covered topics on consumer price statistics in the United Kingdom, the contingency plan overview, response rate and consumer price statistics for April.

The main indices compiled in the United Kingdom are indicated to be CPI, including owner occupiers’ housing costs (CPIH) – currently the main measure of inflation for ONS UK; CPI – the ONS UK Harmonized Index of Consumer Prices, which is produced for Eurostat; and Retail Prices Index – a historic measure which ONS UK is required to produce by law. The ONS UK contingency plan covers CPIH, CPI and the Retail Prices Index. The plan has been developed in line with international best practice provided by Eurostat, and advice provided by the Advisory Panel on Consumer Price Statistics.

The main COVID-19 challenges were indicated to be the shut-down of large parts of the economy, resulting in certain items becoming unavailable for consumers. Furthermore, price collectors were unable to undertake the usual physical price collection in stores owing to the implementation of social distancing policies and movement restrictions brought into effect on 23 March 2020. This mode of collection accounts for approximately 80 per cent of the price quotes used in the consumer price statistics for the United Kingdom.

The collection of prices was therefore restricted to available items, ONS UK extended the collection period from “index day” to “index week”, and price collectors priced as much as possible through websites and via telephone. The imputation methods used, response rate for CPIH and the results of the consumer price statistics for April were presented.

QUESTIONS AND ANSWERS

(a). How can we make use of the “decision tree”?

- The decision tree is a useful system that national statistical offices can take as a reference, even though countries have different markets. The rationale of the decision tree is for the unavailable item imputation not to have an impact on the headline rate as much as possible. If the item is not seasonal, we impute based on the monthly growth rate of the all item index for the available items. If the item is seasonal, we use the annual growth rate because of the seasonality of that index. So, the idea behind both imputation methods is to ensure that unavailable items are not driving any changes in headline inflation.

(b). During a crisis such as COVID-19, how can we tackle the issue of fewer items in the basket?

- One option is that we can compile an “experimental” series using a rescaled basket. At a later time, we may consider re-weighting the basket (excluding the unavailable items) and using a chain-link.

(c). What factor is considered when rescaling the basket?
In terms of the basket in the United Kingdom, it is set at the start of the year and ONS UK uses a range of expenditure data primarily from National Accounts that sets the basket for the year. ONS UK does not plan to update the basket until 2021 because timely, robust and detailed expenditure data are not available.

\textbf{(d). Why do we make use of scanner data?}

- Data are collected by retailers when consumers purchase any goods and services from them. The benefit of scanner data is that it not only collects extensive price information, but the quantity of the product purchased. This allows us to accurately measure changing consumer spending patterns.

\textbf{(e). What measures can be taken to avoid artificial price variation when the collection period is extended?}

- Amid the COVID-19 crisis, the collection period was extended from index day to index week. However, ONS UK will review the collected data and potentially exclude some outliers and prices that are not plausible.

(Compiled by Ali Yedan and Elias Fisseha).
SECTION 2.3 BEYOND CONVENTIONAL IMPUTATIONS: CHALLENGES FOR CONSUMER PRICE INDEX COMPILATION DURING COVID-19 LOCKDOWN

Patrick Kelly

Chief Director of Price Statistics

Statistics South Africa

SUMMARY OF PRESENTATION

The presentation covered the impact of the restrictions on consumer spending in South Africa, the challenges to normal CPI operation and compilation, standard imputation methods, new imputation methods, application guidelines and the calculation of an essential products CPI.

It was indicated that due to COVID-19, hard lockdown was imposed on 27 March 2020, but some easing of the restrictions has been applied since May. The lockdown resulted in the unavailability of certain products or services for sale, a limited number of outlets to be open and travel restrictions, which curtailed purchases, price collection and office activities. Other impacts include the use of a reduced sample and non-traditional collection methods, threat to matched model method, high number of imputations (as standard imputation methods might not be adequate) and delays in the publication of results. In addition, there was concern for the health and safety of data collectors.

The presenter discussed the standard imputation methods used in CPI compilation and the new methods used. The new methods include the use of monthly change of all items to impute banned products, the use of monthly change from year ago and the use all items annual rate. The use of the all items index neutralizes the impact of “non-expenditure” items on monthly CPI, it is economically meaningful and it does not bias the aggregate index. Its disadvantages include: a problem if too many months are imputed, breaks in seasonal patterns and possible contradiction of elementary index trends. The advantages of use of monthly change from year ago are that it only uses data from that specific elementary index, it maintains seasonal trends, and keeps the annual rate stable. A Netherlands study found it to be best performing (its disadvantages – having a problem if too many months are imputed and assumes price trends from the previous year – still apply).

In the applications of the new methods in South Africa, the issues that need to be considered include how much of the index is imputed, whether the CPI is of adequate quality, capacity to change computer programmes, error risk in using Excel or changing systems, proper documentation, impact on index of multiple months of imputation, and transparency and stakeholder communication, including the extent of imputations.

In addition to the April publication of traditional CPI figures, Statistics South Africa has produced and published a consumer price index of essential products having weights less than 20 per cent of CPI basket of goods. The presenter gave an overview of the Weekly CPI Project (EP-CPI), which published a CPI of essential products. EP-CPI has been particularly welcomed by users in the food and agriculture sector, and has created good visibility for Statistics South Africa. In addition, the project carried out two COVID-19 special surveys. However, in undertaking EP-CPI, instructions to the team were not consistently applied, no quality assurance measures were taken, all calculations in Excel might have had potential for errors, and there was uncertainty about the economic meaning and comparison with CPI changes.
(a). How many months are possible to consider for imputation prices?
   • Normally, we should not impute price for more than two months if it is temporarily unavailable. But in this time of uncertainty, sometimes even if it is not available for more than two months, we recommend that imputation continue as there is hardly an alternative. There are two sub scenarios: consumers are not spending at all; and consumers are spending but it is difficult to collect. With the second scenario, it is important to think what the alternative ways are that we can use to collect the price data.

(b). How can we address the issue of when some seasonal products experienced a disturbance in the previous year?
   • It is possible to use averages for several years versus using a single year’s data. This needs an assessment of data to make a decision. In case of some extreme events, for example, drought, it is not a bad idea to average a couple of years back. As long as it is reasonable and communicated with the user community, it is fine.

(c). Have you updated the structure of the basket during this time of COVID-19?
   • It is not advisable to make changes on the structured weights of CPI. There is not much information available during this time. Nonetheless, the situation will be monitored in a dynamic manner and may be discussed at some point if a change is necessary. At present, the weights used by Statistics South Africa are fixed and updated and based on Household Expenditure Surveys or other sources.

(d). What shall we do to make sure there is no confusion when publishing results?
   • It is always important to make methodology notes public on the website, so that people can understand the results properly.

(e). What is the best way of compilation when some items are not available?
   • If consumers are not allowed to purchase certain items (e.g., transport between provinces), headline CPI is more appropriate than Class-Mean. For items sold in the market, we have to have the price at least for two consecutive months.

(f). Can you elaborate more on the impact that imputation has on CPI?
   • In the case of Statistics South Africa, overall weight of the indices that is subject to imputations is approximately 25 per cent (by which the presenter means “high”). But it does not bias the overall index because Statistics South Africa uses the average change of price of all other indices. So, the overall headline index is not going to be negatively affected or biased.

(Compiled by Negussie Gorfe and Sheng Zhao).
SECTION 2.4 CONSUMER PRICE INDEX COMPILATION IN TUNISIA DURING COVID-19 LOCKDOWN

Nejib Haouech  
Price Statistics Expert  
Statistics Tunisia  

SUMMARY OF PRESENTATION

Since 20 March 2020, Tunisia has been under general population lockdown due to COVID-19, which has resulted in the suspension of fieldwork with the exception of reduced-price surveys for CPI compilation. Accordingly, certain challenges had to be met, including the security of staff (price collectors), the sufficient number of observations, regional coverage, price imputations, the quality of the index, compliance with delivery deadlines and better communication.

To this end, some exceptional measures have been taken at the health, administrative and technical levels. These include the wearing of masks, the signing of travel authorizations, the reduction of outlets (points of sale) and the number of price statements.

Data collection was flexible during lockdown and alternative ways were used, such as the Internet, telephone, email, newspaper advertisements, and requesting friends and family to report the prices when they do shopping.

It emerges from this situation that the field survey covered only 52.8 per cent of the prices collected, and only 55 per cent of the outlets. The missing prices were imputed according to the guidelines from the Consumer Price Index manual. In addition, the introduction of online sales was introduced for certain products (the introduction of new varieties must be done carefully – we should separate the delivery services from prices of products).

The imputations of the missing prices were made in accordance with international recommendations, specifically by using the average of the price variations of products of the elementary aggregate which are available, or if the index of an aggregate is entirely unavailable, the imputation is then based on the aggregate which is at the nearest higher level to it.

One of the main challenges was the dissemination and communication of the results. Communication had to be improved by adding notes on the collection status and the description of the imputation methods.

Data dissemination should preserve index structure and publish the usual level of detail. For metadata, all imputed indexes should be flagged for users, with an indication of the number of missing and imputed prices by group and the proportion of outlets available for collection.

For the way forward, it is planned to improve the collection of data by integrating the use of scanner data and Internet data (web scraping), and to improve information technology infrastructure as well as the legal framework.

(Compiled by Emmanuel Ngok).
SECTION 2.5 CONSUMER PRICE INDEX AND CONSUMER PRICE INDEX COMPILATION IN AFRISTAT MEMBER STATES DURING THE COVID-19 PERIOD: CURRENT SITUATION, CHALLENGES AND MAIN RECOMMENDATIONS

Yankhoba Jacques Badji
Price Statistics Expert
AFRISTAT

SUMMARY OF PRESENTATION

The presenter covered the history of the Economic and Statistical Observatory for Sub-Saharan Africa (AFRISTAT), the state of CPI production, price collection and management of missing prices, products without supply and demand, new products, publication and dissemination of results, challenges faced and main recommendations.

AFRISTAT has been mainly involved in the development of methodological documents and in the provision of training and technical assistance (see annex II). The state of CPI production in AFRISTAT member States includes price collection from open outlets through direct interview and from temporary closed outlets through outlets replacement rule, which is relaxed and continues to collect prices online.

For the management of missing data, two types of varieties are considered: heterogeneous varieties, which are mostly manufactured products and services; and homogeneous varieties, which are fresh and unprocessed products similar from one outlet to the other. To impute missing price data of heterogeneous and homogeneous varieties, the iterative estimation method and the exogenous method are used. With regard to iterative estimation, the methods described in the Harmonized Index of Consumer Prices Methodological Manual in the West African Economic and Monetary Union zone (or the Central African Economic and Monetary Community) are preferred.

The iterative estimation method for heterogeneous products is either the monthly average evolution of the other series of the variety in the same type of outlet, or in other points of sale, or of the higher hierarchical aggregate, and exogenous method is obtained from auxiliary information. For homogeneous products, the iterative estimation method is the average unit price for same day, same week or same month; while the exogenous method is the evolution of the higher aggregate. In the case of products that are without supply and demand – the services at outlets such as hotels, air travel agencies, bus stations, concert halls, schools, bars and restaurants – the prices are imputed in accordance with the methodology applicable to heterogeneous varieties. The presenter indicated that there was no introduction of new products in the baskets of member countries during COVID-19.

The presenter pointed out that CPI is published for all levels of COICOP, whether the prices of products are collected or imputed. The weights are kept constant for all levels of aggregation and the CPI has been disseminated to users. However, some delays were observed with regard to deadlines in some countries due to restrictions on travel between regions and poor means of communication. The need to disseminate metadata to users in the interest of transparency and maintaining trust was also indicated.

The main challenges during this period of COVID-19 have been the closure of certain points of sale; the need to protect investigators and respondents from the risk of contamination; the protection of the personnel in charge of the production of CPI; and some logistical aspects to ensure remotely price
collection and remotely staff work. In order to mitigate these difficulties, it was recommended, among other things, to relax certain technical rules such as the replacement of outlets, and to use the imputation methods provided in the methodological guide and include metadata in the publication in order to guarantee transparency.

QUESTIONS AND ANSWERS

(a). Could we have an idea about the rate of imputed prices due to this COVID-19 period? Are the methodological changes and the high imputation rate induced during this period not likely to pose problems of representativeness of the basket of products and point of sale and of inducing problems of break in the time series of CPI? What impact would it have on the CPI forecast and the techniques to be used to adjust it?

- The rate of prices imputed in April 2020 varies from 2 to 45 per cent depending on the country.
- There is no change in methodology during this period, as these cases of missing data are provided for in the Methodological Guide. The representativeness of the basket could be affected if the products in question were to undergo variations which were not properly taken into account.
- The impact that it would have on the forecast of the index cannot be known a priori. But, if such a possibility arises, data for the COVID-19 period would be considered as an outlier and discarded when making the mid-term forecast.

(b). With regard to the methods of imputation presented for the homogeneous varieties, should we not be particularly concerned with the methods of imputation of prices for seasonal products (fresh fish, fruits and vegetables), given that the index is calculated for these seasonal products using the geometric mean formula?

- The variation of the geometric mean of the indices of the other varieties that are of the same aggregate must be used for the imputation of a seasonal product without quotation in the month instead of arithmetic mean.

(c). With the Phoenix tool, which iterative method does the application plan to use for these seasonal products? In addition, what is the number of years to be taken into account if the estimation of the prices of missing seasonal products is made on the basis of changes in previous years?

- The Phoenix tool allows you to use iterative methods from M1 (same outlets at the same day) to M6 (other outlets within the month) for seasonal products with some quotations. For the imputation of the prices of seasonal products without quotations in the month, we will favour the change of the same month of the last year. If the last year is abnormal, the average change of the same months of the last three years will be applied. In all cases, it is the price managers who decide the best period taking into account their knowledge of the environment.

(d). When you are talking about the new products, what prices will be considered on the base period?

- In the case of the introduction of a new product, the base price is estimated applying the price change of a similar product. In case there is none, the base price is estimated by applying the average change in the higher aggregate to the current price. The introduction of a new product should not disturb the trend of the hierarchical aggregate to which it is attached.
(e). Why did you use the Methodological Guide to the harmonized consumer price index to impute missing prices rather than the CPI manual?

- The Harmonized Index of Consumer Prices Methodological Manual takes account of the imputation methods recommended in the CPI manual. In addition, the associated Phoenix application integrates with the imputation methods of the Methodological Manual.

(Compiled by Emmanuel Ngok).
Chapter 3: Data harvesting from the Internet and website

SECTION 3.1 PRICE POPULI: ASSISTING COUNTRIES WITH CONSUMER PRICES INDEX DATA COLLECTION

Andrew Baer

Senior Economist, Real Sector Division of Statistics Department
International Monetary Fund

SUMMARY OF PRESENTATION
The main topics covered in the presentation, included challenges faced, pilot programme, timeline, use cases, crowdsourcing option, access control and respondent data, date of price reports, data transfer process and longer-term plans.

The presenter indicated that the challenges faced by many countries, such as struggling to produce timely and high quality official statistics during the COVID-19 pandemic, traditional in-store price collection is not an option in some countries that have enacted social distancing requirements, and using telephone collection which anecdotal information indicates response rates to be down.

The presenter indicated that in order to meet these challenges, a pilot programme to build a simple data collection instrument hosted on the Internet is being launched. The pilot programme would supplement the current CPI collection and solicit volunteer countries to provide a limited number of specified foods and essential items from their current CPI basket. The pilot countries include Bahrain, the Gambia, Mexico, Sierra Leone, South Africa and Sri Lanka. Countries may choose to employ the tool to assist current CPI price collectors, to expand data collection to a trusted group and to offer the tool to the general public through social and traditional media marketing campaigns as a true “crowdsourcing” effort, which is branded as Price Populi – designed to motivate the general public to participate. The presentation also covered in detail how the Price Populi works that included access control and respondent data, date of price reports, data transfer process and longer-term plans.

QUESTIONS AND ANSWERS
(a). What is the recommendation for particular products that are sold in variable volume, for example, rice? In addition, in many countries, products are sold in non-standard units.
• The collection instrument is being modified to allow an option for respondents to specify the volume of particular products purchased. IMF can work with the country to implement this requested customization.

(b). Does the Tool have semi-automation options?
• Semi-automation options can be a way for the improvement of the Tool, for example, the user can simply upload the itemized receipt and the information can be automatically scraped. This will be an interesting area for future research and improvement, but for the
time being, it is not available due to time constraints and the flexibility countries will need for their own baskets.

(c). How do we control the quality? In many situations, one product is sold in different qualities in the market.

- Countries should provide products from their current CPI basket with specified characteristics so that quality is held constant as much as possible. If there is some judgement going beyond what is described in the Tool, IMF can collaborate with countries to see how to account for that.

(d). What is the final purpose of collecting the prices by IMF?

- The purpose of the Tool is to assist countries to collect CPI but not to do research on the collected prices at IMF. IMF will study how these tools can help countries to collect prices.

(e). How does the Tool work?

- This Tool would be on the web and we could pool data weekly so that whatever data users submit on the website, IMF will transfer it to the secure folder from which countries can download data onto their infrastructure. But in terms of outlier detection in the data, we would assist as requested, but that would be the responsibility of the collecting countries. How to process this data is the countries’ responsibility. The pilot countries to use the Tool are in direct communication with IMF through email. The plan is to provide country data every week.

(f). What are the requirements for using the Tool?

- The Tool needs Internet connection and it is not with GPS tool. It doesn't allow people to report whatever type of price, and IMF can work with countries if there is a request for data filtering. Countries identify items to be collected and included in the Tool. If an item disappears in the Tool due to availability, countries can request to replace anytime. The Tool is completely free for interested countries.

(g). What are the specific functions of the Tool?

- There is no interface to review submitted prices, and on the download side, it will be in Excel spreadsheet; it is a price collection tool only and is not able to calculate CPI; there is no function to match with import or export data; and the Tool can detect duplicates from the same submitter.

(h). Which option of “use cases” is the best?

- It is a matter of how a country is going to use it. For example, Option 3 offers the tool to the general public through social and traditional media marketing campaigns as a true “crowdsourcing” effort. However, it requires greater resources commitment for a country to market the website through social media in the country and to invest in data filtering. It can be a good option if the country has the capacity.

(Compiled by Ali Yedan and Elias Fisseha).
SECTION 3.2 INTRODUCTION TO WEB SCRAPING USING R

Randi Johannessen

Head of Price Statistics

Statistics Norway

SUMMARY OF PRESENTATION

The presentation covered what web scraping is, how a web page is created, R programming language and examples of how to scrape by using R. The presenter shared the experience of Statistics Norway regarding the online harvesting of prices of consumer articles for the purpose of compiling CPI. In her presentation Ms. Johannessen spelled out the principles guiding web scraping exercises and showcased a specific example of price harvesting from a retailer web page.

The pros of web scraping include, the collection of more prices in less time, better quality and less rework, specialists now use the robot tool, work is more interesting, there is no need for organizational changes, and web scrapers are suited for many prices on few websites. On the other hand, the cons are indicated to be not feasible to build a robot for every single site, as this can be too expensive for monitoring and maintenance purposes, and while in traditional price collection prices are verified during collection, web scraped prices need to be controlled after collection. Some websites may even be “closed” for web scraping.

The presenter covered what has to be done before scraping a website, how a web page is created, the use of R programming language, which is an open source, and provided a demonstration of the programme. It was pointed out that the existing RVest package is a built tool for web scraping and can be easily adapted to the user’s own context, and the outputs of the web scraping can be formatted and mapped with required COICOP categories for integration into the CPI compilation process.

QUESTIONS AND ANSWERS

(a). How do we link results of web scraping with COICOP structure?
   • COICOP structure should have the scraped items already. The scraped data need to be structured so as to correspond to COICOP.

(b). What do we have to pay attention to when we use web scraping?
   • After web scraping, it is very important to always check the prices after collection (as contrary to conventional data collection, which we can check during the process).

(c). How can we export data from web scraping to other statistical analysis software such as STATA?
   • Other statistical analysis software such as STATA can be used to process web scraping data. Text files can be read into STATA, it just needs some structuring on the columns.

(d). How can countries benefit from web scraping when local prices are not posted to the website?
   • There is nothing to scrape if there is no price on the website available. In this case, we can ask them to send data rather than going there physically. In addition, the usual practice is to scrape national prices for the whole country, not any “local prices” for a specific region.

(e). What kind of errors do we expect to come from web scraping data?
   • To date, there hasn’t been so much error noticed for scraping, but during the pandemic price availability will be a problem (for example, suspension of flights so prices may not
be available at websites). An additional note on flights: International flights are from the consumer’s perspective booked normally several months in advance, so prices for a specific flight are scraped many times – five months, four months, three months, two months, one month, two weeks, one week prior to departure date. An international guideline by Eurostat is available to deal with flights and other "moving prices".

(f). How should we deal with the difference between web-scraping prices and real transaction prices?

• It is always superior if we can get transaction prices, but it is not always possible. So, web scraping is an alternative.

(Compiled by Negussie Gorfe and Sheng Zhao).
Chapter 4: Data collection through telephone surveys

SECTION 4.1 ALTERNATIVE METHODS OF PRICE COLLECTION

Valentina Stoevska

Senior Statistician, Department of Statistics

International Labour Organization

SUMMARY OF PRESENTATION

The presenter began by explaining two main price collection methods: local price collection in which prices are obtained from outlets around the country by personal visits and by telephone; and central price collection in which prices are collected at the central office with little or no field work involved. The restriction on movement due to the COVID-19 pandemic led to a suspension of collection of data through personal visits in most of the countries. To accommodate the disruptions in personal data collection, there is need to move to alternative modes.

The switch to a particular mode depends on the availability of skilled data-collection staff, the availability of survey respondents, availability of items, and the technological and logistical capacity of a national statistical office. On a short-term basis, the solution may be a collection of prices from a subsample of outlets chosen to be representative of the full sample, and (or) a collection of prices from a subsample of products which are representative of the CPI basket.

In cases where no fieldwork can be performed, use of alternative sources is required. The alternative sources include data collection by telephone, by mail or email correspondence, from retailers’ websites and scanner data, and newspaper advertisements, catalogues and list prices. The presenter covered the advantages and limitations of all the alternative methods indicated, alongside practical issues that need to be considered for collecting prices without visiting the outlets.

The presenter recommended that in order to maintain adequate coverage and ensure the quality of the price data collected and the quality of CPI, there is a need to move to alternative price collection processes. It was suggested that the reliance on traditional methods of collecting price data should be lessened in the future – initially, in many countries, telephone enquiries were the only way to collect data previously collected by personal visits. In case of prolonged restrictions on movement, there is a need to take advantage of the new technology to collect prices and to train staff on using these technologies.

QUESTIONS AND ANSWERS

(a). How can we deal with the issue that the Internet and mobile phones are rarely available in open market, from which a large part of our CPI basket is drawn?

- The use of alternative methods of data collection may not be possible from street vendors. In the absence of contact numbers, the only option would be in-person data collection providing that price collectors have access to these open markets. When they are doing their personal shopping in these markets, they have to collect prices. Otherwise, data collectors...
may by asking friends and relatives to report the prices when they do shopping in the open markets.

(b). How can we treat the change in consumption patterns during the COVID-19 pandemic?

- The consumption pattern of consumers is changing in almost all countries, because of the various restrictions linked to the COVID-19 pandemic. IWGPS, in its Business Continuity Guidance, does not recommend changing the weight for CPI for this year [2020] because the current expenditure data are not available, and ad hoc weight adjustments are not consistent with the fixed basket approach used as the basis for compiling CPI. Furthermore, it is not known how long this lockdown will last. As the situation is evolving, and in many countries, outlets are re-opening, it is possible that the consumption pattern will go back to normal. However, if the lockdown continues for a prolonged period of time, there might be a need to conduct a new household income and expenditure survey or use other sources to find out what is the current consumption pattern.

(c). Is there a possibility of machine learning tools for structuring data from web scraping to COICOP structure?

- Right now, there is no machine learning tool for structuring data from web scraping. But Statistics Netherlands has developed a semi-automated scraping tool that makes it easier and faster to monitor online prices for selected products on the web.

(d). Is there guidance on how to compute CPI in the context of COVID-19, including on how to collect data?

- Many national statistical offices have developed guidance notes on issues relating to the impact that COVID-19 has had on CPI compilation. ILO is compiling this information from national statistics offices. Many international and regional organizations and IWGPS have prepared guidance notes on CPI. For any interested group, ILO can provide the link to these notes.

(Compiled by Negussie Gorfe and Elias Fisseha).
SECTION 4.2 CHALLENGES OF MOVING FROM FACE-TO-FACE SURVEY TO THE TELEPHONE

Jo Bulman
Living Costs and Food Survey Manager
Office for National Statistics, United Kingdom

SUMMARY OF PRESENTATION
The presentation was made by Ms. Jo Bulman, Living Costs and Food Survey Manager at ONS UK. The outline of the presentation covered Living Costs and Food Survey (LCFS), changes to the data collection process, outputs and next steps.

LCFS is designed to collect household expenditure data through a random sample of households in the United Kingdom and there are two stages to data collection. These are: face-to-face interview involving all household members with 80-minutes duration on average; and each member of the household completes a two-week expenditure diary in which the interviewers may visit respondents’ homes multiple times. The uses of LCFS are to compile a retail prices index and a consumer prices index; and to have information on the spending patterns of the population, household expenditure for gross domestic product (GDP), the effect of taxes and benefits, and on food consumption and nutrition.

In response to the COVID-19 lockdown, face-to-face interviewing was paused on 17 March 2020, and the shifting of the data collection to telephone interview was made. As a result of this shift, there was an agreed strategy to reduce the length of the questionnaire, remove the need for respondents to use show cards, and to administer the diary over the telephone-receipts and diaries posted back to the head office. The shift of the whole survey process has had an impact on advance materials, making contact, data collection instruments, response rates, use of incentives, processing of data and outputs. The presenter highlighted the responses and feedbacks obtained, compared the rates of face-to-face survey with telephone survey and the actions taken to improve telephone surveys.

It was indicated that the output helped ONS UK to understand the effects of estimation and bias, to analyse the characteristics of the responding sample; and in understanding mode effects. In addition, it helped in the analysis of diary data to understand changes in spending patterns, in particular food; and to compare high-level trends in LCFS data to other available sources of expenditures, such as retail sales inquiry and market research data.

To conclude, the presenter covered issues on how to explore online solutions during telephone interview by focusing on income data collection first, then expenditure; challenges to reduce the length of questionnaire further; development of systems to support data collection; and on diary by exploring use of Smartphones App, and browser-based versions.

QUESTIONS AND ANSWERS
(a). How can we cope with budget constraint for paying incentive for response? Is this practice ethical?
   • Paying incentive can increase response rate and the quality of data, but many countries have budget constraints. The practice itself is acceptable. In the case of the United Kingdom,
it is a gift as opposed to payment. It is given in the form of a voucher the respondents can use to buy items from grocery.

(b). Will we miss some data with reduced sample questionnaire and duration in the phone interview?
   • Yes, we may. We can identify blocks of questions that are interdependent aspects of the questionnaire and we presented them for our internal users. In addition, we can collect only more important data from the customers, and we may only ask independent questions.

(c). Will the smartphone application replace the telephone survey and face-to-face survey?
   • With the popularity of smartphone apps, they may replace the diary data collection so the face-to-face with telephone element would remain that records payment for housing cost and utilities will be covered with face-to-face interviews. It is possible that the smartphone application will replace the diary data collection but not the telephone survey.

(d). What is the percentage of response rate face-to-face vis-à-vis telephone?
   • When we moved from face-to-face survey to telephone survey, the response rate dropped. In the case of the United Kingdom, in April 2019 the response rate was 42 per cent and for April 2020, when ONS UK moved to telephone survey it was 18 per cent, which was much lower.

(e). Is there a minimum age limit for a member of the household to respond to the questionnaire?
   • There is age restriction in the questionnaire (16 years of age for the United Kingdom), and it is good to have age limits and ranges for different questionnaires.

(f). How can we collect information on food items via telephone?
   • For food data collection, it happens using the diary. Before COVID-19, respondents would attach the receipt they purchased to the diary and ONS UK would get information from the receipt, which is then manually typed into the computer programme. Thus, if the respondent purchases the good online, we will ask for the receipt and we pick up the date from the receipt.

(g). How can the collection process deal with weight or volume of food items?
   • We ask respondents to provide details of the weight or volume of good items. When the survey is carried out, face to face interviewers can provide weighing scales to respondents to enable them to weigh food items. Administering the survey remotely we are not able to do this at present, which may reduce the quality of this detailed information.
SESSION 4.3 SELECTED SURVEY DESIGN TOPICS ON TRANSITION FROM FACE-TO-FACE TO PHONE

Zeina Mneimneh

Director, Survey Research Center

International Unit, Institute for Social Research

University of Michigan

SUMMARY OF PRESENTATION

The presenter introduced the organizational structure of the Institution for Social Research – Survey Research Center. It was established in 1946 by the University of Michigan Board of Regents and is a multidisciplinary and interdisciplinary research organization devoted to the discovery of, and insight into, major issues in the social and behavioural sciences. The Survey Research Center is an international leader in research involving the collection and analysis of sample surveys, administrative and other non-survey data. In addition, its faculty specializes in cutting-edge theory and research on key questions facing society.

The presenter informed the participants of the topics that have been covered by the Survey Research Center, which included frame coverage considerations between face-to-face and telephone surveys and sampling for telephone surveys; non-response and weighting for telephone surveys; survey of consumer attitudes case study – moving from landline frame, to landline and cell phone, to cell phone only; measurement considerations in transitioning to telephone surveys; and staffing and infrastructure for telephone surveys. In addition, the presenter covered total survey error perspectives, which included measurement error, processing error, coverage error, sampling error, non-response error and adjustment error.

QUESTIONS AND ANSWERS

(a). In telephone data collection, one of the tools that we use to create a sampling frame is random digit dialling (RDD). What is the advantage and disadvantage of this tool compared with telephone directories from telephone companies or previously collected telephone numbers from face-to-face survey?

- As long as it is implemented correctly, RDD provides a full coverage of all telephone numbers in the target population, whereas a telephone directory might consistently exclude phone numbers of households with different characteristics, thus inducing to some potential coverage bias in the estimates. The quality of a telephone survey using previously collected telephone numbers from a face-to-face survey depends on the quality of such survey and on the degree of success of obtaining such numbers among the survey respondents. If there is some differential non-response in terms of the telephone number information, non-response bias may also arise. On the other hand, a pure RDD procedure will tend to produce a very inefficient sample of telephone numbers, in which a large amount of them will be non-working or inactive. Telephone directories and telephone numbers collected from previous face-to-face surveys tend to present a smaller portion of such type of numbers.
(b). In the case of telephone survey framework selection, how do you identify and treat “dual frame line”? Is it identified by telephone companies or by the surveying institution?

• Typically, you would identify them by asking respondents questions about what other telephone numbers they have and what type (cell or landline). There are a few methods to deal with dual users, among those, the two most common ways are: using a screening design, in which you only conduct interviews with dual users coming from one of the frames (landline or cell) and screen out if they come from the other frame; and weighting adjustment, in which you adjust the weights for the fact that dual users have a higher probability of selection than cell-only or landline-only respondents.

(c). The landline survey is conducted within the household. If any member of the household responds, does it lead to selection bias? What is the implication?

• If you are using a landline phone number, you will need to do household listing in which you list all members of the households and then randomly select one. The randomly selected individual becomes the respondent that needs to be interviewed.

• If the within-household selection for landline samples is not conducted using some sort of randomization and a respondent is selected using some systematic rule (such as, whoever attends the phone), there can be some selection biases, depending on the association of the survey characteristics and the selection procedure. For example, if the survey selected whoever answers the phone, in general, unemployed people will be more likely to answer the phone, and if you are measuring employment, you can be over-estimating the unemployment rate this way.

(d). In a household that a landline survey is conducted, how do we prevent the same number of the household being interviewed again by cell-phone?

• Assuming that you are using a dual frame design (landline and cell phone), the chance that the same household will be interviewed by landline and cell phone is low. The larger problem, however, is the unequal probability of selection. In this sense, a household that has a landline and cell phone has twice the probability of selection than a household with one or the other. Thus, when conducting a survey on the telephone, you can collect this information by asking how many phone lines the household and the respondent have.

• In a general population survey, in which the population tends to be very large, the probability of this happening is extremely small. In practice, we are rarely concerned about that.

(Compiled by Negussie Gorfe and Elias Fisseha).
SECTION 4.4 COVERAGE AND SAMPLING FOR TELEPHONE SURVEYS

Raphael Nishimura

Director of Sampling Operations

Survey Research Center

University of Michigan

SUMMARY OF PRESENTATION

The presenter covered sampling frames. A sampling frame is a set of materials and (or) procedures used to identify and select units and elements from the population. Sampling frames can be a set of lists, maps and procedures. It was pointed out that face-to-face data collection usually relies on clustering to reduce survey costs, while telephone data collection generally does not need clustering.

In addition, the presenter focused on coverage error, coverage considerations between face-to-face and telephone, coverage considerations between landline and cell phone, and sampling for telephone surveys. The coverage factors that need to be considered when transitioning from face-to-face to telephone includes telephone penetration and differences between populations with and without telephones, such as degree of urbanization of dwelling areas, wealth, age and gender gap. The coverage considerations between landline and cell phone need to cover dual-frame design and, depending on the cell phone coverage, to consider a cell-only design; landlines are considered household devices, while cell phones are generally assumed to be an individual’s device and individuals with more than one phone have a higher chance of being selected than those with only one phone.

The sampling for telephone surveys is done by using RDD, which uses the knowledge of the telephone system to sample banks of numbers assigned to residential service. It was indicated that the Mitofsky-Waksberg method improves the efficiency of simple RDD. In the first stage, it selects an RDD element sample of primaries and in the second stage a subsample of numbers at random from within the “working clusters”.

QUESTIONS AND ANSWERS

(a). In a country that is used to face-to-face survey but the survey is pending because of the COVID-19 pandemic, is it possible to conduct a household survey method such as the living conditions survey, to collect daily expenditure using telephone survey?

- In general, telephone survey usually takes place for shorter surveys. The survey shouldn't take more than 20–30 minutes, but household surveys usually take a much longer period (45–60 minutes). In case national statistical offices want to conduct household surveys through telephone, we might need to shorten the questionnaire.

(b). Concerning the nationwide survey using the telephone survey method, as there is no cluster, is it possible for regions and States?

- It will depend on how we can identify telephone numbers in this geography. For example, the area code in the United States is a very good proxy of geography, which can be used for stratification to allow for regional and state-wide estimates. This will depend on the telephone system in each country that allows us to select a sample from a State or region.

(c). How can these considerations be applied in the case of price data collection?
The general coverage and sampling considerations apply for any type of survey, including price data collection. Some modifications, depending on the country and type of data collection, might need to be made, but it varies from case to case.

(Compiled by Negussie Gorfe and Elias Fisseha).
SECTION 4.5 NON-RESPONSE AND WEIGHTING FOR TELEPHONE SURVEYS

Raphael Nishimura
Director of Sampling Operations
Survey Research Center
University of Michigan

SUMMARY OF PRESENTATION
The presenter covered the four groups of outcomes and response rates. The outcomes include interviews, eligible but no interview, unknown eligibility, and not eligible. The responses of the survey could be complete interview, partial interview, refusal and break off, non-contact, other, unknown if household is occupied and unknown (other). An estimated proportion of cases of unknown eligibility that are eligible can be computed for response rate calculations. Six different response rate definitions, according to the American Association for Public Opinion Research Standard Definitions (www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx), were also presented.

The presenter discussed, in detail and with examples, survey weighting and general steps used in weighting; unequal selection probabilities adjustment, adjustments to be made for unknown eligibility in the cases of single class adjustment and class-based adjustment, multiplicity adjustment; frame integration in dual frame designs; within-household selection adjustment; non-response weighting; calibration, including post-stratification and raking; and trimming.

Multiplicity happens when a household has multiple telephones and therefore a higher probability of selection as it could have been selected through a variety of sample elements. Thus, the household weights should be adjusted to deal with the increased probability of selection. In order to make the necessary adjustments on multiplicity, the information about the multiplicity should be collected. The presenter indicated that the multiplicity problems could be dealt with before the data collection.

QUESTIONS AND ANSWERS
(a). What non-response weighting method can be used (response propensities adjustment) if the item of the survey has little correlation with the characteristics available but is highly correlated with unavailable variables (for example, income)?

• The non-response adjustment is only going to work for mitigating non-response bias when variables used are correlated with survey variables. So, the mitigation of non-response bias depends on the correlation. It is difficult to adjust effectively for non-response bias in reality. The second last step adjustment is recommended (calibration in the presentation). Demographic variables, such as gender, age, education, income level, in HH survey tends to be more reliable. Not missing at random non-response cannot be solved by using weights, but by statistical modelling (imputation might be useful).

(b). What is the treatment of non-response weighting if the survey is targeted for businesses and establishments? Any difference from a survey on individuals and households?

• The general weighting procedure is similar to businesses and establishments. Overall, in such type of survey there are more auxiliary variables available for non-response adjustments, but the general steps are the same.
(c). Can you please elaborate more on the “composite estimator” and in what situation is it preferred more than “single-frame estimator”?

- Generally speaking, the composite estimator approach assigns a value between 0 and 1 to a mixing parameter, which is then used to combine the sample estimates of each frame. Essentially, the sample cases from one frame receives a weighting adjustment proportional to such mixing parameter, and the sample cases from the other frame receive an adjustment proportional to the complement of this mixing parameter. The composite estimator approach is preferable when you have available information to estimate the mixing parameter that minimizes the sampling variance of the survey estimates.
SESSION 4.6 SURVEYS OF CONSUMERS: A CASE STUDY OF TRANSITIONING FROM LANDLINE TO CELL PHONE SAMPLING FRAME

Z. Tuba Suzer-Gurtekin
Assistant Research Scientist
Institute for Social Research
University of Michigan

SUMMARY OF PRESENTATION
The presentation covered the dual-frame pilot study, dual-frame field study and the transition from dual-frame to cell phone only.

The presenter indicated that the basic surveys of consumer attitudes (SCA) are monthly RDD surveys that are used to measure consumer expectations and are undertaken in a rotating panel design. The transition from landline (RDD) sampling to dual (RDD) sampling was started in July 2012, and to cell (RDD) sampling frame in January 2015. The addition of the cell phone sampling frame was considered due to the growing percentage of cell phone only users. The characteristics of cell phone only households in the general population include, adults renting their homes, adults aged between 25 and 29 years, and adults living with unrelated roommates (Blumberg and Luke, 2010).

The SCA dual-frame pilot study included supplementary cell phone samples to the monthly SCA RDD samples, its objectives were on how to weight SCA pilot data, and to determine the effect of conducting more cell sample cases (El Kasabi, and others, 2011). The project started with 100 sample lines from the cell phone frame. The conclusions from the SCA dual-frame pilot study indicated that, while due to increase in variance in weights (which is measured by 1+L) and as a result of loss in efficiency, the respondent demographic distribution were closer to the population demographic distributions reported by Current Population Survey (March supplement). Three commonly used designs were simulated to further understand the implications of using a dual frame. The simulations clearly showed that the landline only design had a larger relative difference on all measures that were considered – specifically, the differences in the younger age group were motivating a dual frame. The research staff decided to follow an overlapping design because of the cost.

In the second phase, the research staff identified methods to improve efficiency using pre-screening and investigated the effects of using single cell phone frame (Jiang, and others, 2014; Jiang, and others, 2015). The interviews from the cell phone frame provided a better sample balance on demographic distribution throughout the course of the production period. Furthermore, it allowed for interviewing a larger proportion of younger age group in addition to the increase in nationwide coverage in the United States of America.

Switching from dual-frame estimation to single frame estimation removes the complexity in the weighting method (with a single frame, the weighting steps were easier to be investigated). Although the effort (as measured by average number of calls to complete an interview) to complete an interview from a cell phone frame is higher, the operational findings depend on the observational data and there are no widely known findings from the randomized experiments.
The presenter concluded by highlighting the following conclusions: to transition from a dual frame to a single cell phone frame, reasons are simple, well-motivated, well-documented and empirically supported (substantial coverage error, complexity in weighting method); revision in the design require critical thinking on sample design and operations – changes in the sample design (from within household selection to any adult selection, household head versus any adult (18+ years of age), target population versus survey population, portability); and methods to improve efficiency require ongoing monitoring of questions for operations design (call-back rules, how to determine call back, call outcome).
References


(Compiled by Isidore Kahoui and Tesfaye Belay).
SESSION 4.7 MEASUREMENT CONSIDERATIONS IN TRANSITION FROM FACE-TO-FACE TO TELEPHONE INTERVIEWS

Zeina Mneimneh

Director, Survey Research Center

International Unit, Institute for Social Research

University of Michigan

SUMMARY OF PRESENTATION

The presenter focused on the total survey error framework, which is one of the contributions of the University of Michigan to the Regional Seminar. The framework provides assistance in understanding all types of errors that occur at various levels of survey undertakings and is likely to affect the final estimations of survey variables. Understanding the framework is essential to the design and implementation of a survey in such a way to minimize error.

The main focus of the presentation was on measurement error in the total survey error framework. It was indicated that measurement error occurs when true answer to the questions is different from reported answer to the questions. Measurement error has two components — measurement bias and measurement variance. Response errors may be due to events not recorded in respondents’ memory, respondents may misunderstand the question, respondents may forget relevant events, respondents may take shortcuts, or respondents may intentionally misreport.

The presenter also covered topics on general questionnaire design considerations when designing a telephone survey that includes the need to reduce the number of survey items to essential ones, show cards and scales, labelled scales, sensitive questions, and interviewer effects. It was indicated that telephone interviewers and the interaction between interviewers and respondents could have an impact on the survey. In telephone surveys, interviewers need to rely only on aural cues to identify whether the respondent has any issues with comprehension, retrieval of information, among other things; interviewer ethnicity, language, cultural characteristics that may have an impact on their response; larger workload magnifies the effect; in decentralized versus centralized telephone gains in increased supervision is reduced; and interviewer expectations, attitudes, and skills (some differences in skill sets for face-to-face and computer-assisted telephone interviewers).

QUESTIONS AND ANSWERS

(a). What is SCA?

• SCA stands for Survey of Consumer Attitudes, which, for a number of years, has been gathering the consumer sentiment in relation to the economy. It is a major indicator for the United States economy and has been replicated in many other countries. Results from SCA are highly correlated with other economic indicators. There is a dedicated website for SCA, and results are published every month. The sample they use include some they have interviewed before and some additional fresh samples. The survey was purely done through landline, but now the survey has transitioned from landline to cell phone.

(b). How can we apply telephone survey in CPI data collection given that it involves some of the items, and how do you deal with response bias?
With regard to dealing with response bias, this depends on the type of response bias. For example, one type of bias is called recency effect, in which respondents are more likely to select the lower item on the scale (there are different kinds of response bias). One way to deal with this bias, other than to reduce the number of response options, is to randomize the options. When there is order in the response – for example, strongly agree, agree, and so on – it will not work in this case. If the bias in question is the social desirability bias, it might help to avoid the involvement of an interviewer by pushing the question to a web or an app in which the respondents administer the interview themselves.

(c). How can we avoid bias in the household answers when using the telephone?

- The issue of recall bias is always there and keeping diary might be a possible solution. Implementing the diary in an app might work for a telephone survey. However, it is always a good practice to undertake a feasibility study before implementing the telephone survey method.

(d). In a telephone survey, couldn't the respondent's answers be biased because she or he wants to quickly end the survey? How can we solve this problem?

- In telephone surveys, there could be bias due to the fact that the respondent wants to quickly end the survey. It is an issue. However, there is not much empirical evidence that cell phone responders concentrate less than landline responders, thereby affecting the measurement quality of data. What can be done is to train the interviewer to model a certain pace that the respondent needs to follow. Rushing to answer some questions before the interviewer finishes reading them is a problem that can happen even in a face-to-face interview. Interviewers should be trained to insist on finishing the question before the respondents provides their response.

(e). How can we solve the problem of interviewer effect by cell phone?

- Decentralization helps in recruiting interviewers locally who know the culture, language and other characteristics of the respondents. However, in a telephone survey, this might create a problem in that the number indicating the locality might not agree with the owner of the phone number. In a telephone survey, matching of interviewer and respondent can be done but it depends on the facility and infrastructure that is made available. In addition, increased supervision can help to reduce interviewer effects.

(f). In a telephone survey, how can we make sure that sensitive questions do not embarrass respondents? Isn't it better in certain telephone surveys that a man interviews a man and a woman interviews a woman? Would this cause potential bias?

- If it is possible to match the interviewer and the respondent by gender on topics that are gender-related or that are affected by the gender of the interviewer, that approach should definitely be followed. In fact, gender matching might be easier to do on the telephone, but first you have to identify whether the respondent is a female or male and then transfer them to the right interviewer. This requires more logistical arrangements but can be done.

(g). For face-to-face surveys, it is possible to reduce the geographic coverage errors of “enumerated area” by having a good household numbering (before getting the household selected). What about telephone surveys?

- For a face-to-face survey, proper identification of household will greatly help to reduce geographic coverage error. With a telephone survey, it depends how you select your telephone samples. The problem is that there might be a geographic coverage problem with using the telephone if telephone access to rural areas is limited. We may need to
supplement with another frame if possible. If telephone surveying is the only option, as in the current situation, we might need to do some weight adjustment to mitigate some of the problems. But there is no 100 per cent solution for this. (Compiled by Isidore Kahou and Tesfaye Belay).
SESSION 4.8 MOVING FROM IN-PERSON TO TELEPHONE DATA COLLECTION: STAFFING AND INFRASTRUCTURE CONSIDERATIONS

Grant Benson

Director of Data Collection Operations

Data Collections Unit, Survey Research Operations

Survey Research Center

University of Michigan

SUMMARY OF PRESENTATION

The presentation covered the key considerations for transitioning from in-person to telephone data collection, which include production, supervision and support, quality assurance and respondent confidentiality.

The production adjustments include modifying scheduling strategies, contact efficiencies, and doing without non-verbal cues. Scheduling strategies include adjusting for the contact objectives, such as eligibility screening or re-contact effort. To increase contact efficiency, especially if you don’t have the benefit of in-person visits, it is more important to ensure that interviewers work through different times of the day and days of the week until contact is made. Furthermore, when doing without non-verbal cues, it is important to be prepared to make an impression during the first few seconds of an introduction; to listen carefully; to read verbatim; to have voice clarity, pace and cadence; and to have neutrality and to use probing. It was pointed out that the length of the questionnaire has an impact on telephone interviews – it increases on average by 3.6 per cent, owing to the lack of non-verbal cues.

In addition, it was indicated that social connectedness covers coping with isolation, to check in with the individual at least once weekly, to lead small group discussions every one or two weeks, and to have front-line resources in place (such as a team leader) for any questions. Some key metrics may include attrition (quitting the job), attendance (showing up late, leaving early, or not showing up at all), hours per interview, honoured appointment rate (returning a call to a household at the appointed hour), dials per hour, and using contact windows to distribute calls. The equipment necessary for conducting decentralized telephone interviews are headsets, dedicated phones and dedicated office space.

The presenter covered quality assurance, which deals with data quality consistency, data collection call-back and data validation. Quality assurance is about compliance with standards, the assessment of investigators, verification of survey conditions and content, and validation of data quality. The need to provide respondents with absolute protection against disclosure risk was also indicated.

QUESTIONS AND ANSWERS

(a). What is the rate of non-response and the rate of dropout during interview?

• If we keep the interview length under 35 minutes, we can generally keep the dropout rate to below 5 per cent. If the interview extends for an hour, we definitely get a much higher dropout rate. It depends on the type of questions being asked. In a CPI type of questions, usually around 5 per cent drop out within 35 minute interview duration.
(b). Does the gender (gender of interviewer, gender of respondents, between interviewer and respondent) have an impact on the response rate and data quality for phone survey? What are the results that you have on this issue?

- The telephone surveys are very different from the face-to-face survey. In-person surveys in general have a significant gender effect. In the United States, when sending a male interviewer, we have a higher rate of people not answering the door in the first place. In telephone surveys for general survey questions, if the interviewer speaks in a non-monotone way, the gender of the interviewer will not have any impact on the response. However, depending on the sensitivity of the questions asked, the gender of the interviewer may have an impact on the outcome.

(c). How can we manage funding for telephone surveys as they require more funding?

- For funding management, cell phones are charged at the rate of a few cents per interviewer hour in order to accumulate enough funds to purchase cell phones. These phones are used for various projects and every project needs to contribute to its refill. Getting a cell phone is not a huge cost, especially when considering the efficiency one gets from phone interviewing, compared with face-to-face interviewing.

(d). How can we deal with the issue of scamming?

- There is definitely concern about all the scams conducted over the phone, which makes it harder to establish legitimacy. At the University of Michigan, we tried to work with a company that whitelists numbers with all the major carriers. The idea of whitelisting is that it certifies a number as being legitimate and not a scamming number. In principle, this prevents our number from getting blocked. However, a recent controlled study that we conducted showed that whitelisting has no impact on contact rate at all for cell phone numbers (that is, the respondents had cell phones). During this time of COVID-19 our contact rate has doubled and response rate has gone up. This might be because people are home at all times and are willing to talk more now.

(e). In a situation when the unemployment rate is high, especially the household survey that requires us to call or send an email to the respondent, it is always difficult to get their concentration and they tend to think the survey is a waste of time. How can we handle such a situation?

- When we approach a low-income group during a survey, we explain to them that the government needs accurate data to make decisions and what we are doing is getting those data and their cooperation will help in this regard.

(f). Do we have confidentiality issue in the practice of telephone surveys?

- For cell phone survey, confidentiality of respondent is less of an issue, as cell phones are owned by the individual respondent or by a close family member. Our screener questions also focus on the characteristics of interest of the respondent and not on specific name. Our institution review board has said that this is an acceptable risk if this is done within a household and certainly if it is a cell phone interview as it is limited within a narrow group of people.

(g). Does the COVID-19 lockdown period increase stress and non-response rate compared to normal period?

- Stress levels among the general public have increased during the COVID-19 lockdown. This is mainly due to isolation and confinement in homes. However, our non-response rates have gone down.

(h). Is it practical or recommended to compensate respondents for the time lost during the interview?
Regional Seminar on Data Collection for Compilation of Consumer Price Index

- Although we would like to give a token of appreciation to the respondent for a telephone interview, if we don't have a means of mailing ahead of the interview, it has no impact on the pickup rate. It may be helpful to reference the American Association for Public Opinion Research Task Force on Telephone Interviewing (www.aapor.org/Education-Resources/Reports/The-Future-Of-U-S-General-Population-Telephone-Sur.aspx). For household surveys, it is a good idea if we can pre-mail the token of appreciation.

(i). Can respondents give us the answers we need 100 per cent via phone?

- It depends on the questions if they will be properly answered. Studies show that if we simplify the questions in a way that doesn't need respondents to refer to a respondent booklet or other pieces of paper, then yes, we get a high response rate. If we train interviewers to probe, the “don’t know” and refusal rates are not that different from in-person to telephone interviews.

(Compiled by Emmanuel Ngok and Tesfaye Belay).
Summary and conclusions

(1). The IWGPS-CPI business continuity guidance drafted by IMF and developed in conjunction with IWGPS (Eurostat, ILO, IMF, OECD, ECE, and the World Bank) provides guidance to national statistics organizations to ensure the continued publication of their consumer price indices in the face of challenges due to COVID-19. It provides specific suggestions and recommendations on how to deal with missing data, how to collect data from different outlets and sectors; what modes of data collection to be used, and what are the best practice in terms of compilation, imputation and dissemination of the CPI data. In addition, it provides some issues to consider in preparing to work remotely. The guidance has been followed by ONS UK to develop its own contingency plan, by INSEE France to compile methodological notes, and by Statistics South Africa to conduct imputation for CPI.

(2). IWGPS-CPI is preparing the 2020 CPI manual, an update of the 2004 manual. The key updates of the manual are for it to become more prescriptive, reflect advances in technology that have given rise to e-commerce and the digital economy, reflect improvement made to the concepts and methods used to compile CPIs, reflect evolving data user needs, ensure broader consistency with the 2008 System of National Accounts, eliminate repetition and ensure consistency between chapters; and incorporate the Practical Guide to Compiling Consumer Price Indices into the new manual. Each chapter now concludes with a summary of the key points and identifies key recommendations, standardizing terminology to create more uniformity and consistency throughout countries, and elementary aggregate has been defined more clearly.

(3). IMF has developed and is testing Price Populi, an online tool developed to support countries amid the challenges to maintain price data coverage for CPI compilation. This tool, which is being testing in a handful of countries, is meant to supplement current CPI collection and is optimized for mobile devices. With the focus of Price Populi on a limited number of specified food and essential items from their current CPI basket, the tool is aimed at motivating the general public to participate in price data gathering for CPI. A positive outcome of the testing could spur interest for its adoption by African countries. The manual will be supplemented by a Theory Publication on price statistics.

(4). National statistical offices are increasingly resorting to scanner data as a type of administrative source for price statistics. This source of data provides real transaction prices and feeds well among price data for CPI. It is being used in France, Norway and South Africa, and features among ILO recommended alternative sources for CPI data. Although a good source for a wide range of consumption basket items, it may have inherent selection bias and under coverage factors that need to be considered when it is used in CPI compilation.

(5). IWGPS-CPI has elaborated guidelines that advise on imputation techniques for missing data in CPI production. These techniques are applicable to situations such as the current health crisis. However, challenges arise to their systematic and full application in individual country cases. In particular, the principle of fixed weighting system and comparable basket of products remain difficult to apply with economic meaning. Adaptation measures have been taken to mitigate the loss of quality to ensure that the best possible CPI services are provided to data users, including targeted imputation methods. To go around the issue of comparability, simultaneous computation and release of Laspeyres and Paasche types of indices have been envisaged; furthermore, special group of product indices have been introduced so as to maintain the relevance of the statistics office and retain the trust of the public in statistics. Information about imputations should be
provided with the release of CPI; indices with full or significant imputations should be flagged; and if possible, to address the impact on the overall quality of CPI.

(6). Telephone surveys constitute another possible way to adapt data collection to an evolving environment. Telephone surveys for price data collection have been advised by ILO and used in France, South Africa, the United Kingdom and other countries to supplement price data. Among its advantages, the telephone survey reduces costs in comparison with personal in-store visits, as fewer price collectors are needed and there are no travel costs involved; it produces more timely results and allows quick and inexpensive follow-ups for additional information. Furthermore, it can help to reach remote places that are difficult to reach by in-person mode. Owing, however, to the burden that it places on respondents, only a limited number of products can be priced and only a limited number of product characteristics can be collected, with limited scope for verifying accuracy of data. Telephone surveys have other selection and error-generation factors that need to be closely analysed and accounted for.

(7). Data collection by email, from retail web pages (web scraping) and from magazines has been implemented elsewhere but is heavily dependent on each country’s infrastructural setting and may face challenging implementation barriers in African countries. Nevertheless, insofar as possible these approaches could be explored for their cost effectiveness for price data gathering.

(8). The survey respondent’s attitude during the interview is an important quality element to be taken into consideration. Accounting for this element needs to start at the earliest design stage of the questionnaires and be pursued during the interview phase, by considering, for example, the respondent’s burden or capacity to recall. The cue as to the respondent’s attitude or need for clarification depends on the questionnaire administration mode used, whether through face-to-face or by telephone. Many other factors that can influence the type of answers chosen by respondents need to be carefully analysed and taken into account by interviewers. The Survey Research Center of the University of Michigan has valuable resources in these topics and can be of help to countries’ survey investigations.

(9). National statistics offices are encouraged to adhere to the international standards and recommendations on CPI; to mitigate coverage issues arising from the current crisis by enlarging the coverage and comparability of CPI basic price data, making use of the alternative solutions, imputation techniques and aggregations; to follow good practices in dissemination of official statistics; to be transparent to ensure public trust in CPI by disclosing methodological methods and procedures used in the production of the CPI data; and to take advantage of technical support availed by development partners and international organizations to organize the transformation of their business processes through adoption of novel and improved approaches to data collection, processing and computation and analysis processes, which are more resilient and anchored on new technologies.
Annex I: Countries and agencies registered for the seminar with number of participants

<table>
<thead>
<tr>
<th>Sq.</th>
<th>Member State</th>
<th>Number</th>
<th>Sq.</th>
<th>Agency</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum 369</td>
<td></td>
<td>Sum</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Angola</td>
<td>10</td>
<td>1</td>
<td>African Development Bank</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Benin</td>
<td>6</td>
<td>2</td>
<td>Economic and Statistical Observatory for Sub-Saharan Africa (AFRISTAT)</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Botswana</td>
<td>13</td>
<td>3</td>
<td>African Union Commission</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Burkina Faso</td>
<td>7</td>
<td>4</td>
<td>Economic Community of Central African States</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Burundi</td>
<td>34</td>
<td>5</td>
<td>Central African Economic and Monetary Community</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Cabo Verde</td>
<td>6</td>
<td>6</td>
<td>Common Market for Eastern and Southern Africa</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Cameroon</td>
<td>13</td>
<td>7</td>
<td>Demographic and Health Survey Programme</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Central African Republic</td>
<td>3</td>
<td>8</td>
<td>East African Community</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Chad</td>
<td>3</td>
<td>9</td>
<td>Eastern Africa Statistical Training Centre</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Comoros</td>
<td>12</td>
<td>10</td>
<td>Economic Commission for Africa</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>Congo</td>
<td>2</td>
<td>11</td>
<td>Economic Community of West African States</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Côte d’Ivoire</td>
<td>5</td>
<td>12</td>
<td>Economic and Social Commission for Western Asia</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Democratic Republic of the Congo</td>
<td>4</td>
<td>13</td>
<td>Eurostat</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Djibouti</td>
<td>4</td>
<td>14</td>
<td>Food and Agriculture Organization of the United Nations</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>Egypt</td>
<td>1</td>
<td>15</td>
<td>ICF International</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Equatorial Guinea</td>
<td>15</td>
<td>16</td>
<td>Intergovernmental Authority on Development</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>Eswatini</td>
<td>4</td>
<td>17</td>
<td>International Monetary Fund</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>Ethiopia</td>
<td>5</td>
<td>18</td>
<td>National Institute of Statistics and Economic Studies (INSEE), France</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>Gabon</td>
<td>9</td>
<td>19</td>
<td>Makerere University</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>Gambia</td>
<td>2</td>
<td>20</td>
<td>Moldova National Statistics Office</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>Ghana</td>
<td>9</td>
<td>21</td>
<td>New Partnership for Africa’s Development</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>Guinea</td>
<td>3</td>
<td>22</td>
<td>Overseas Development Institute</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>Kenya</td>
<td>12</td>
<td>23</td>
<td>Office for National Statistics, United Kingdom</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Country</td>
<td>6</td>
<td>24</td>
<td>Southern African Development Community</td>
<td>1</td>
</tr>
<tr>
<td>---</td>
<td>------------------</td>
<td>---</td>
<td>----</td>
<td>----------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td>24</td>
<td>Lesotho</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Liberia</td>
<td>2</td>
<td>24</td>
<td>West African Economic and Monetary Union</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>Libya</td>
<td>4</td>
<td>26</td>
<td>Arab Maghreb Union</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>Madagascar</td>
<td>4</td>
<td>27</td>
<td>United Nations Environment Programme</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>Malawi</td>
<td>2</td>
<td>28</td>
<td>World Bank</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>Mali</td>
<td>4</td>
<td>29</td>
<td>York University, Toronto, Canada</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>Mauritania</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Mauritius</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Morocco</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Mozambique</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Namibia</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Niger</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Nigeria</td>
<td></td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Rwanda</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Sao Tome and Principe</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Senegal</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Seychelles</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Sierra Leone</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Somalia</td>
<td></td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>South Africa</td>
<td></td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>South Sudan</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Sudan</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Togo</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Tunisia</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Uganda</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>United Republic of Tanzania</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Zambia</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Zimbabwe</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Annex II: Results of post-seminar survey on areas needed for further capacity-building and technical assistance

<table>
<thead>
<tr>
<th>Country</th>
<th>Phone/mail</th>
<th>Web scraping</th>
<th>Imputation</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>28</td>
<td>28</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>Benin</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Botswana</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Updating of weights, rental surveys and Owner-Occupied Housing</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burundi</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Cabo Verde</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Cameroon</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Central African Republic</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Chad</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Comoros</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Technique de prévision</td>
</tr>
<tr>
<td>Djibouti</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Sur la rénovation de l’IPC</td>
</tr>
<tr>
<td>Equatorial Guinea</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Eswatini</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Household consumption survey methodology, data collection, processing and analysis</td>
</tr>
<tr>
<td>Gabon</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambia</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ghana</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Activity</td>
<td>Activity</td>
<td>Activity</td>
<td>Activity</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Kenya</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Lesotho</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Madagascar</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Réolution de pondération et modèle de prévision IPC</td>
</tr>
<tr>
<td>Malawi</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Mauritania</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>CPI data analysis, deriving of CPI weights, report writing, graphic design, regional/provincial CPIs compilation</td>
</tr>
<tr>
<td>Namibia</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Sao Tome and Principe</td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Senegal</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seychelles</td>
<td></td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Somalia</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Somalia – Puntland State</td>
<td>Y</td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Somalia – Somaliland</td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>South Sudan</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Designing the country production system and CPI catalogue.</td>
</tr>
<tr>
<td>Togo</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td></td>
<td>Y</td>
<td></td>
<td>Collecte de données à partir des données scanner (Scanner Data)</td>
</tr>
<tr>
<td>Uganda</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

***