

Regional Seminar on Data Collection for Compilation of CPI
Teleconference, 18-28 May 2020, Addis Ababa, Ethiopia

GUIDANCE NOTE

June 2020



International
Labour
Organization



Office for
National Statistics



Statistisk sentralbyrå
Statistics Norway



stats sa
Department
Statistics South Africa
REPUBLIC OF SOUTH AFRICA



إحصائيات تونس
STATISTIQUES TUNISIE



United Nations
Economic Commission for Africa



UNECE

Table of Contents

Background and Introduction	3
Chapter 1: International Standards and Recommendations	5
Section 1.1: Business Continuity Notes on Producing CPI during COVID-19 by the IWGPS	6
Section 1.2: Updating and Modernizing the 2004 CPI manual	8
Chapter 2: Country Experience and Practice	11
Section 2.1: How to Produce CPI in a COVID-19 Context? The French Experience	12
Section 2.2: UK Contingency Plan for Producing Consumer Price Statistics during the COVID19 Pandemic	15
Section 2.3: Beyond Conventional Imputations: Challenges for CPI Compilation in Times of COVID-19 Lockdown	17
Section 2.4: CPI Compilation in Tunisia During COVID-19 Lockdown	19
Section 2.5: Situation of CPI / CPI Compilation in AFRISTAT's Member States During The COVID-19 Period: Current Situation, Challenges and Main Recommendations	20
Chapter 3: Data Harvesting from Internet and Website	23
Section 3.1: Price Populi – Assisting Countries with CPI Data Collection	24
Section 3.2: An Introduction to Web Scraping Using R	26
Chapter 4: Data Collection through Telephone Surveys	28
Section 4.1: Alternative Methods of Price Collection	29
Section 4.2: Challenges of Moving a Face-To-Face Survey to the Telephone	31
Session 4.3: Selected Survey Design Topics on Transition from Face-To-Face to Phone	33
Section 4.4: Coverage and Sampling for Telephone Surveys	35
Section 4.5: Nonresponse and Weighting for Telephone Surveys	37
Session 4.6: Surveys of Consumers: A Case Study of Transitioning from Landline to Cell Phone Sampling Frame	39
Session 4.7: Measurement Considerations in Transition from Face-To-Face to Phone	41
Session 4.8: Moving from In-Person to Telephone Data Collection: Staffing and Infrastructure Considerations	44
Summary and Conclusions	47
Appendix 1: Countries and Agencies Registered to the Seminar with Number of Participants	49
Appendix 2: Results of Post-Seminar Survey on Areas Needed for Further Capacity Building and Technical Assistance	50

Background and Introduction

The outbreak of COVID-19 has brought big challenges to national statistical systems and their operations: while there is an increasing demand for statistics and data as the policy-makers and the public want to know how the outbreak is affecting various aspects of the economy, the ongoing social-distancing and lockdowns measures that are introduced to curb the spread of the virus have also imposed restriction on possibilities for statistical data collection activities as most of the censuses and surveys are conducted through face-to-face interviewing. While the challenges to each country might diverge due to different national circumstances and patterns, there are also 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 even after COVID-19.

At a recent teleconference with the Directors-General of National Statistical Offices (NSOs) in Africa, countries requested from Pan-African institutions and development partners an organized setting where they can share and exchange good practice and experience. They further requested advice and guidance on how to continue the CPI data collection through new and alternative approaches and techniques. In this regard, UNECA along with partner institutions, organized a Regional Seminar via teleconference on Data Collection for Compilation of CPI in context of COVID-19 in the weeks of 18-28 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.

The main results of the survey conducted with regards to the impacts of COVID-19 on the activities of the national statistical offices (NSOs) in Africa. In his opening remarks, Mr. Oliver Chinganya, Director of African Centre for Statistics, ECA, indicated that the results of the survey have been sent to the Directors-General of NSOs recently. He further indicated that exchanges of experiences are required on how to address the challenges faced by the COVID-19 pandemic. He stressed the need to look for alternative ways of data collection, that price data watch will be launched during the coming month, and the importance of having frank discussions so as to share the expertise of participants. He finally informed the participants that the report of the Regional Seminar will be shared with all of them.

The purpose of the Seminar is indicated to be 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: 1) better understand of the impacts and challenges of Covid-19 on the activities and operations of the data collection and compilation of CPI ; 2) exchange on countries' experiences of mitigation strategies and techniques; 3) identify and evaluate innovative data collection approaches and methods to ensure consistency over time and comparability across countries.

The online regional seminar spanned for 8 days, usually with 2 sessions for each day. There were 16 sessions in total, including one opening, one closing, and 14 substantive session covering 1) International Standards and Recommendations; 2) Data Harvesting based on Website and Data Science; and 3) Data Collection through Telephone Surveys. Below institutions and development partners joined ECA to give presentations:

- AFRISTAT
- International Labour Organization (ILO)
- International Monetary Fund (IMF)
- National Institute of Statistics and Economic Studies, France (INSEE France)
- Office for National Statistics, United Kingdom (ONS UK)
- Statistics Norway

- Statistics South Africa
- Statistics Tunisia
- University of Michigan (U of Michigan)
- United Nations Economic Commission for Europe (ECE)

There are 432 participants from 51 member States and 29 agencies and institutions registered in this online regional seminar.

Chapter 1: International Standards and Recommendations

Section 1.1: Business Continuity Notes on Producing CPI during COVID-19 by the IWGPS

Mr. Carsten Boldsen
Chief, Economic Statistics Section
United Nations Economic Commission for Europe (UNECE)

Summary of Presentation

The presentation gave an overview of the guidance notes for producing CPI by the Inter-secretariat Working Group on Price Statistics (IWGPS). The Guidance note provide guidance on data collection, imputation methods and publication.

For price collection from outlets that remain open alternative data sources include outlet's website, newspapers and advertisements, telephone, e-mail, in person price collection by NSO staff and scanner data. For restaurants and cafes, take-away or delivery menus may be priced if comparable to sit-down menus usually followed. For airlines, hotels and package holidays it may often be possible to collect prices from the web and/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. However, if justified, the last observed price may be carried forward for a limited period. Methods should be 'self-correcting', i.e. the index must show the correct change from the last period where prices were observed to the period where prices again can be collected.

Where 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 basked approach and will create breaks in time series.

When deciding on imputation methods, the main uses of the 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 the CPI. Documentation and explanations of methods should be made available to users. NSOs should continue to publish CPI sub-indices, also if these are imputed. Information about imputations because of COVID-19 should be provided with the release of the CPI. Indices with full or significant imputations due to COVID-19 should be flagged. If possible, address the impact on the overall quality of the CPI.

Questions and Answers

- (a) What need 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; should only be used for products with stable price development and for a limited period.
 - Document methods and procedures to ensure 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 to make use of data from on-line website?
- On-line prices may be used to replace physical outlet prices for comparable products; quality differences should in principle be adjusted for.
- (c) How to collect data from informal markets?
- In practice it a challenge to collect prices from informal markets. To address this problem different methods are followed by different countries. Alternative data sources may be used.
- (d) How to handle the prices of banned/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 scope of the CPI and should in principle be covered. However, for practical measurement reasons, this may not always be possible.
- (e) How to 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 with the normal CPI. Furthermore, it is important to be transparent and provide the users with information on how the 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 CPI manual

Mr. Brian Graf
Senior Economist, Real Sector Division
International Monetary Fund (IMF)

Summary of Presentation

The presentation provided an overview of why the manual was 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 developments in methods and practices as well as theory and research since 2004; updating material on data sources, data collection methods and related calculation methods; reflecting developments since 2004, such as e-commerce and digital economy and emerging data sources like scanner data, web scraping; and reflecting evolving user needs.

The need to update the Manual was agreed during the 2014 UNECE-ILO CPI Group of Experts Meeting. In 2014 the Inter-Secretariat Working Group on Price Statistics (IWGPS) endorsed the updating of the CPI Manual, with the IMF as the lead agency. Because the Manual would be presented for endorsement by the United National Statistical Commission (UNSC) as an international standard, the 2004 Manual was split into two publications – *Consumer Price Index Manual: Concept and Methods* and *Consumer Price Index Theory*. The Manual benefited from multiple layers of review before submission to the UNSC. The Manual was endorsed as an international statistical standard by the UNSC 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 improvements made to the concepts and methods used to compile CPIs; reflecting evolving data user needs; ensuring broader consistency with the 2008 SNA; 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 the chapters 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 presentation concluded with a discussion of on the way forward and planned outreach to promote the Manual. The Theory publication will be finalized by the end of 2020; however, because the UNSC does not endorse theory, it will not be presented to the UNSC for endorsement.

Questions and Answers

- (a) Are we going to get a hardcopy of the Manual and in different languages?
 - Countries hope there will be more than one copy sent to each NSO.
 - It takes time to translate and review also, 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 discusses consumption in SNA and CPI. SNA provides a conceptual framework for CPI. It guides us when we have issues.
 - Chapter 5 has detailed discussion on all methods, including telephone collection.
 - Chapter 10 is on scanner data provides a detailed overview of how to implement scanner data into a 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 of the Manual is about special cases like e-commerce issues, owner occupied housing, seasonal items, etc.
- (c) How can NSOs prepare to adopt COICOP 2018 from COICOP 1999?
- In Chapter 2 of the Manual, differences of the two are discussed.
 - A few countries have adopted COICOP 2018. Issues are raised based on experience. There is need for descriptive guidance. The change does affect CPI.
 - Two countries in the past months implemented COICOP 2018. They didn't update the series backwards.
 - In the Manual, countries can compile internally back one year. IMF will have more definitive guidance in the coming months. The Manual itself focuses more on COICOP classification. Match between CPC and COICOP can be discussed as part of the research agenda.
- (d) What's the best method for owner occupied housing?
- There is no consensus on methods for owner occupied housing. The Manual presents the different approaches (rent equivalent, net acquisitions approach, user cost, etc.) and discusses 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?
- Yes, 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 with the plan to conduct in-person training when safe to do so. Countries can request technical assistance from IMF.
- (f) What happens to the Practical Guide with the publishing of the Manual?
- Practical Guide is discontinued. It has been incorporated into the Manual and, going forward, the Manual serves as the single international standard on compiling CPIs. Significant effort has been made to ensure 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 (ICP) to consumer price indices?
- There is an ICP working group addressing CPI and ICP synergy. The working group will issue their recommendations and the IWGPS will review these recommendations. An appendix on the ICP is included in the Manual.
- (h) Is there guidance on data collection and compilation for the CPI in a crisis like COVID-19?
- Continuity Notes are issued. CPI continuity plans in general shall be in place not only for COVID-19 but also 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 CPI in a COVID-19 Context? The French Experience

Ms. Marie Leclair
Head of Consumer Prices Division
National Institute of Statistics and Economic Studies, France (INSEE France)

Summary of Presentation

The topic covered the consequences of the lockdown due to COVID-19 pandemic which resulted in the disappearance of a lot of consumption segments as well as the adjournment of the price collection on the field; thus, no information could be obtained on prices. To address the problem, very useful guidelines were released by international organization (Eurostat, UNECE) on new way to collect data during the crisis, on imputation methods in case of missing prices or of disappearance of consumption segments.

Among the new processes to collect prices during the Covid-19 crisis, INSEE used more than usually scanner data, online price collection and price collection by phone. The main idea was to keep the basket of products used before the lockdown and to find the prices of these products thanks to one of these new collection processes. However, this was done carefully because price level may be different from one data source to another one (especially with scanner data, due to the different treatment of special offers, or online prices) and the quality of the products sold may be different.

The presentation also pointed out the difficulties that were anticipated when returning to the field to collect prices after the lockdown which include to not be welcomed in the outlets, the impossibility to touch the products and therefore to collect crucial information like quantity, a lot of queues, closed shops and important changes in the pattern of consumption.

With regards to imputations the presentation pointed out four methods that were followed according to Eurostat guidelines which included estimation based on available prices for the same product; nearest aggregate estimation (for instance, fast food and take away food services were used to impute restaurants); estimation based on the all-item index (when there is no similar consumption: tourism, cars, craftsmen, etc.); and the carry-forward method (mainly for products purchased on a yearly basis). It was pointed out that when using these methods, one can face some special issues, for instance, (i) with highly seasonal products (plants, tourism) ; (ii) prices collected in advance for services (airfare, train, etc.) and (iii) 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 learnt. An important shock was noticed on the consumption structure due to the fact that some outlets were closed (restaurants, museums, etc.), households changed their consumption habit: more lunch at home, less transport, etc. In order to assess the impact of the crisis, INSEE proposes a now casting exercise twice a month by using new available data sources that are compared to more traditional data sources in a macroeconomic model.

The main lesson learnt was indicated that both Laspeyres-type and Paasche-type index may be useful during the pandemic. The importance of communicating to the users about the quality of the 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 to use month-on-month change from 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. But we have to take into consideration the availability of long time series to apply that method.
- (c) How to 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 purpose 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 to 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 a specific products whose price was collected before on the field. We were able to do that because for each product in our basket we collected also the barcodes. But once you matched the price collected on 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 are the data quality assurance protocols you put in place to build the confidence of data users?
- To give confidence of the users, we always give the 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 the data and the different issues we face in the production of the CPI. We can also try to produce different indices with different assumptions, and we make clear the consequence of those assumptions to our users.
- (f) How do you deal with the missing prices from mini markets?
- In the French CPI, we tried to be representative for each consumption segment cross type of outlets. With the covid-19, we have fewer information that usual. Therefore, we were no more representative at that level, but we think that we were still representative for the whole products sold in minimarket or for a whole consumption segment (without taking into account the type of outlet). Moreover, our estimation for each product in super and hypermarket was very precise as we usually used for these segments' extensive scanner data.
 - Consequently, we performed a double imputation for minimarket. First, we calculated the index for minimarket. Then we imputed the index for a consumption c in minimarket by the index for this consumption segment in super and hypermarket. Finally, we adjusted all these indices (consumption segment cross minimarket) by homothetically such that their aggregation was consistent with the minimarket index 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 do the usual quality adjustment as bridge overlap method, meaning 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 should continue for more than one month.
- (h) How can NSOs make decisions for a crisis like COVID-19?
- COVID-19 is an unexpected crisis. NSOs have to take quick decision and shall regularly perform risk analysis and prepare action plans in case we fail to get certain data. COVID-19 also brought up importance of new products like masks and gloves, which were not in our basket as their weight were below 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: UK Contingency Plan for Producing Consumer Price Statistics during the COVID19 Pandemic

Mr. Michael Hardie
Deputy Director, Prices Division
Office for National Statistics, United Kingdom (ONS UK)

Summary of Presentation

The presentation covered the topics on UK consumer price statistics, the contingency plan overview, response rate and April's consumer price statistics.

The main indices compiled in UK are indicated to be Consumer Prices Index including owner occupiers' housing costs (CPIH) - currently ONS's main measure of inflation; Consumer Prices Index (CPI) - ONS's Harmonized Index of Consumer Prices which is produced for Eurostat; and Retail Prices Index (RPI) – historic measure which ONS is required to produce by law. ONS's contingency plan covers CPIH, CPI, and RPI. 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 that large parts of the economy have been shut down, resulting in certain items becoming unavailable for consumers. Furthermore, price collectors were unable to undertake the usual physical price collection in stores due to the implementation of social distancing policies and movement restrictions brought into effect on 23 March 2020 in the UK as a result of the coronavirus (COVID-19) pandemic, this mode of collection accounts for approximately 80% of the price quotes used in the UK's consumer price statistics.

The collection of prices was therefore restricted to available items, ONS extended the collection period from 'index day' to index week, and price collectors priced as much as possible through websites and via telephone. Furthermore, the imputation methods used, response rate for CPIH and the results of April's Consumer Price Statistics were presented.

Questions and Answers

- (a) How can we make use of the “decision tree”?
- The “decision tree” is a useful system that NSOs can take reference, although 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 decided to 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 like COVID-19, how can we address 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 chainlink.
- (c) What factor is considered when rescaling the basket?

- In terms of the basket in the UK, it is set at the start of the year and ONS uses a range of expenditure data primarily from National Accounts that sets the basket for the year. ONS did not plan to update the basket until 2021 because timely, robust and detailed expenditure data is not available.
- (d) Why do we make use of Scanner Data?
- The data are collected by retailers when consumers purchase any goods and services from them. The benefit of scanner data is not only collecting extensive price information, but also the quantity of the product purchased. This allows us to accurately measure changing consumer spending patterns.
- (e) What measures can be taken to avoid artificial price variation when the collection period is extended?
- Amid COVID-19, the collection period was extended from 'index day' to index week. However, ONS 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 CPI Compilation in Times of COVID-19 Lockdown

Mr. Patrick Kelly
Chief Director of Price Statistics
Statistics South Africa

Summary of Presentation

The presentation covered the restrictions to 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 while some easing of the restrictions has been applied since May. The lockdown resulted in the unavailability of certain products or services for sale at all, a limited number of outlets to be open and travel restrictions which curtail purchases, price collection and office activities. Moreover, the pandemic may impact the health and safety of data collectors and outlets. , 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.

The presentation also covered 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 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 cons on the other hand include a problem if too many months are imputed, breaks in seasonal patterns and possible contradiction of elementary index trends. The pros of use of monthly change from year ago include that it only uses data from that specific elementary index, maintains seasonal trends, and keeps the annual rate stable. A Dutch study found it to be best performing. Its cons are 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, is CPI 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. The presentation also covered the Weekly CPI Project (EP-CPI) which published Consumer Price Index of essential products. EP-CPI is welcomed by users especially in the food and agriculture sector, created good visibility for Statistics South Africa together with two other COVID-19 special surveys done. However, in undertaking EP-CPI, instructions to the team were not consistently applied, no quality assurance measures taken, all calculations in Excel may have potential for errors, and it is unsure of the economic meaning and comparison to CPI changes.

Questions and Answers

- (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 the time of uncertainty, sometimes even if it is not available for more than two months, imputation is recommended to continue as there is hardly alternative. There are two sub scenarios: 1) consumers are not spending at all; 2) consumers are spending but it is difficult to collect. For the case of 2) 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 where 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 assessment of the data to make a decision. In case of some extreme events, for example, drought, it's 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 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 discuss at some point if a change is necessary. Currently, Statistics South Africa's weights are fixed, and updated only 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 put Methodology notes public on the website, so people can understand properly.
- (e) What is the best way of compilation when some items were 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 on CPI from imputation?
- In Statistics South Africa's case, overall weight of the indices that is subject to imputations is around 25% (by which the presenter means "high"). But it doesn't bias the overall index because Stats SA 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: CPI Compilation in Tunisia During COVID-19 Lockdown

Mr. Nejib Haouech
Price Statistics Expert
Statistics Tunisia

Summary of Presentation

Since March 20, Tunisia has been under general population lockdown due to COVID-19, which resulted in suspending field work with the exception of reduced-price surveys for CPI compilation. Certain challenges had therefore been to be met, including the security of the 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 levels of health, administrative and technical. 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 way can be used: Internet, telephone, email, newspaper advertisements, friends and family

It emerges from this situation that the field survey covered only 52.8% of the prices collected, and only 55% of the outlets. The missing prices were imputed according to the guidelines from the Consumer Price Index manual. Moreover, the introduction of online sales has been introduced for certain products. The introduction of new varieties must be careful, we should separate the delivery services from prices of products

The imputations of the missing prices were made in accordance with international recommendations, namely by using the average of the price variations of the 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 immediately 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 usual level of detail
For metadata, all imputed indexes should be flagged for users. Indicate number of missing and imputed prices by group. Proportion of outlets available for collection

On the way forward, it is planned to improve the collection of data by integrating the use of scanned data, and Internet data (web scraping) and to improve the IT infrastructure as well as the legal framework.

(Compiled by Emmanuel Ngok)

Section 2.5: Situation of CPI / CPI Compilation in AFRISTAT's Member States During The COVID-19 Period: Current Situation, Challenges and Main Recommendations

Mr. Yankhoba Jacques BADJI
Price Statistics Expert
AFRISTAT

Summary of Presentation

The presentation covered the history AFRISTAT, the state of production of CPI, price collection, management of missing prices, products without supply and demand, new products, publication and dissemination of results, challenges faced and main recommendations.

It was indicated that AFRISTAT has been mainly involved in the development of methodological documents as well as in the provision of training and technical assistance. The state of production of CPI at 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 also continue 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 Exogenous Method are used. It was pointed out that with regards to iterative estimation, the methods described in the Methodological Guide to the harmonized consumer price index in the WAEMU zone (or CEMAC) is preferred.

The iterative estimation method for heterogeneous products is either 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 either same day, same week, or same month while the exogenous method is the evolution of the higher aggregate, etc. In the case of the products which are without supply and demand, i.e. the services at outlets such as hotels, air travel agencies, and bus stations, concert halls, schools, bars, restaurants etc., the prices are imputed in accordance with the methodology applicable to heterogeneous varieties. It was further indicated that there was no introduction of new products in the baskets of member countries during COVID-19.

The presentation pointed out that the 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 dissemination to users. However, some delays were observed related 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 interests of transparency and maintaining trust was also pointed out.

The main challenges faced during 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 the 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 Covid19 period? Are the methodological changes and the high imputation rate induced during this Covid19 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 the IPC? What would be the impact 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% depending on the country.
 - There is no change in methodology during the 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 on the forecast of the index cannot be known a priori. But, if such a possibility arises, the data for the Covid19 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), since 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 of the same aggregate must be used for the imputation of a seasonal products without quotation in the month instead of arithmetic mean.
- (c) With the Phoenix tool, the application plans which iterative methods 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 favor the change of the same month of the last year. If this 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 manager who decides the best period taking into account his knowledge of the environment.
- (d) When you are talking about the news products, what prices will be considering 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 Methodological Guide to the harmonized consumer price index takes account of the imputation methods recommended in the CPI manual. In addition, the associated Phoenix application integrates the imputation methods of the Methodological Guide.

(Compiled by Emmanuel Ngok)

Chapter 3: Data Harvesting from Internet and Website

Section 3.1: Price Populi – Assisting Countries with CPI Data Collection

Mr. Andrew Baer
Senior Economist, Real Sector Division of Statistics Department
International Monetary fund (IMF)

Summary of Presentation

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

The presentation indicated 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 where anecdotal information indicates response rates to be down.

The presenter indicated that to address the 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 food and essential items from their current CPI basket. The pilot countries include Bahrain, 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 on 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’s the recommendations for particular products that are sold in variable volume, for example, rice? Also, 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-automatic options?
 - Semi-automation options can be a way for the improvement of the Tool, for example, 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 flexibility countries needed 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 the 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 where countries can download the 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 communicated directly with IMF through email.
 - The plan is to provide countries 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 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.
 - It 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: An Introduction to Web Scraping Using R

Ms. Randi Johannessen
Head of Price Statistics
Statistics Norway

Summary of Presentation

The presentation covered what web scraping is, how a webpage 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 consumer price index. 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, 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.

Furthermore, the presentation covered what has to be done before scraping a website, how a webpage created, about R programming language which is an open source, and a demonstration of the programme. It was also pointed out that the existing R Vest package is a built tool for web scraping and can be easily adapted to one’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 to be corresponded 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 the data from web-scraping to other statistical analysis software like 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 where local prices are not posted to the website?
 - There is nothing to scrape if there is no price on website available. In this case, we can ask them to send the data rather than going there physically. Also, 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?
 - So far, there hasn't been so much error noticed for scraping, but during the pandemic the availability is a problem, e.g. suspension of flights so prices may not be available at websites. And additional note on flights: International flights are from the consumers perspective booked normally several months in advance, so prices for a specific flight are scraped for many times: 5 months, 4 months, 3 months, 2 months, 1 months, 2 weeks, 1 week prior to departure date. 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

Ms. Valentina Stoevska
Senior Statistician, Department of Statistics
International Labour Organization (ILO)

Summary of Presentation

The presentation started by explaining two main price collection methods: (a) local price collection where prices are obtained from outlets around the country by personal visits and by telephone; and (b) central price collection where prices are collected at the central office with little or no field work involved. The restriction on movement due to 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, the availability of items, the technological and logistical capacity on NSO. On a short-term basis, the solution may be a collection of prices from a sub-sample of outlets chosen to be representative of the full sample, and/or a collection of prices from a sub-sample 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/email correspondence, from retailers' websites and scanner data, and newspaper advertisements, catalogues and list prices. The presentation covered the advantages and limitations of all the alternative methods indicated, as well as practical issues that need to be addressed for collecting prices without visiting the outlets

It was recommended that in order to maintain adequate coverage and ensure the quality of the price data collected, as well as 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. It was pointed out that; initially, in many countries, telephone enquiries were the only way to collect data previously collected by personal visits. In case of prolonged restrictions on movements, there is a need to take advantage of the new technology to collect prices; but also to train staff on using these technologies.

Questions and Answers

- (a) How can we deal with the issue that internet and mobile phone are rarely available in open market, where large part of our CPI basket is from?

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 to 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 of consumption pattern during COVID-19?

Because of the various restrictions linked to COVID-19 pandemic, the consumption pattern of consumers is changing in almost all countries. The Inter-secretariat Working Group on Price Statistics in its Business Continuity Guidance, does not recommend changing the weight for CPI for this year 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, and it is not consistent with the fixed basket approach used as the basis for compiling CPI. Also, 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 the usual normal pattern. 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 the data from web-scraping to COICOP structure?

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

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

Many NSOs have developed guidance notes regarding the issues relating to the impact of COVID-19 on CPI compilation. ILO is compiling this information from statistical offices. Many international and regional organizations as well as the Inter-secretariat Working Group on Price Statistics have prepared guidance notes on CPI. For interested group, ILO can provide the link to these notes.

(Compiled by Negussie Gorfe and Elias Fisseha)

Section 4.2: Challenges of Moving a Face-To-Face Survey to the Telephone

Ms. Jo Bulman
Living Costs and Food Survey Manager
Office for National Statistics, United Kingdom (ONS UK)

Summary of Presentation

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

The LCFS is designed to collect household expenditure data through a random sample of households in the UK and there are two stages to the data collection. The two stages are face-to-face interview involving all household members with 80-minutes duration on average, and each member of the household completes a two weeks' expenditure diary where the interviewers may visit respondents' homes multiple times. The uses of LCFS are to compile Retail Prices Index & Consumer Prices Index, as well as to have information on the spending patterns of the population, household expenditure for GDP, effect of taxes and benefits, and on food consumption and nutrition.

In response to COVID-19 the face-to-face interviewing was paused on March 17, and the shifting of the data collection to telephone interview was made. As a result of the 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/diaries posted back to the Head Office. The shift of the whole survey process has impacts 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 to telephone survey and the actions taken to improve telephone survey. It was further indicated that the output helped to understand the impacts of estimation and bias; to analyze the characteristics of the responding sample; in understanding mode effects; 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.

Finally, the presentation covered on how to explore online solutions during telephone interview by focus on income data collection first, then expenditure, challenges to reduce the length of questionnaire further, and development of systems to support data collection; and on diary by exploring use of Smartphones App, and Browser based version.

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 UK, it is a gift as opposed to payment. It is given in the form of voucher the respondents can used 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. We can also collect only more important data to 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 interview. It is possible that the smartphone application will replace the diary data collection but not the telephone survey.
- (d) What is the percentage of respondent's face-to-face vis a vis telephone?
- When we moved from face-to-face survey to telephone survey, the response rate dropped. In the case of UK, in April 2019 the response rate was 42% and for April 2020 when ONS moved to telephone survey it was 18% 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 for UK). And it is good to have age limits/ranges for different questionnaires.
- (f) How can we collect information on food items through telephone?
- For food data collection, it happens using the diary. Before COVID-19, respondents will attach the receipt they purchased to the diary and ONS get information from the receipt. Then it is manually typed into computer programme.
 - If the respondent purchases the good online, we will ask the receipt and we pick up the date from the receipt.
- (g) How can the collection process deal with weight/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 currently able to do this which may reduce the quality of this detailed information.

(Compiled by Isidore Kahoui and Elias Fisseha)

Session 4.3: Selected Survey Design Topics on Transition from Face-To-Face to Phone

Dr. Zeina Mneimneh
Director, Survey Research Center
International Unit, Institute for Social Research
University of Michigan

Summary of Presentation

The presentation introduced the organizational structure of the Institution for Social Research – Survey Research Center (SRC). She pointed out that the SRC was established in 1946 by the University of Michigan Board of Regents; SRC is a multi- and inter-disciplinary research organization devoted to the discovery of and insight into major issues within the social and behavioral sciences; SRC is an international leader in research involving the collection and analysis of sample surveys, administrative and other non-survey data; and its faculty specialize in cutting-edge theory and research on key questions facing society.

The presenter informed the participants the topics that have been covered by SRC during the Regional Seminar on CPI Compilation, which included frame coverage considerations between face-to-face and phone surveys & sampling for telephone surveys; nonresponse and weighting for telephone surveys; survey of consumer attitudes case study: moving from landline frame, to landline & cell phone, to cell phone only; measurement considerations in transitioning to telephone surveys; and staffing and infrastructure for telephone surveys. The presentation also 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 Dialing (RDD). What is the advantage and disadvantage of this tool as compared to telephone directories from telephone companies or previously collected telephone numbers from face -to-face survey?
- As long as 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, hence 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 nonresponse in terms of the telephone number information, nonresponse 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 number of them will be non-working or inactive. Telephone directories and telephones 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 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: 1) 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, 2) 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, where 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 (whoever attends the phone, for instance), 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 unemployment rate this way.
- (d) In a household where landline survey is conducted, how do we prevent the same number of the household is being interviewed again by cell-phone?
- Assuming that you are using a dual frame design (i.e. 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 that a household with either. Thus, when conducting a survey on 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

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

Summary of Presentation

The presentation covered sampling frames. It was indicated that 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.

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

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

Questions and Answers

- (a) In a country where survey used to face-to-face are pending because of COVID-19 pandemic. Is it possible to conduct household survey method like 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 to 30 minutes, but household surveys usually take a much longer period, like 45 minutes and hour. In case NSOs want to conduct household surveys though telephone, we might need to shorten the questionnaire.
- (b) Concerning Nationwide survey using telephone survey method, since 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 statewide 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: Nonresponse and Weighting for Telephone Surveys

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

Summary of Presentation

The presentation 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 breakoff, non-contact, other, unknown if household is occupied, 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 ([https://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx)), were also presented.

The presentation also covered, 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; nonresponse weighting; calibration, including post-stratification and raking; and trimming. Moreover, the presenter pointed out that multiplicity happens when a household has multiple telephones and therefore a higher probability of selection as it could have been selected through different sample elements. Thus, the household weights should be adjusted to address the increased probability of selection. In order to make the necessary adjustments on multiplicity, the information about the multiplicity should be collected. It was further indicated that the multiplicity problems could be addressed 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 correlations with these 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 like gender, age, education, income level etc. in HH survey tends to be more reliable. Not missing at random nonresponse cannot be solved by using weights, but by statistical modeling (imputation might be useful).
- (b) What's 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. In general, in such type of survey there are more auxiliary variables available for nonresponse 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 over “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

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

Summary of Presentation:

The presentation covered Dual-Frame Pilot Study, Dual-Frame Field Study and the transition from Dual-Frame to Cellphone Only.

The presentation indicated that the basic Survey of Consumer Attitudes (SCA) are monthly Random Digital Dialing (RDD) surveys which are used to measure consumer expectations and are undertaken in a rotating panel design. The transition from landline (RDD) sampling to Dual (RDD) Sampling has started in July 2012 and to Cell (RDD) sampling frame in January 2015. The addition of cell-phone sampling frame was considered due to the growing percent of cell phone only users. The characteristics of cell phone only households in the general population include adults renting their homes, adults aged 25 to 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 are on how to weight the SCA pilot data, and to also determine the effect of conducting more cell sample cases (El Kasabi et al., 2011). The project started with 100 sample lines from the cell phone frame. The conclusions of 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 loss in efficiency, the respondent demographic distribution were closer to the population demographic distributions reported by Current Population Survey (CPS) March supplement. Three commonly used designs were simulated to further understand the implications of using a dual frame. Clearly 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 prescreening and investigated the impact of using single cell-phone frame (Jiang et al., 2014). The interviews from the cell phone frame provides a better sample balance on demographic distribution over the course of the production period and especially it allows to interview a larger proportion of younger age group in addition to the increase in the nationwide coverage in the US. Switching from a dual-frame estimation to single frame estimation removes the complexity in the weighting method. The steps in weights for a dual-frame method is presented in the first part of the presentation, by switching to 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 single cell phone frame, reasons are simple, well-motivated and 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 sample design (from within household selection to any adult selection, household head versus any adult (18+), target population versus survey

population, portability); methods to improve efficiency require ongoing monitoring of questions for operations design (call-back rules, how to determine call back, call outcome).

References:

Mahmoud Elkasabi, Zeynep Tuba Suzer-Gurtekin, James M. Lepkowski, Richard Curtin, Rebecca McBee, Boyang Chai. (2011). The Impact on Accuracy of Estimates of Increasing Cell Telephone Sample to Correct Coverage Error in the Survey of Consumer Attitudes. AAPOR Annual Conference Phoenix, Arizona, May 12-15, 2011.

Charley Jiang, James M. Lepkowski, Richard Curtin, Dan Zahs. Comparisons Between Landline and Cell Phone Samples in The Survey of Consumer Attitudes. AAPOR Annual Conference Anaheim, California, May 15-18, 2014.

Charley Jiang, James M. Lepkowski, Tuba Suzer-Gurtekin, Michael Sadowsky, Richard Curtin, Rebecca McBee, Dan Zahs. AAPOR Annual Conference Hollywood, Florida, May 14-17, 2015.

(Compiled by Isidore Kahoui and Tesfaye Belay)

Session 4.7: Measurement Considerations in Transition from Face-To-Face to Phone

Dr. Zeina Mneimneh
Director, Survey Research Center
International Unit, Institute for Social Research
University of Michigan

Summary of Presentation

The presentation covered Total Survey Error (TSE) Framework, which is one of the contributions of the University of Michigan to the Regional Seminar. TSE helps to understand all types of errors that occur at various levels of survey undertakings and is likely to affect the final estimations of survey variables. Understanding TSE framework is essential to design and implement a survey in such a way to minimize the error for a given survey. 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 which are measurement bias and measurement variance. It was further indicated that response errors may be due to events not recorded in respondent's memory, respondents may misunderstand the question, respondents may forget relevant events, respondents may take shortcuts, or respondents may intentionally misreport.

The presentation also covered topics on general questionnaire design considerations when designing a telephone survey that include the need to reduce the number of survey items to essential ones, show cards and scales, labeled scales, sensitive questions, and interviewer effects. Moreover, it was indicated that survey could be impacted by telephone interviewers and the interaction between interviewers and respondents. In telephone surveys, interviewers need to rely only on aural cues to identify whether the respondent has any issues with comprehension, retrieval of information, etc.; interviewer ethnicity, language, cultural characteristics may impact 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 CATI interviewers).

Questions and Answers

- (a) What is SCA?

SCA stands for Survey of Consumer Attitudes. It is a survey that is being undertaken for many years now. It is a way of gathering the consumer sentiment as related to the economy. It is a major indicator for US economy and has been replicated in many other countries as well. 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 land line but now the survey has transition to cell phone.

- (b) How can we apply telephone survey in CPI data collection since it involves some of the items and how do you deal with response bias?

With regard to dealing with response bias, it depends with the type of response bias. For example, one type of bias is called recency effect where 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 reduce the number of response options is to randomize the options. The situation where there is order in the response like strongly agree, agree, etc. doesn't work in this case. If you are dealing with the social desirability bias, it might help to avoid the involvement of an interviewer by pushing the question to a web or an app where the respondent administers it to herself/himself.

- (c) How can we avoid the bias in the household answers from telephone?

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

- (d) In a telephone survey, can't the respondent's answers be biased by the fact that she/he wants to quickly end the survey? How can we solve this problem?

In telephone surveys, there could be bias by the fact that the respondent wants to quickly end the survey. It is an issue. However, there hasn't been a lot of empirical evidence that cell phone responders concentrate less affecting the measurement quality of the data more than land line. What can be done is that 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. The interviewer should be trained to insist on finishing the question before the respondent provides his 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 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 telephone survey, matching of interviewer and respondent can be done but it depends on the facility & infrastructure that is made available. Increased supervision also helps with reducing 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, definitely we should go for it. In fact, gender matching might be easier to do on 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 face-to-face survey proper identification of household will greatly help in reducing geographic coverage error. With Telephone survey, it depends how you select your telephone samples. The problem is that there might be a geographic coverage problem with telephone if telephone access to rural areas is limited. We may need to supplement with another frame if possible. If what can be done is only

telephone survey, as in the current situation, we might need to do some weight adjustment to mitigate some of the problems. But there is no 100% solution for this.

(Compiled by Isidore Kahoui and Tesfaye Belay)

Session 4.8: Moving from In-Person to Telephone Data Collection: Staffing and Infrastructure Considerations

Mr. 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 that 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 day and days of the week until contact is made. Moreover, when doing without non-verbal cues it is important to be prepared to make impression during the first few seconds of an introduction; to listen carefully; to reading verbatim; to have voice clarity, pace, and cadence; to have neutrality and to use probing. It was pointed out that the impact of the length of the questionnaire increases on average by 3.6% for telephone interviews due to the lack of non-verbal cues.

It was further indicated that social connectedness covers coping with isolation, to check in with the individual at least once weekly, to lead small group discussions every 1-2 weeks, and to have front line resources in place for any questions such as a Team Leader. Some key metrics may include attrition (quitting the job) and attendance (showing up late, leaving early, or not showing up at all), hours per interview, honored appointment rate (returning a call to a household at the appointed hour), dials per hour, and using contact windows to distribute calls. Equipment that are necessary for conducting decentralized telephone interviews are headsets, dedicated phones, and dedicated office space. The presentation further covered quality assurance which dealt 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, as well as 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 below 35 minutes, we can generally keep the dropout rate to below 5%. If the interview extends for an hour, we definitely get a much higher dropout rate. It also depends on the type of questions being asked. In a CPI type of questions usually about 5% dropout within 35-minute interview duration.
- (b) Does the gender (gender of interviewer, gender of respondents, between interviewer and respondent) have impact on 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 significant gender effect. When sending a male interviewer, we have a higher rate of people not answering the door in the first place in the U.S. Telephone surveys for general survey questions, if the interviewer speaks in a non-monotone way, the gender of the interviewer doesn't have any impact on the response. However, depending on the sensitivity of the questions asked, gender of interviewer may have an impact.
- (c) How can we manage funding for telephone surveys as they require more funding?
- For funding management, cell phones are charged a couple of cents per interviewer hour in order to accumulate enough funds to purchase cell phones. These phones are used across 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 to face-to-face interview.
- (d) How can we deal with scamming issue?
- 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 "white lists" numbers with all the major carriers. The idea of "white listing" 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 white listing has no impact on contact rate at all for cell phone numbers (that is, the respondents had cell phones). However, during COVID-19 our contact rate has doubled, and response rate have also gone up. This might be because people are home at all times and are willing to talk more now.
- (e) In a situation where the unemployment rate is high especially the household survey where we have to call or send an email to respondent its always difficult to get the concentration, and they tend to think the survey is a waste of time. How can we handle such situation?
- When we approach a low-income group during a survey, we shall explain to them that 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 Covid19 period increase stress and non-response rate compared to normal period?
- During COVID-19, stress level among the general public has 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?
- Although we would like to give a token of appreciation to the respondent for 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 AAPOR Task Force on Telephone Interviewing (<https://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 % 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 respondent to refer to a respondent booklet or other piece of paper, 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. **IWGPS-CPI business continuity guidance** drafted by the IMF and developed in conjunction with Inter Secretariat Working Group on Price Statistics (Eurostat, ILO, IMF, OECD, UNECE, and the WB) provides guidance to national statistical organizations to ensure the continued publication of their consumer price indexes in the face of challenges of 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 as well as 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. The 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 improvements made to the concepts and methods used to compile CPIs, reflect evolving data user needs, ensure broader consistency with SNA 2008, eliminate repetition and ensure consistency between chapters, incorporate the Practical Guide to Compiling Consumer Price Indices into the Manual, each chapter now concludes with a summary of the key points and identifies key recommendations, standardizing terminology to create more uniformity and consistency across countries, and elementary aggregate has been defined more clearly.
3. The 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 that is being testing in a handful of countries is meant to supplement current CPI collection and is also 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 aims to motivate 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 statistics offices are resorting more and more 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, South Africa, and features among ILO's 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. The IWGPS-CPI has elaborated guidelines to 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 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 indices have been envisaged; furthermore, special group of product indices have also been introduced, so as to maintain the relevance of the statistical office and retain the trust of the public in statistics. Information about

imputations should be provided with the release of the CPI; indices with full or significant imputations should be flagged; and if possible, to address the impact on the overall quality of the CPI.

6. **Telephone surveys** constitute another possible way to adapt data collection to evolving environments. Telephone surveys for price data collection have been advised by the ILO and used in France, South Africa, United Kingdom, and other countries to supplement price data. Among its advantages, telephone survey reduces costs in comparison to 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 are relatively quick and inexpensive. Furthermore, it can help reach remote places that are difficult to reach by in-person mode. However, due to the burden that it represents on respondents, only limited number of products can be priced, and only limited number of product characteristics can be collected, with limited scope for verifying accuracy of data. Telephone surveys have other selections and errors generations factors that need to be closely analyzed and accounted for.
7. **Data collection by email, from retail webpages (web scraping), and from magazines** have been implemented elsewhere but are heavily dependent upon 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. Survey respondent's attitude during the **interview** is an important quality element to be taken in consideration. Accounting for this needs to start at the earliest design stage of the questionnaires and be pursued during the interview phase, by considering, for instance respondent's burden or capacity to recall. The cues as to the respondents' attitude or need for clarification depends on the questionnaire administration mode used, whether through in face-to-face or through telephone. Many other factors that can influence the type of answers chosen by respondents need to be carefully analysed and taken in account by interviewers. The Survey research center of the University of Michigan has valuable resources in the 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 the 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 the data collections, processing and computation and analysis processes, which are more resilient and anchored on new technologies.

Appendix 1: Countries and Agencies Registered to the Seminar with Number of Participants

Sq.	Member State	Number
	Sum	369
1	Angola	10
2	Benin	6
3	Botswana	13
4	Burkina Faso	7
5	Burundi	34
6	Cabo Verde	6
7	Cameroon	13
8	Central African Republic	3
9	Chad	3
10	Comoros	12
11	Congo	2
12	Côte d'Ivoire	5
13	Democratic Republic of the Congo	4
14	Djibouti	4
15	Egypt	1
16	Equatorial Guinea	15
17	Eswatini	4
18	Ethiopia	5
19	Gabon	9
20	Gambia	2
21	Ghana	9
22	Guinea	3
23	Kenya	12
24	Lesotho	6
25	Liberia	2
26	Libya	4
27	Madagascar	4
28	Malawi	2
29	Mali	4
30	Mauritania	4
31	Mauritius	5
32	Morocco	2
33	Mozambique	6
34	Namibia	5
35	Niger	2
36	Nigeria	24
37	Rwanda	5
38	Sao Tome and Principe	9
39	Senegal	4
40	Seychelles	4
41	Sierra Leone	6
42	Somalia	27
43	South Africa	26
44	South Sudan	2
45	Sudan	2
46	Togo	9
47	Tunisia	10
48	Uganda	3
49	United Republic of Tanzania	5
50	Zambia	1
51	Zimbabwe	4

Sq.	Agency	Number
	Sum	63
1	AfDB	2
2	AFIRSTAT	2
3	AUC	2
4	CEEAC	2
5	CEMAC	1
6	COMESA	1
7	DHS Program	1
8	EAC	1
9	EASTC	4
10	ECA	5
11	ECOWAS	1
12	ESCWA	1
13	Eurostat	3
14	FAO	7
15	ICF International	1
16	IGAD	1
17	IMF	4
18	INSEE France	4
19	Makerere University	3
20	Moldova NSO	2
21	NEPAD	1
22	ODI	2
23	ONS	5
24	SADC	1
25	UEMOA	1
26	UMA	2
27	UNEP	1
28	World Bank	1
29	York University	1

Appendix 2: Results of Post-Seminar Survey on Areas Needed for Further Capacity Building and Technical Assistance

Country	Phone/Mail	Web-scraping	Imputation	Others
Count	28	28	30	8
Benin	Y	Y	Y	
Botswana	Y	Y	Y	Updating of weights, rental surveys and Owner-Occupied Housing
Burkina Faso		Y	Y	
Burundi	Y		Y	
Cabo Verde	Y	Y	Y	
Cameroon	Y	Y	Y	
Central African Republic	Y	Y	Y	
Chad	Y	Y	Y	
Comoros			Y	
Cote d'Ivoire	Y	Y	Y	Technique prevision
Djibouti	Y	Y	Y	Sur la rénovation de l'IPC
Equatorial Guinea	Y		Y	
Eswatini	Y	Y	Y	
Ethiopia	Y	Y	Y	Household consumption survey methodology, data collection, processing and analysis
Gabon	Y	Y	Y	
Ghana	Y	Y	Y	
Kenya	Y	Y	Y	
Lesotho	Y	Y	Y	
Madagascar	Y	Y	Y	Résolution de pondération et modèle de prévision IPC
Malawi	Y	Y	Y	
Mauritania	Y	Y		
Morocco		Y		
Mozambique	Y	Y	Y	
Namibia	Y	Y	Y	CPI data analysis, Deriving of CPI weights, Report writing, Graphic design, Regional/Provincial CPIs compilation
Sao Tome and Principe			Y	
Senegal	Y	Y		
Seychelles		Y		
Sierra Leone	Y		Y	
Somalia	Y	Y	Y	
Somalia - Puntland State	Y		Y	
Somalia - Somaliland			Y	
South Sudan	Y	Y	Y	designing the Country production system and CPI catalog.
The Gambia	Y			
Togo	Y	Y	Y	
Tunisia		Y		Collecte de données à partir des données scanner (Scanner Data)
Uganda		Y	Y	
