Statistical Commission
Forty-ninth session
6 – 9 March 2018
Item 3(j) of the provisional agenda
Agricultural and rural statistics

Background document
Available in English only

Food data collection in household consumption and expenditure surveys. Guidelines for low- and middle-income countries.

Prepared by the
Inter-Agency and Expert Group on Food Security, Agricultural and Rural Statistics

Inter-Agency and Expert Group on Food Security, Agricultural and Rural Statistics

Food Data Collection in Household Consumption and Expenditure Surveys
Guidelines for Low- and Middle-income Countries

Final version -February 22, 2018

The document is the result of contributions from a large team of experts from academia, international organizations and national statistical offices. The group was coordinated by the World Bank and FAO, under the aegis of the Inter-Agency and Expert Group on Food Security, Agricultural and Rural Statistics (IAEG-AG). The report draws on new material as well as material that members of the group have authored and published elsewhere. Full acknowledgment of the contributors and the material cited or re-used will be included when the document is finalized. The work of the group was supported in part by the Global Strategy for Agricultural and Rural Statistics. The document will be professionally edited for language and formatted after the discussion at the 49th Session of the UN Statistical Commission.

Preface

The measurement of food consumption and expenditures is a fundamental component of any analysis of poverty and food security, and hence the importance and timeliness of devoting attention to the topic cannot be over-emphasized as the international development community prepares to meet the challenges of monitoring progress towards Agenda 2030.

In 2014 the International Household Survey Network (IHSN) published a desk review of the reliability and relevance of survey questions as included in 100 household surveys from low- and middle-income countries. The report was presented in March 2014 at the 45th session of the UN Statistical Commission (UNSC), in a seminar organized by the Inter-Agency and Expert Group on Food Security, Agricultural and Rural Statistics (IAEG-AG).

The assessment painted a bleak picture in terms of heterogeneity in survey design and overall relevance and reliability of the data being collected. On the positive side, it pointed to many areas where even marginal changes to survey and questionnaire design could lead to a significant increase in reliability and thus great improvements in measurement accuracy. The report sparked a lot of interest from development partners as well as UNSC member countries, and the IAEG-AG was encouraged to pursue this area of work with the ultimate objective of developing, validating and promoting scalable standards for the measurement of food consumption in household surveys.

The work started with an expert workshop that took place in November 2014 in Rome. Successive versions of the Guidelines were drafted and discussed at various IAEG-AG meetings, and in another expert workshop organized in November 2016 in Rome. The Guidelines were put together by a joint FAO- World Bank team, with inputs and comments received from representative of national statistical offices, international organizations, survey practitioners, academics and experts in different disciplines (statistics, economics, nutrition, food security analysis). A full list of contributors will be included in the final version of the document. In December 2017 a draft of the Guidelines was circulated to 148 National Statistical Offices from low to high income countries for comments. The comments received during this global consultation were then incorporated in the document by the drafting team, and are reflected in the current version of the document. The process received support from the Global Strategy for Agricultural and Rural Statistics.

The document is now being submitted to the 49th Session of the UNSC for endorsement. The objective is for the Guidelines to become a useful reference document for practitioners, but also a tool for promoting the international harmonization of survey instruments and the comparability of the resulting data.

Executive Summary

Food constitutes a key component of a number of fundamental welfare dimensions, such as food security, nutrition, health, and poverty. It makes up the largest share of total household expenditure in low-income countries, accounting on average for about 50 percent of the household budget (USDA, 2011), and thus constituting a sizeable share of the economy. Proper measurement of food consumption is therefore central to the assessment and monitoring of various dimensions of well-being of any population, and hence of interest to multiple international, national, and local agencies, and to several development domains – social, economic, and human.

Food consumption data are needed to monitor global targets such as the newly adopted Sustainable Development Goals. Its measurement is crucial to assess and guide FAO's mandate to *eradicate hunger, food insecurity and malnutrition*, or the World Bank's twin goals of *eliminating extreme poverty and boosting shared prosperity*. It is also important for national accounts measures of the overall size of the economy, such as GDP. Finally, it is of interest to national and local governments, as well as non-governmental organizations, to guide local and regional analysis and policy, as the mismeasurement of food consumption can lead to misallocation of funds and compromise the design, monitoring, and evaluation of relevant policies and programs.

The main vehicle used to collect information on food consumption for these purposes are household consumption and expenditure surveys (HCES). However, current practices for collecting consumption data differ widely across types of surveys, between countries and over time, compromising both the quality and comparability of resulting data and measures. In the interest of improving food consumption measures, and to ensure that data collected respond to the needs of a wide range of users, several development partners have come together around a common agenda aimed at harmonizing practice and recommendations for design of food consumption modules in HCES. This report proposes a preliminary set of internationally agreed recommendations to adopt in future HCES in order to collect food data aiming at improving the measurement of food consumption.

Taking stock of current practices

The last two decades have witnessed unprecedented progress in the quantity of household consumption and expenditure data across the developing world. In 1990, the World Development Report published by the World Bank was based on data from only 22 countries, and no country had more than one survey (Jolliffe et al., 2014). Today, there are at least 137 countries with consumption or expenditure information, and many of them have multiple surveys, adding to a total of more than 845 surveys. The number of countries with no poverty data (which is primarily estimated from consumption expenditure data) over a 10-year period declined from 33 percent to 19 percent since the 1990s, whereas the share of countries with 3 or more data points over a 10-year period increased from 27 to 41 percent over the same period (Serajuddin et al., 2015).

Depending on their primary objective, the surveys that collect information on household consumption or expenditure can take different forms, including Household Budget Surveys (HBS), Income and Expenditure Surveys (IES), or 'multi-purpose' or 'integrated' household surveys, like the Living Standards Measurement Study (LSMS) surveys. We refer to this family of surveys, which are often nationally representative, as Household Consumption and Expenditure Surveys. While the variety of purposes naturally translates into different designs, the dramatic increase in the number of household surveys in developing countries has been associated with a proliferation of instruments and methods in the collection of food data that cannot

be explained only on the basis of their different purposes or of country-specificity. To a large extent, this is a manifestation of the lack of globally agreed standards to guide the collection of food consumption in household surveys, which is evident by the heterogeneity in methods we observe within the same type of survey both across countries as well as within countries across time.

A recent desk review, jointly led by the International Household Survey Network (IHSN), FAO and the World Bank (Smith et al., 2014), identified the multiple purposes these household surveys serve, proposed a method to assess the reliability and relevance of survey questions, and applied the method to 100 household surveys from low- and middle-income countries, a sample that resulted from selecting the most recent nationally representative household survey from each developing country, with the only condition for inclusion of having enough documentation. The assessment found large differences in data collection methods across surveys and paints a bleak picture with regards to both the comparability and quality of resulting measures. It points to many areas where survey design and questionnaires can be significantly improved. The multiplicity in methods, variation in quality, and diversity of uses, are increasing concern both in academic and policy circles, leading to calls for a more systematic approach to survey design and a better understanding of its consequences when it comes to measuring food consumption or expenditure.

Survey design does matter

The international statistical community is well aware that poverty and hunger estimates are inconsistent across countries and over time, and that the lack of harmonization of survey methods contributes to this situation. The notion that survey design matters is not new, as evident by the work by Mahalanobis and Sen (1954) or Neter and Waksberg (1964). While the issue has been largely neglected in the economics literature for a long time (Browning et al., 2014), in recent years there has been a surge of interest in the measurement of household expenditure also among economists, mostly spurred by two factors: First, an increasing body of evidence suggests that inferences from comparisons of survey estimates across space and time can be seriously compromised by variation in survey design and practice. Second, the persistent fall in quality that consumer expenditure surveys are experiencing across several developed countries, particularly associated with non-response and under-reporting.

With respect to evidence from developing countries, Deaton and Grosh (2000) provide a comprehensive review of the issues and data needs for the measurement of consumption in household surveys, drawing on the lessons derived from LSMS surveys. A recent special issue in the journal *Food Policy* includes case studies from a diverse set of developing and OECD countries analyzing, both theoretically and empirically, how different surveys design options affect the quality of the data being collected and, in turn, the implications for statistical inference and policy analysis (Zezza et al., 2017). Several other studies have analyzed the implications of particular aspects of survey design for total expenditures, poverty and inequality measures. For example, Joliffe (2001) and Pradhan (2001) evaluated the impact of varying the length of the consumption list in El Salvador and Indonesia respectively; Gibson et al (2003) looked at the effect of changing the length of the data collection period for the case of China; Beegle et al (2012) compared results from 8 questionnaire designs which include variations in methods of data capture, level of respondent, length of reference period, number of items in the recall list, and nature of the cognitive task required of the respondent; and Backiny-Yetna et al (2017) compared different data collection methods, which include 7-day recall period, a 7-day diary, and a 'usual month'. All of these studies find that the design and implementation of survey instruments for collecting food consumption impacts resulting measures considerably.

As economies develop, implementing quality surveys is bound to become more difficult and require changes to current approaches in many countries. First of all, respondents tend to become less compliant and harder to survey as their incomes increase leading to greater non-response, attrition and bias. Second, inequality rather than poverty *per se* becomes a more common objective for measurement efforts; even countries who have escaped mass poverty have policy concerns over inequality, and poverty itself becomes more sensitive to inequality the more the poverty rate falls. This change in focus places additional demands on surveys. Meanwhile, with lower levels of poverty, errors have a larger influence on the measurement of poverty and hunger, especially when attention turns to the contribution of transitory poverty and transitory hunger. Last, existing survey methods that may do well for staple ingredients do less well for income elastic food items and for changes in consumption patterns. For instance, as consumption of food away from home increases the person responsible for reporting the household's food consumption is less aware of the amount and/or type of food consumed by other household members.

On the road to better food data

There is arguably a need for more systematic research before general recommendations for the collection of food consumption or expenditure data in HCES can be provided. While the lack of broadly accepted practices affects all aspects of household consumption or expenditures analysis, it is significantly pronounced for the case of food consumption. This is not only the case because few existing papers focus explicitly on food (e.g. the work by Backiny-Yetna et al., 2017), but also because the research has almost exclusively looked at *mean* expenditures, and poverty or inequality measures. However, the impact of survey design on other measures is less-well understood, and there is an increasing interest in using the data to analyze other dimensions of well-being, such as food security, health, and nutrition.

The UN Statistical Commission (UNSC) endorsed, in 2012, the creation of the Inter-Agency and Expert Group on Agricultural and Rural Statistics (IAEG), as part of the governance Framework of the Global Strategy to Improve Agricultural and Rural Statistics (GS), with the mandate of "*guiding methodological developments in statistics for food security, sustainable agriculture, and rural development*" (UNSC, 2012: p.10). In the spirit of this mandate, a workshop was convened in Rome in November 2014 bringing together experts in the measurement of food consumption derived from HCES. The workshop presented the latest methodological developments, discussed the main limitations in current practices, and identified priority areas for future research. Following this workshop another expert consultation took place in November 2016 in Rome to discuss possible guidelines and form a first set of recommendations. A first draft of the present guidelines document was then discussed by the members of the IAEG-AG and presented in an Open Seminar during the 48th Session of the UN Statistical Commission in New York in March 2017. The results of those discussions are reflected in this document.

Conclusion and recommendations

The recommendations in these guidelines aim at assisting practitioners in improving survey design, while taking into account the cost of organizing a survey and the constraints statistical offices in low- and middle-income countries face. To stay relevant, these recommendations will have to evolve over time as additional research and new technologies become available. For instance, changes in consumer behavior and in how food systems change may lead to a rethinking of the underlying survey design principles, such as the priority placed on collecting data on ingredients and replying primarily on quantities reported. The main recommendations of the guidelines can be summarized as follows:

- ***Recall vs diary, and length of reference period.*** HCES should adopt a 7-day recall period for food consumption measurement. In low- and middle-income countries recall surveys are generally preferable to diary surveys which should only be deployed with careful and continuous supervision and should not exceed 14 days. While a *well-implemented* diary is generally held as the gold standard for food expenditure data collection, there is ample evidence that in low-income settings with a prevalence of illiterate respondents, diaries are often implemented as a series of short recall interviews, with issues of respondent and enumerator fatigue affecting data quality, and with unsustainable implementation costs. Any change in the recall period or method should be accompanied by an experimental component that assesses the impact of methodological changes on survey estimates and enables temporal reconciliation.

- ***Seasonality and number of visits.*** HCES surveys should spread data collection over time in order to capture seasonal variation in food consumption and expenditure patterns. Two main approaches can be adopted: one visit per household and split the sample over 12 months or conduct two visits per household over 12 months.

- ***Acquisition and consumption***. HCES should collect data on all the main modes of acquisition (food purchases, own production and food received in-kind) and make it clear to respondents, enumerators, and data users whether the survey is collecting data on food acquisition and/or food consumption (with or without waste). Survey design should avoid sources of incomplete or ambiguous enumeration (e.g., food consumed from own production vs food harvested; routine month vs. specific months; rule-out leading question).

- ***Meal participation***. Information should be collected on the number of meals and the number of individuals (household and non-household members) who participated in each meal. It is recommended that all HCES should consider adding an individual household member-based meal module.

- ***List of food items***. Data should be collected on all types of foods and beverages that make up a country's human diet. Lists should be kept up to date to take into account changes in dietary habits and should be drawn having in mind that products that account for minimal budget shares can have particular nutritional values. A list of general principles that can guide the design of a food list is included in the guidelines, and includes the following criteria:
    - Ensure that survey food lists are sufficiently detailed to accurately capture consumption of all major food groups making up the human diet. To facilitate data integration and analysis, the categories used in COICOP, FoodEx2 but also in food composition tables, should be considered;
    - Include exclusively food (no other commodities);
    - Processed foods (from moderately to highly processed) need to be adequately represented;
    - All food groups need to be represented and include a reasonable number of food products;
    - Broad categories (such as fish, meat, fruits, vegetables) should be avoided and for each basic food group list most common items and add "other" category as needed. Items from subsidized programs, food fortification programs, and micronutrient rich foods should be listed individually.

- ***Non-standard units of measurement.*** Surveys should allow respondents to report in both standard and non-standard units (NSU), according to what they are most familiar with for each item reported.

It is critical to establish (define or collect) conversion factors for all NSUs that will be used. Additional features to improve the accuracy of reported NSU quantities – such as market surveys to establish accurate NSUs and conversion factors, photo reference aides, and on-the-spot value verification using Computer Assisted Personal Interviewing (CAPI), may also be considered. NSOs and implementation partners should work together to establish NSU databases that can be used across surveys, effectively increasing the standardization of the units while also limiting the cost of their implementation. To this end, survey implementers should thoroughly document all NSU protocols and related conversion factors and make them publicly available.

- ***Food away from home***. The practice of collecting food away from home (FAFH) information with just one question should be discontinued. The importance of FAFH warrants the design of a separate module, based on a clear definition of FAFH. In particular, surveys should be clear in identifying how to collect information on potentially ambiguous categories of food: "food prepared at home and consumed outside" and "food prepared outside and consumed at home". The latter can be integrated in the food at home module (e.g. take-out food) provided there is clarity to enumerators, respondents and data users that that is the case.

  Data collection should be organized around meal events, including snacks and drinks. At a minimum, surveys should collect information on the monetary value of each meal consumed away from home (breakfast, lunch, dinner, solid snacks and drinks). The meal events list should be adapted to the local context. FAFH is best collected via individual level interviews of adults. A proxy respondent can be used to report on children's meals away from home and other adults. Surveys should use for FAFH the same reference period used in the food consumed at home module. The data to estimate FAFH related nutrient content, when feasible, will have to come from other data sources integrated to the HCES, such as a survey of food establishments or administrative data on the content of public meals (e.g. schools, social programs).

1. Introduction

1.1. Background and motivation

Food is an important component of many fundamental dimensions of welfare, such as food security, nutrition and health. It comprises the largest share of total household expenditure in low-income countries, accounting for about 50 percent of the average household budget (USDA, 2011) and it is thus key for consumption and poverty analysis. Low levels of food access play a role in explaining why around 815 million individuals were estimated to be chronically undernourished in 2016 (FAO, WFP, IFAD, UNICEF, WHO 2017). Data on food consumption and expenditure underpin the most widely used measures of poverty and of food security. The collection of high quality food consumption data is therefore central to the assessment and monitoring of the well-being of any human population, and is of interest to national governments, international agencies, and to anyone concerned with monitoring and understanding trends in social, economic, and human development.

Data on food consumption are needed, for example, to build the indicators and monitor some of the targets set for Sustainable Development Goals 1 and 2 (end poverty and hunger). Similarly, data on food consumption are needed to assess and guide FAO's mandate to help eradicate hunger, food insecurity, and malnutrition, as well as the World Bank's twin goals of eliminating extreme poverty and boosting shared prosperity[1]. Even more importantly, national and local governments as well as non-governmental organizations need high-quality food consumption data to guide local and regional analysis and policy, since the mismeasurement of food consumption can lead to misallocation of funds and may compromise the design, monitoring, and evaluation of relevant policies and programs.

Box 1. The concepts of food consumption

Food data collected in HCES can be diverse, and often refer to diverse concepts. Even the term "food consumption" lends itself to multiple meanings. When the focus of the analysis is expenditure, the term "consumption" can designate the purchase of foods, disregarding the end-use of what was purchased. At the opposite end, analyses and surveys that focus on nutrition use the term "food consumption" to designate the intake of a food, possibly net of unusable parts. Throughout this document the term "food consumption" is used in a general sense and encompasses concepts or data that include food consumption, acquisition, expenditure, and intake. Additional descriptives are specifically used in places where we address or contrast their specific meanings, or for details that relate to that precise terminology.

The last two decades have witnessed unprecedented progress in the production and dissemination of household consumption and expenditure data across the developing world. In 1990, the World Development Report published by the World Bank was based on data from only 22 countries, and

[1] For a list of indicators that can be derived from food data collected in HCES see Moltedo et al., 2014 and Foster et al. 2013.

no country had more than one survey (Jolliffe et al., 2014). Today, there are at least 137 countries with consumption or expenditure information, and many of them have multiple surveys, adding to a total of more than 845 surveys[2] (Ferreira et al., 2016). The number of countries with no poverty data (which is primarily estimated from food consumption data) over a 10-year period declined from 33 percent to 19 percent since the 1990s, whereas the share of countries with 3 or more data points over a 10-year period increased from 27 to 41 percent over the same period (Serajuddin et al., 2015).

When the FAO started publishing estimates of the prevalence of undernourishment to monitor the hunger targets of the World Food Summit in 1999, it used food consumption data collected in household surveys for only 15 countries, relying on indirect estimates of the distribution of food consumption for most countries. The latest assessment, conducted in 2017, used data from 57 surveys for 39 countries.

Despite this progress, there are still 29 countries without a survey between 2002 to 2011 that could be used to measure poverty, and another 28 which have had just one survey in that time frame that would enable estimating national poverty figures (Serajuddin et al., 2015). Without such data, it is impossible for these countries or for international development actors to analyze trends and progress (or lack thereof) in poverty eradication, something that has prompted the World Bank President to pledge to assist 78 of the poorest countries in the world to conduct at least one such survey every three years. At the same time, the UN Statistical Commission has established an Intersecretariat Working Group on Household Surveys at its forty-sixth session "to foster coordination and harmonization of household survey activities across agencies and member countries" (United Nations, 2014). These initiatives will result in a surge in the number of household surveys in developing countries in the coming years, underscoring the urgent need for more rigorous guidance on survey design.

A challenge facing low- and middle-income countries is that implementing quality surveys is bound to become more difficult in many respects, as their economies develop. One aspect NSOs need to prepare for is that people become less compliant and harder to survey as their incomes increase. Another is that, with per capita income growth, policy concerns relate more to issues of inequality rather than to poverty *per se*. Moreover, even when poverty remains a concern, measures of poverty become increasingly sensitive to measures of inequality as the poverty rate falls. Thus, from a measurement point of view this means that inequality and variance of welfare measures become more important. At the same time, with lower levels of poverty, errors have a larger influence on the measurement of poverty and hunger, especially as interest grows in the transient components of these welfare shortfalls. Furthermore, existing survey methods that may do well for staple ingredients and the traditional view of the household sharing food from a common pot, do less well for income elastic food items and for changes in consumption patterns. For instance, as an increasing share of food consumption happens away from home, as the result of actions and decisions of individual household members, it will become increasingly difficult for 'the main food preparer' to report on the content and value of those consumption episodes (meals, snacks).

[2] These figures are based on a count of consumption surveys available from PovcalNet, Eurostat and US Bureau of Labor Statistics as of September 2017.

Depending on their objectives, surveys collecting information on household consumption take different forms and labels such as Household Budget Surveys (HBS), Income and Expenditure Surveys (IES), or 'multi-purpose' or 'integrated' household surveys, like the Living Standards Measurement Study (LSMS) surveys. We collectively refer to this family of surveys, which are almost always nationally and sub-nationally representative, as Household Consumption and Expenditure Surveys (HCES).

The primary objectives of HCES include measuring poverty and consumption levels, deriving consumption patterns needed for the calculation of consumer price indices, and providing inputs to the compilation of national accounts. In many cases, they are the only available source of information to estimate the distribution of food consumption in the population. While the variety of HCES purposes naturally translates into different survey designs, the dramatic increase in the number of household surveys in developing countries is associated with a proliferation of approaches and methods of food data collection, which is not only due to different purposes or country-specific considerations. This is because international guidelines and recommendations for the design and implementation of each of the distinct types of HCES surveys, when they exist, are specific to each type of survey, are generally not prescriptive, often lack coherence, and usually leave much flexibility to national survey statisticians and to the consultants and donors who may support their survey efforts. Consequently, we observe heterogeneity in methods, both across countries as well as within countries over time.

HCES are also increasingly being used to address the food and nutrition information gap, despite not necessarily being designed for this purpose, because they contain a wealth of information about food acquisition and consumption; are conducted with increasing frequency in an increasing number of countries (Serajuddin et al, 2015); have large samples; and are often statistically representative at subnational levels. These multi-purpose surveys are also much less costly than other (stand-alone) dietary assessment data sources as they are already being conducted and paid for by government agencies (Fiedler, 2013). Increasingly, HCES data are repurposed and used to calculate food security indicators[3], compile food balance sheets, plan and monitor food-based nutrition interventions, serve the information needs of the private sector, and contribute to other research work. The degree to which a survey dataset is "fit for purpose" is specific to each one of these particular uses (Smith et al., 2014).

While there has been a surge of interest in HCES analyses of food security and nutrition issues, to date, food security analysts and nutritionist have been overwhelmingly passive user of already-collected HCES data. As a result, the full potential of these particular types of repurposing of HCES has yet to be realized. There is a lack of awareness among non-economists about what these data contain and there is a need for further research and action to improve the quality and utility of these

[3] Food consumption data are used in the calculation not only of FAO's prevalence of undernourishment indicator, but also of indicators such as the food budget share, and dietary diversity indicators such as simple food counts or counts of food groups. Other common food security indicators such as the Food Insecurity Experience Scale (FIES), the Household Food Insecurity Access Scale (HFIAS), the Household (and Individual) Dietary Diversity indicators (HDD and IDD), and the World Food Program's Food Consumption Score (FCS) are collected using purposely designed survey modules.

data and the methods that can be applied to analyze them from a food security and nutrition lens. If the food security and nutrition community—with its unique skills and experiences—were to become more proactively involved in the design, implementation, and analyses of HCES, these could be strengthened substantially as tools for evidence-based food and nutrition programming and policymaking.

There are a few recent antecedents to the present document in terms of attempts to create a common ground for international household survey data on poverty and consumption. One highly relevant document is the report of the Seventeenth International Conference of Labour Statisticians on Household Income and Expenditure Statistics (ILO, 2003). This report is very useful as a reference for internationally agreed-upon definitions and concepts, and does an excellent job of discussing many of the survey design issues the present document is concerned with. It stops short, however, of providing specific recommendations on survey design, which is the main objective of the present document.

Other documents exist that provide guidelines for the implementation of high quality consumption surveys in Europe, such as the guidelines provided by Eurostat for implementation of the European Union Statistics on Income and Living Conditions (EU-SILC)[4], and by the Household Finance and Consumption Network (HFCN)[5]. Both surveys have sections attempting to measure consumption, and both are supported by documents providing specific guidelines for survey design. While both surveys and the supporting documentation can provide ideas on how to design consumption surveys elsewhere, they are at best only partially relevant for the issues faced by statistical offices and food consumers in low- and middle-income countries[6].

The UN *Handbook on Poverty Statistics* (UNSD, 2005) aimed to assist practitioners by providing clarity around poverty measurement issues and the key decisions involved in choosing and applying methods for poverty measurement. Though the handbook explicitly discussed issues with consumption data (Gibson, 2005), it did not provide specific guidance for survey design. The LSMS 'Blue Book' (Grosh and Glewwe, 2000) on the other hand provided such guidance in the chapter on Consumption authored by Deaton and Grosh (2000). The current document builds on that chapter, revisiting the same themes but with new evidence accumulated over almost two decades of experience since that publication. The focus here is also somewhat different as these Guidelines also look to provide guidance that is relevant to the different users of HCES surveys (including food security and nutrition analysists, consumer price index (CPI) and national accounts compilers),

[4] Available online at http://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions

[5] Available online at https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html

[6] FAO guidelines are also available for nutrition assessments (e.g. EFTA, 2014; visit FAO website: http://www.fao.org/nutrition/assessment/en), but those are less relevant for designing HCES.

whereas Deaton and Grosh were mainly targeting an audience of poverty economists using LSMS data.

This document presents a set of recommendations for practitioners engaged in the collection of food consumption data in HCES. In developing these recommendations, it draws on the results of a desk assessment of questionnaires from 100 nationally-representative surveys conducted in low- and middle-income countries, commissioned by the International Household Survey Network (IHSN), the World Bank (WB) and the Food and Agriculture Organization of the United Nations (FAO) (Smith et al., 2014). That assessment revealed an unwarranted level of variation in survey practices around the world, identifying clearly suboptimal practices still being used in a number of countries. The assessment concluded that while many basic reliability criteria are met by the large majority of current HCES, only 13 percent of the questionnaires analyzed met all seven criteria used to classify a survey as reliable.

As a result, FAO and the World Bank initiated a collaborative effort to identify and disseminate best practices for food consumption measurement through HCES, which includes both methodological work and the publication of guidance material, of which this document is the first output. This work program was presented at the first meeting of the Inter-Agency and Expert Group on Food Security Agricultural and Rural Statistics (IAEG-AG) in February 2015, and endorsed by the IAEG-AG members. The guidelines were finalized through an extensive consultation process, drawing on inputs from experts on relevant disciplines and from representative national statistical authorities from all developing regions and presentations at the margin of the United Nations Statistical Commission (UNSC) Sessions in 2015 and 2017. The current version of these guidelines also draws on expert workshops organized by the Global Strategy for Agricultural and Rural Statistics in Rome in November 2014 and in November 2016, and by the World Bank in Washington DC in November 2015.

The guidelines stop short of articulating recommendations in a strongly prescriptive manner, particularly given that for several survey design decisions, there is no consensus among experts and/or not enough evidence from methodological research for there to be a clear best practice to be recommended for adoption[7]. While more methodological research is needed in most of the areas covered by these guidelines, survey practitioners and national statistical authorities need to implement surveys, and international organizations are requested to provide advice. These guidelines aim to offer users a set of recommendations around which a group of experts and national statistical offices agreed upon given the current status of knowledge. At the same time, the document explicitly recognizes that as more knowledge is generated through survey experience and methodological research, and with technological advances in survey data collection, the guidelines will need to be periodically updated to keep or increase their relevance.

Additionally, it should be noted that since food consumption data are used for different purposes, not all decisions matter equally for all users. In that sense Smith et al. (2014) distinguish issues around

[7] The guidelines are also limited to issues around questionnaire design and major decisions in field implementation (timing, number of visits). Other survey implementation features that are as important for data quality (training, data entry etc.) are not covered by the guidelines. Important emerging issues where consolidated international practices are still lacking, such as the measurement of food waste, are also not covered in the document.

the reliability of survey data (i.e. "whether the survey design and method complies with good practice") and their relevance (i.e. whether the data provide the information or indicators needed by different users). For instance, users interested in consumption data for the calculation of the consumer price index will be less interested in details of the quantities of specific food items consumed compared with users interested in analyzing nutrient consumption and deficiency. Following these guidelines in full would allow for HCES data to satisfy the needs of all the main uses (poverty, inform CPI, national accounts, food security and nutrition).

Given the costs involved in data collection, there are clear advantages to ensuring that food consumption data include information relevant to as many users as possible, yet without compromising quality. Survey design requires striking an appropriate balance between these competing demands. It is important that all core users of the data are consulted while developing the questionnaire, sampling plan and fieldwork. There should also be a realistic and pragmatic recognition, however, of the extent to which repurposing is feasible. A HCES will never be a substitute for an assessment of dietary intake, and analyses of food consumption based on HCES data will always have limitations compared to analyses based on purposely designed surveys.

1.2. Objectives and audience

These guidelines are put forward as a draft document for consultation among the international statistical community. As noted above, the objective is for the draft guidelines and the consultation process to lead to a finalized set of guidelines the international statistical community can agree on and possibly submit to the UN Statistical Commission (UNSC) for endorsement by member countries at its next session. This would allow the use of the guidelines not only as a reference document for practitioners, but also as an instrument for promoting international harmonization of survey instruments and comparability of the resulting data. Against the backdrop of this institutional context, the guidelines have multiple aims:

- First, they will provide survey practitioners tasked with designing and implementing HCES in low-income settings with a harmonized set of guiding principles. The aim is to inform the main decisions that need to be taken when designing a HCES, factoring in the objective of serving a wide range of users, without compromising data quality.

- Second, by putting forward a vision for some of these principles, these guidelines can serve as the basis for an international dialogue between practitioners and data users coming from different disciplines and looking for different features in the data.

- Third, a set of guidelines that can be widely shared and agreed upon, can increase the harmonization of the surveys that are implemented (and the resulting data) and give greater coherence to the advice that national statistical offices receive from the international statistical community. Often, different users and institutions push in different directions, resulting in countries adopting very different approaches. Resulting survey design can end up

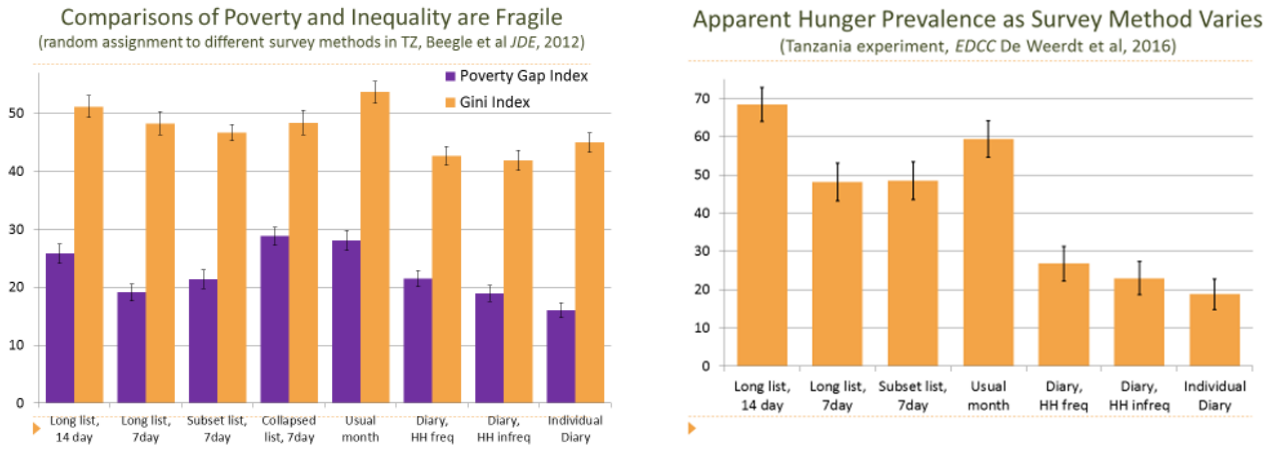
reflecting priorities of donors rather than those of countries and detract from the comparability of data across countries and with other surveys within the same country.

- Fourth, by identifying areas where the consensus is based on a limited evidence-base, the guidelines can serve the purpose of charting the way for an internationally-agreed survey methodology research agenda. Importantly, the guidelines can be an entry point for sustaining an interdisciplinary dialogue for the advancement of this agenda, where statisticians, economists, food security analysts, and nutritionists can come together to contribute to an effective repurposing of HCES that can increase their 'value for money'.

1.3. Emerging issues

The international statistical community is well aware that poverty and hunger estimates are inconsistent across countries and over time, and that lack of harmonization of survey methods plays an important part in that. During the consultative process that generated these draft guidelines, NSOs in several low- and middle-income countries expressed that they do not feel well-equipped to adapt survey practices to the changes that are happening in their societies and food systems, making it difficult to continue collecting data that are relevant and reliable. More concretely, the urgency to address outstanding and emerging methodological issues in survey design for measuring food consumption is immediately apparent if one looks at how adopting different survey design approaches results in markedly different estimates of poverty and hunger, as exemplified in Figure 1 (based on data from the Tanzania SHWALITA study).

Figure 1: Differences in poverty, inequality and hunger measures associated with different survey design options



Source: Beegle et al. JDE, 2012 (left panel); De Weerdt et al., 2016 (right panel)

It is also important for this document to be explicit about a number of areas where the current state of knowledge is not able to provide satisfactory solutions. Some of these challenges are bound to become increasingly important for statistical offices as economies grow and poverty is reduced. To contribute to addressing such questions (some long-standing, some recently emerging) the document also identifies priorities for future methodological research. Deaton and Grosh (2000) identified a number of outstanding methodological research issues in measuring consumption in household surveys and for many of those there has been remarkably little research since their study was published. Furthermore, the research that has been carried out comes mostly from a few countries and datasets (e.g. Bangladesh, Tanzania), casting some doubts on the validity of the findings to other settings. More research is therefore needed to pin down all the relevant survey design issues. The stance taken in this document is that while these gaps in knowledge are being filled, surveys still need to be implemented and advice given; therefore, it is useful to be pragmatic and agree on what is practical, while also being mindful that some perspectives may change as new evidence emerges.

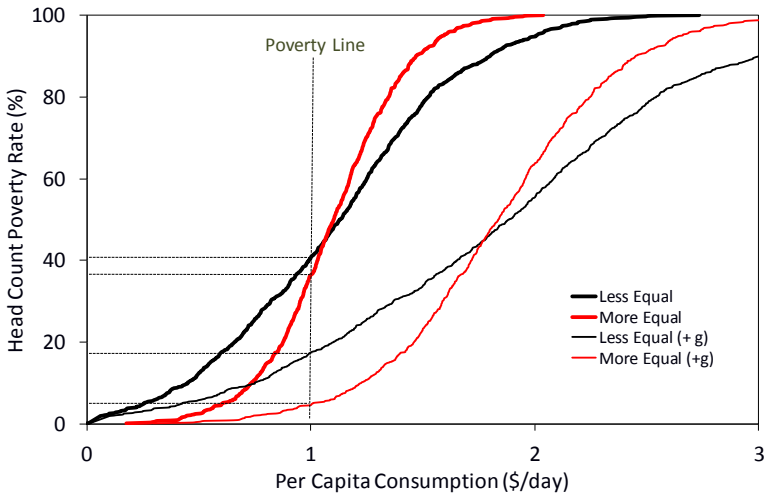
What will be increasingly challenging for national statistical offices and researchers alike, is to keep up with some of the emerging concerns. First, some aspects of the measurement task get harder, not easier, with rising affluence. Respondents become less compliant and harder to survey, particularly in the higher income strata, and NSOs will have to learn to deal more with non-response than is currently the case in low- and middle-income countries. Lessons can be learned from countries that have dealt with these issues in the past, but some degree of new learning will need to happen. To the extent that the following recommendations included in these Guidelines entails a greater burden on respondents, this might aggravate issues of item and unit non-response, and increase implementation costs. Countries should therefore carefully evaluate these trade-offs when implementing changes to their data collection methods. Training, enhance field supervisions, and the use of technology, on the other hand, can help keep some of those concerns in check when complexity increases.

From a policy perspective, inequality is higher on the global agenda than it previously was, but what matters even more for a measurement agenda is that as countries escape mass poverty, headcount poverty measures become more sensitive to differences in the distribution. The intuition for this argument is conveyed by Figure 2. The comparison of the distributions in the figure shows that the same degree of inequality has a much larger impact on the poverty headcount (given a poverty line) when average per capita consumption is higher. Therefore, as economies grow and people rise out of absolute poverty it will become more important for surveys to be able to do a good job of measuring inequality as well as mean consumption. Most of the available methodological survey work is informative about the impact of survey design on central tendency measures, but much less so on what happens to the shape of the distribution[8].

[8] This statement might be somewhat less true for high income countries with the preoccupation with capturing the

Compounding that is the fact that, with per capita income growth, diets become more diverse – income elastic consumption items and foods not eaten from a common household pot become rapidly more important, and these are the food (and non-food) items that surveys are generally less well-equipped at capturing. That is associated with larger measurement error, and measurement error has a greater bearing on distribution-sensitive measures of welfare (such as the severity of poverty and hunger) which are also the measures receiving increasing political attention. It is also difficult to separate measurement error from transitory fluctuations, and as more people escape chronic poverty it is likely that more attention will be paid to transitory welfare shortfalls. These guidelines single out some of the key questions survey methods research should prioritize in order to provide practitioners with the tools to address these challenges.

Figure 2: Headcount becomes more sensitive to inequality as poverty rate falls



Source: John Gibson

The audience for the guidelines is comprised of all those who have an interest in the design, implementation and analysis of high quality, relevant HCES in low- and middle-income countries. First and foremost, these are the national statistical authorities which are tasked with implementing HCES to generate high quality data, to serve the formulation and monitoring of national policies and programs, and to meet the challenges of monitoring the SDGs under the 2030 agenda. The guidelines are also directed at staff in international and national technical organizations providing advice to national statistical authorities in HCES design and implementation.

As mentioned in the previous section, while there are some antecedents to this document, currently there is no single reference document that experts can refer to when advising countries on survey design choices regarding the collection of food consumption data. This results, on the one hand, in shortcomings of survey design (such as those identified by Smith et al., 2014), and on the other, in a

income at top of the distribution is a long-standing one. For low- and middle-income countries measurement error at both ends of the consumption distribution is much less well researched.

high degree of heterogeneity in survey designs which limits comparability. Finally, the guidelines are also directed at survey method researchers and other parties interested in engaging in a global research agenda to advance the knowledge on the data quality implications of different survey design choices from the perspective of different users. The vision is to update these guidelines periodically as new knowledge is generated by methodological research and survey practice, as new challenges to survey design and implementation emerge, and as new opportunities are offered by advances in technology.

The remainder of this document is organized in two parts. Section 2, first briefly summarizes the main uses of food consumption and expenditure data, and desirable features for each of these uses. It then reviews the arguments, theory, and empirical evidence that can inform specific survey design choices. This section is organized around a selection of key choices in the process of designing a quality survey to collect food consumption or expenditure data. Based on that discussion and evidence Section 3 offers practical guidance, summarizing the main findings and offering specific recommendations.

2. Review of the evidence and summary of the main issues

A comprehensive review of the different uses of HCES is provided in Smith et al. (2014). These include poverty measurement, informing food security assessment, providing inputs in the compilation of food balance sheets, providing information for the planning and monitoring of nutrition interventions, informing the compilation of national accounts, and collecting data for compilation of the CPI. These different uses, and the constituencies of users associated with them, have different demands from the data, and depending on the exact nature of the HCES being designed, there are going to be different sets of constraints and opportunities for repurposing. Any attempt at adjusting the design of a survey needs to take into account the analytical needs of the different users. In this document, the main uses considered in setting the criteria for guiding survey design are food security assessments, poverty measurement, and nutrition policy and programming.

Some key issues in the measurement of poverty and food security, and for the monitoring of nutrition interventions, that are useful to understand the data needs connected to these uses are presented in the online Annex 1[9]. In what follows, the document summarizes the literature on key choices that confront practitioners as they design and implement HCES questionnaires. These aspects were identified as priority areas by the review of Smith et al. (2014) and by the experts that participated in the consultation process convened by the IAEG-AG and led by FAO and the World Bank. Several of those issues are also treated, both theoretically and empirically, in a recent issue of the journal *Food Policy*. The volume includes case studies from a diverse set of developing and OECD countries analyzing how different surveys design options affect the quality of the data being collected and, in turn, the implications for statistical inference and policy analysis (Zezza et al., 2017).

[9] Online Annex 1 can be downloaded at http://goo.gl/2EpmXm .

2.1. Recall vs diary and length of reference period

Data on household food consumption (and acquisition) is commonly collected either by asking households to keep a diary over a reference period (e.g. days or weeks) or through interviews in which respondents are asked to recall consumption for a specific period (e.g. last week or last 30 days). A large body of evidence from research and practical experience shows that the method chosen can matter significantly for the resulting estimates of consumption.

We refer to the *recall period* as the period over which respondents are asked to recall the consumption of food items. The recall period differs from the *reference period* when households are interviewed multiple times across multiple visits to the household (Smith et al., 2014). For example, if households are interviewed about their food consumption in the last seven days over four weekly visits, the recall period would be seven days and the reference period 28 days.

In diary methods, households are generally asked to record consumption at the moment in which it takes place (e.g. at meal times or at time of purchase). However, in practice, households often fill in information about their consumption at the end of the day, or during supervision visits from enumerators (for example, for 2-day recall periods if visited every other day). This blurs the line between diary and recall methods, especially when respondent illiteracy is high and supervisors support completion of the diary with visits every few days.

The choice of recall period has long been a critical element of survey design for which there has been limited agreement and evidence on best practice. Scott and Amenuvegbe (1990) suggested the "wide variations [in recall period] reflect the almost total absence of evidence for developing countries on the level of recall error and its relation to recall duration." Similarly, Deaton and Grosh (2000) commented that "there are no definitive answers about the optimal recall period (…). In the meantime, however, surveys must be designed".

This uncertainty is reflected in the large variation observed in the choice of recall periods across surveys. The review of 100 HCES undertaken by Smith et al. (2014) reveals that of the 56 surveys using exclusively interview methods, 26 surveys were using multiple recall periods depending on the source of acquisition or the nature of the purchase (frequently or less frequently purchased). Of the 30 surveys using only one recall period, 13 use a recall period of 7 days, 4 use a recall period of 14 to 15 days, 2 use a recall period of one month, 5 use the 'usual month' or 'usual week' approach, and the rest use a different recall period.

The 'usual month' or 'usual week' approach uses a recall period longer than the month (usually the last 12 months) and aims at capturing seasonality and other short-term fluctuations in food consumption. Households are asked to recall their average monthly or weekly consumption over the last year, sometimes by breaking this into questions about the number of months per year that they consume (or acquire) the food in question, the times per month that they acquire it in those months and the typical quantity and value on each acquisition occasion. Thus, the recall period is the year and the reference period is meant to be the typical month within that year, although there is

evidence that respondents anchor their answers in the economic conditions of the most recent month (Gibson, 2007).

For recall surveys, the challenge is to choose an effective method for measuring the concept of interest while avoiding biases resulting from two main sources: memory decay and telescoping. A longer recall period may be desirable to better capture items consumed infrequently and to obtain a better sense of the true distribution of consumption over a longer time period (addressing the seasonality of consumption). However, one common effect of longer recall periods is memory decay (or 'progressive forgetting' on the part of the respondent, to use the terminology of Deaton and Grosh, 2000), which can lead to underreporting of consumption. Scott and Amenuvegbe (1991) investigated the magnitude of recall error in LSMS-style surveys in experiments with the Ghanaian Living Standards Survey. For 13 frequently purchased items, expenditures reported in the survey fell an average of 2.9 percent per additional day of recall. For seven-day recall, expenditures were 87 percent of what they were for a single day recall; after two weeks, the recall error levelled out at around 20 percent. Similarly, the Indian National Sample Survey conducted experiments on recall period using "last week" and "last month" in which it was found that weekly recall gives food expenditure estimates more than 20 percent higher than monthly recall (NSSO, 2003).

Although a shorter recall period reduces error caused by memory decay, choosing a short recall period introduces another set of problems. As noted by Deaton and Grosh (2000), even under perfect recall, when the recall period is shorter than the period used for analysis, the measure will include variance that does not reflect the true distribution of living standards. A short recall period of one day may eliminate bias in the mean[10], but poorly reflect the distribution of expenditure and consumption over a longer time period (such as month or year), which generally is the key statistic of interest for household surveys. While the 'usual month' approach was advocated by Deaton and Grosh as a way to structure a long recall period to make it more feasible for respondents to answer, while providing analysts with a measure of more typical living standards than is available from a short recall period, the evidence is that this method is not able to overcome the tension between what is feasible to ask and what is desirable to know. In particular, the 'usual month' method has proven to be cognitively burdensome, and it therefore introduces educational-related inequality into the measure of consumption inequality, takes almost twice as long to field as a fixed recall survey over the same foods, and introduces errors on both the extensive and intensive margins (Friedman et al, 2017).

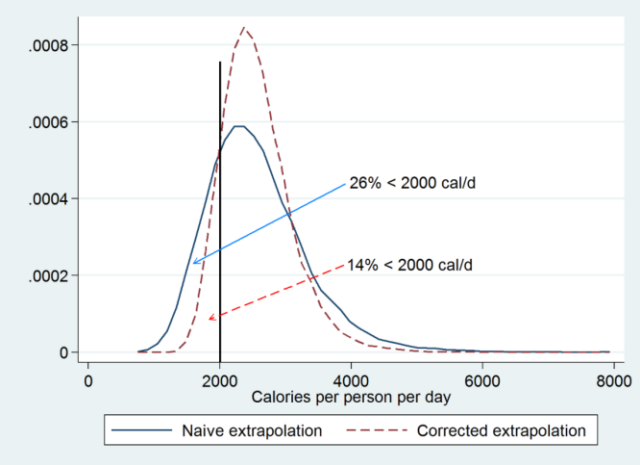
Another strategy used in some surveys is to break the longer reference period into a series of short, adjacent, recall periods. For example, in the Ghana LSMS households are visited up to ten times over one month, so that there are only short recall periods between each visit. A similar design is present in many diary-keeping surveys for illiterate respondents, who may be visited every day or second day over the 14 to 28-day reference period. While there may be some novelty for a respondent being interviewed the first time, a high frequency of repetitious interview visits is likely to induce non-

[10] Assuming the sample of household is sufficiently distributed throughout seasons

compliance, and clear evidence of that is shown in the Ghana LSMS by Schündeln (2017) who finds that data quality is highest for the first interview and falls monotonically with each successive interview. Thus, measured food poverty would be 13 percentage points higher if all ten interviews of the same household are used, compared with using just the data from the first visit.

The tradeoff between shorter reference periods that allow recall over that same period to be less prone to forgetting, but which provide a poor guide to typical, long-run, living standards also affects studies focused on estimating average daily per capita dietary energy consumption (DEC).[11] In particular, shorter reference periods are found to affect the variability in energy and nutritional estimates. Using data for Myanmar collected over two monthly rounds per household approximately 6 months apart, Gibson (2016) annualizes estimates of daily calories per capita from each survey round in two ways: a naïve extrapolation that multiplies estimates from each round by six and then adds them, and a corrected extrapolation which is based on the intra-year correlation in daily calorie per capita across survey rounds of 0.45. The implications for measures of hunger by doing this correction is exemplified in Figure 3. Given two distributions with the same median calories per capita per day, the one based on naïve extrapolation of the monthly data will have a greater dispersion compared to the adjusted one, resulting in a greater incidence of hunger for a given threshold (2000 kcal/day in this case)[12].

Figure 3: Chronic Hunger Overstated by Naïve Extrapolation from Monthly Calories to Annual



Source: Gibson (2016).

An analysis performed by FAO Statistics Division using the 2010 data of the Bangladesh Household Income and Expenditure Survey (BHIES) illustrates how the variance of the per capita DEC is significantly reduced over longer observation periods. (Box 1).

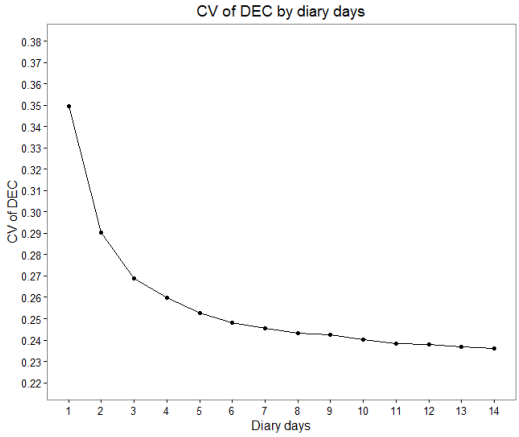
[11] This is a key variable in the measurement of undernourishment, as estimated by the FAO and reported in the *State of Food Insecurity in the World Report* series.

[12] To estimate the prevalence of undernourishment (SDG indicator 2.1.1), FAO is using the Minimum Dietary Energy Requirement that depends on the age, sex structure of the population, the fifth percentile of the Body Mass Index distribution for adults, the height of the individuals in the population and the average of the sedentary lifestyle range for physical activity levels in the country. This threshold is not fixed and varies from one country to the other.

Box 2. Case study: Bangladesh 2010 Household Income and Expenditure Survey

The Bangladesh Household Income and Expenditure Survey 2010 (BHIES) was carried out from February 2010 to January 2011. Within these 12 months of investigation, the survey was divided into 18 periods, each 20 days long. Food consumption was collected through a diary over a period of 14 days. Throughout the period households were visited frequently (from 7 to 14 visits).

The Dietary Energy Consumption (DEC) was estimated for each day of the diary and the figure below shows the coefficient of variation (CV) of DEC obtained using different numbers of diary days. From the plot it can be clearly seen that the variability is highest (CV=35 percent) when using observations from the first day only, and decreases convexly and converges at a value of CV of around 24 percent after the 7th day. After the first week, the variability does not seem to decrease much, suggesting that a reference period of 7 days might be enough for estimating the variability of DEC.



Source: Grünberger (2017a).

Short recall periods can also lead to telescoping, where respondents report consumption that took place outside the reference period, also causing a bias in estimates. Several studies suggest that there cannot be one optimal recall length since telescoping and recall decay will behave differently depending on the type of good and the frequency of consumption (Bradburn 2010, Hurd and Rohwedder, 2009). In general, telescoping is more likely for large and infrequently purchased/consumed items under shorter recall periods, while a longer recall period leads to recall decay and underreporting of more common and frequent purchases (Deaton and Grosh, 2000; Moltedo et al, 2014; Neter and Waksberg, 1964). The Indian National Statistical Office (NSSO) designed an experimental survey including three types of data collections: daily visits with direct measurement (benchmark), seven-day recall, and 30-day recall by food group. The report of NSSO (2003) shows that optimal recall period depends on the food group and frequency of consumption. The 30-day recall works better than 7-day recall in measuring staple food like cereals and is not inferior in measuring high-frequency items. One explanation for these patterns is that staple foods, and other high-frequency items, lend themselves to more accurate 'rule of thumb' reporting, based on their regularity (Friedman et al, 2017), so strictly speaking these are being 'estimated' rather than

'recalled', and with a 30-day period the effect of any telescoping is diluted compared with a 7-day period.

A possible solution for dealing with telescoping is a bounded recall approach, suggested by Neter and Waksberg (1964). In this approach, a first visit to the household by the enumerator is used to establish the bound of the recall period for a second visit which is when the actual interview takes place. The enumerator can thus ask the respondent about consumption since the first interview, reducing the likelihood of respondents reporting consumption or acquisition taking place outside the recall period. This approach can however be costly to administer as it requires two visits to the same household, and despite having a long history (including within the LSMS program) there have not been enough studies to provide convincing evidence that it brings a significant advantage in terms of data quality (Deaton and Grosh, 2000; Gibson, 2005). This is therefore an area where future methodological research could usefully focus in order to establish whether such advantage exists in practice.

Diaries present an approach which in theory can deal with important shortfalls of regular (longer) recall methods such as telescoping and recall bias. They are in fact the method of choice and are successfully implemented in many countries for collecting data on food and other frequent expenditures. However, they can be practically challenging to implement in the conditions prevalent in many low- and middle-income countries. Diaries are far more demanding in terms of supervision, especially with illiterate respondents when they are often implemented as a series of short recall interviews, and so become more expensive and demand higher capacity. While a well-implemented diary is generally considered the 'gold standard' for measuring consumption, poorly implemented ones are often inferior to a good recall survey. Even in the US context, where the set of challenges for diaries and recall may be different than in lower income countries, evidence suggests that recall surveys might outperform diaries (Bee, Meyer and Sullivan, 2012).

A growing body of research has shown how the diary method causes considerable response burden and fatigue particularly when the length of the diary increases, ultimately impacting data quality and reliability. Studying the Canadian Food Expenditure Survey, Ahmed et al. (2006) find a decrease in reporting due to 'diary exhaustion' with reporting decreasing by 10 percent from the first to the second week of filling diaries. Similarly, studying the US, Stephens (2003) finds significantly higher values in the first diary week as well as on the first day of each diary week relative to the remaining days, attributable to respondent fatigue. Analyzing the 2009/10 data of Papua New Guinea, Gibson (2013) also finds that the total value of consumption transactions declined by 4.4 percent per day of diary-keeping period. A large set of other studies, such as Kemsley (1961), Turner (1961), Sudman and Ferber (1971), McWinney and Champion (1974), and Silberstein and Scott (1991) find similar evidence of fatigue and decay in information collection in diaries over time.

With high levels of supervision and careful implementation, diaries can and are being implemented in some countries, with good results. In analyzing the Bangladesh 2010 HIES, FAO Statistics Division does report a negligibly low decrease of DEC due to fatigue, likely because of very good respondent

supervision practices, with enumerator visits taking place every 1 or 2 days[13]. Such levels of supervision lead to a mix of diary and interview methods and are, however, not likely to be commonly affordable; diaries do not therefore appear to be the most suitable method for resource-constrained statistical offices in low- and middle-income countries. Furthermore, even for well-implemented diaries the evidence clearly suggests that longer periods of implementation do not add to the quality of information (they actually detract from it) while entailing higher implementation costs. The expanding mobile phone coverage throughout the world opens possibilities for remotely assisting diary completion (as well as recall interviews) at a fraction of the cost. This is an emerging trend (or an established one is some high-income countries), but not something for which there is enough experience at scale in low-income settings for it to be recommended as common practice at the time these Guidelines are being formulated.

One final aspect of diary implementation is that often the analyst is presented with data that have been already to some extent aggregated (e.g. by adding up the 7 or 14 days of data) which does not allow for detecting and correcting possible patterns in the data such as diary fatigue (Troubat and Grünberger, 2017). When diaries are implemented it is therefore important that they be reported together with full metadata, allowing the user to evaluate the data collection process, including the role of the enumerator in aiding the data collection processes, the number and timing of supervision visits, and similar details.

In the LSMS Handbook Deaton and Grosh (2000) provided a discussion of the issues we have just outlined and concluded by recommending only changes on the margins of the LSMS status quo. Specifically, that meant using bounded recall for purchases, coupled with a usual month question for both purchases and consumption of food from own production, plus one 12-month recall question on the value of food gifts received by the household. Deaton and Grosh had already observed, however, a decline in the actual use of bounded recall in LSMS survey practice for reasons related to the added cost and burden (for enumerators and respondents) of the additional household visit. While pointing to the pros and cons of the usual month and of shorter recall periods (progressive forgetting, telescoping, difference from the 'true' variance) as discussed above, in recommending the usual month approach they also recognized that this was based on weak and often contradicting evidence, and mostly motivated by the desire to modify the 'status quo' at the margins, in the absence of stronger evidence in favor of a particular approach.

Despite the lack of conclusive evidence lamented by Deaton and Grosh, and their call for "every survey [to] have a budget for experimentation" there has been a limited number of new studies undertaken in low- and middle-income countries focusing on these methodological questions. One that has been particularly influential is the SHWALITA study ("Survey of Household Welfare and Labour in Tanzania") (Beegle et al., 2012; Gibson et al., 2015; de Weerdt et al., 2016). New evidence has also been produced through the work reported by Backiny-Yetna et al (2017) in Niger. Based on these studies and increased practical experience, practitioners in living conditions surveys have come

[13] Food consumption reporting dropped on average by less than 0.1% per diary day.

to favor a 7-day recall period over longer reference periods. Deaton and Grosh had already noted signs of the bounded approach falling out of fashion with practitioners due to its higher complexity.

The SHWALITA study (Beegle et al., 2012) provides convincing evidence, from an experimental setting, that recall interviews asking about 'usual' monthly consumption food underestimated household consumption expenditure when compared to the benchmark assisted individual diary (see Figure 1 above), whereas the 7-day recall was reasonably close to the benchmark. At the same time, the usual month interviews also had the longest completion times (76 minutes compared to just under 50 minutes for the 7- and 14-day recall), and were not associated with a significantly smaller coefficient of variation when compared to the shorter recall methods. Besides the resource implications of longer fieldwork time, the longer completion time for the usual month approach is suggestive of a greater burden on the respondent who, with the enumerator, needs to engage in a demanding estimation procedure to work out the response for a typical month starting from recalling consumption episodes over a 12 month period. Taken together this evidence indicates that the usual month may be a lose-lose proposition if it is both less accurate and more cumbersome to implement when compared to a 7-day recall. This is possibly the most important single development in the evidence base since the publication of Deaton and Grosh.

Importantly, another plea made by Grosh and Deaton two decades ago remains unanswered and just as valid today. Since changing the recall period or method will lead to incomparability issues with previous surveys using other methods, changes in survey methods over time should be accompanied by an experimental study to allow reconciling the figures produced by the survey before and after the change in methods. Experiments such as Beegle et al (2012) and Backiny-Yetna et al (2014) have provided good practical examples for how changes in methods can be assessed and thus allow for valid comparisons when methods are changed.

2.2. Seasonality, number of visits

Consumption and expenditure patterns often show seasonal variations that are, often linked to the agricultural production season, cyclical events such as floods and droughts, or cultural events (e.g. Ramadan, Christmas), which affect food availability and prices, as well as customary consumption practices. The existence of seasonality in food consumption patterns is well-established (Paxson 1992, 1993; Alderman, 1996) but its extent depends greatly on context.

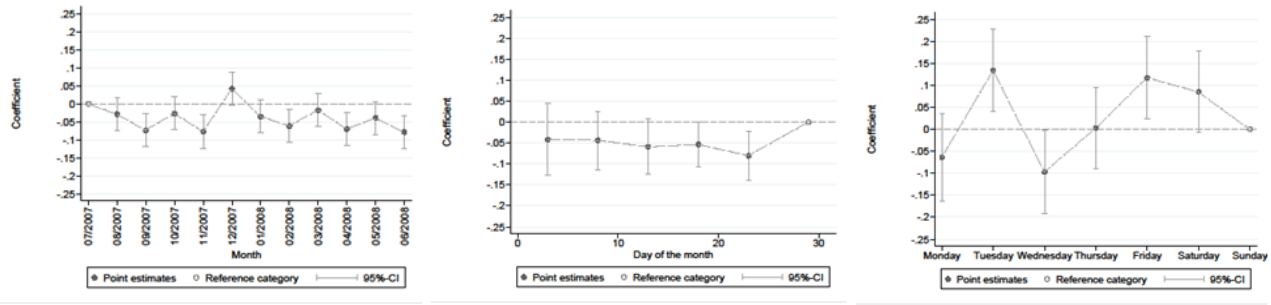
Seasonality can be particularly important for food consumption, because seasonal variations in dietary patterns, overall quantities of food consumed, and the consumption of particular nutrients can be pronounced, partly due to the relationship with food production cycles (Coates et al., 2012). D'Souza and Jolliffe (2012) found that household consumption in Afghanistan can be as much as one third lower in the lean season compared with the post-harvest season. The different levels of consumption, if taken at face value, would result in estimates of the poverty headcount doubling

from 23 percent in the fall to 46 in the following summer (D'Souza and Jolliffe, 2012). A key concern is the seasonality in food prices which is found to be significant and can affect estimates of poverty and consumption (Gilbert et al, 2016). That is of course a major issue for surveys collecting data for the calculation of the CPI.

Seasonal variations can also originate from increased expenditures during festivals and holidays. In the US, it has been established that consumption is higher during holidays and summer months (Barsky and Miron, 1988). In low- and middle-income countries, expenditures can vary significantly with holidays, festivals, and religious observances. Jolliffe and Serajuddin (2015), using data for Jordan, note that during the period of Ramadan consumption levels are 11 percent greater than during other periods of the year. These festive expenditures can be difficult to capture in surveys because it is often difficult for survey fieldwork to operate as normal during festive periods. A few surveys, such as the Living Standards Surveys in Vietnam, use a special recall module for food consumption during festive periods, with analysts typically then spreading this over the consumption estimates for the households observed in the rest of the year.

Within-year temporal variation can originate from other patterns such as those associated with payment schedules for wages or social assistance (Stephens, 2003). Troubat and Grünberger (2017) in studying the urban subsample of the 2008 Socio Economic Survey of Mongolia found cyclical variations in food acquisition levels not only between months but also within months and weeks. Independence Day, a major national holiday occurring in December (see Figure 4, left panel) results in the mean consumption in that month being substantially higher than for the other months of the year. A systematic pattern is also apparent for data collected at the end of each month, where food consumption is significantly higher (center panel). Households tend also to spend significantly more on food on Tuesday and Fridays than they do on Mondays and Wednesdays (right panel).

Figure 4 - Variation in mean food consumption by month, day of the month, and day of the week. Urban Mongolia.



Source: Troubat and Grünberger, 2017.

Even if such patterns are difficult to generalize given the context specificity, they remain an important example of sources of bias that should be mitigated to the extent possible in survey design (Fielder and Mwangi, 2016). If seasonality is not taken into account when there is marked seasonal variability in food consumption, the use of short reference periods will bias the estimates of both the mean and the standard deviation of the distribution of habitual food consumption in the population.

Recorded mean consumption might be higher or lower, depending on the season when the data are collected, and estimates of the CV might be biased by the confounding effect of seasonal variation.

A survey carried out at a specific time of the year (say a season, month, or week), will miss seasonal variation in consumption and risk being unrepresentative of typical consumption across the year, even when it manages to accurately capture consumption over the period of data collection. Also, surveys that are not adequately capturing the entire year will pose problems for international comparability. Comparisons of consumption data for a country conducting a survey in the lean season and one conducting it in the harvest season will be hard to make in the absence of elements allowing to gauge the habitual consumption levels in both countries. Even over-time comparisons of surveys undertaken within the same country and during the same period of the year might be invalidated if major events correlated with consumption patterns move in and out of the survey implementation period. This may happen with Ramadan, the dates for which move from year to year, or when harvest periods are delayed or pushed forward by weather events. For all these reasons it has been recommended that HCES should cover a full-year to properly capture seasonal variations in expenditures (ILO, 2003), although this is by no means a universal practice.

Deaton and Grosh (2000) suggested the use of a "usual month" approach to overcome seasonal variation, but in the previous section we have seen how the reliability of that approach appears questionable, and is associated with longer interview times and heavier cognitive burden on the respondents. Deaton and Zaidi (2002) suggest that for capturing household consumption, the optimal survey implementation and design is the one likely to provide the most precise estimate of annual consumption for each household, not just for households on average. Based on this objective, the ideal design is one where households are visited each "season" and habitual consumption is then derived as an average over the year of seasonal consumption. For variance-based measures, these intra-year revisits allow the possibility of corrected extrapolation, along the lines of what is shown for estimates of hunger in Gibson (2016). A drawback of revisiting the same households is the cost and the trade-off with overall sample size since for any given sample size, the survey costs increase with the number of multiple visits.

Just over half (53 percent) of the surveys reviewed by Smith et al (2014) took seasonality into account, by using one of two approaches. The first (adopted by 41 percent of assessment surveys) is to distribute data collection throughout a year by surveying subsets (usually one-twelfth of households in the sample) in each month of the year, with subsamples representative nationally for each quarter. This approach (which conforms to the ILO recommendation) requires careful planning of the sampling strategy and survey implementation, but can ensure that the seasonal variation across space and time is captured, at least for a 'synthetic' household albeit not for any particular household in the sample, and represents a lower burden on households as they are visited only in one period of the year. This method can present advantages in terms of organization when survey staff are employed to just work on one survey, as it smooths the need for the workforce over the survey year and can allow working with smaller teams hence ensuring tighter supervisions.

The second approach (used in 12 percent of the surveys reviewed) is to conduct two to four visits during the year on the same households. It has been noted above how additional visits to the same households come at a cost, pose logistical challenges, and increase the burden on the respondents. This option can, however, be attractive when the survey also has other objectives, such as collecting data on agricultural activities, to the extent that the visits can be timed around salient moments for both objectives; for example different points in the agricultural season (e.g. post-planting and post-harvest) as it done in the Nigeria National Panel Survey (2010-11, 2012-13, 2014-15) and in Niger's Enquête Nationale sur les Conditions de vie des Ménages et l'Agriculture (2011, 2014). Finally, three of the surveys reviewed by Smith et al. collected data in four visits over a 12-month period. This approach is widely deemed as very difficult to implement and excessively cumbersome in terms of organization, with the burden on respondents at its highest. Some countries implement two rounds of data collection in different periods of the year but on a different cross-section of households. That does away with the added interview burden for the household, but only allows controlling for seasonality for the sample aggregate and not for each specific household, and does not provide an opportunity for correcting variance-based measures for excess variability due to intra-year fluctuations.

Forty-seven percent of the surveys reviewed by Smith et al. (2014), do not satisfactorily account for seasonality in their design, as they are implemented via one household visit over the span of a few months. This approach returns data that are subject to all the biases discussed at the beginning of this section, but it is quite common because of its ease of implementation. When staff are employed by different surveys implemented by the same statistical office, the approach allows the NSO to move them on to the next project when the HIES fieldwork is completed. It is often also motivated by the idea that if the period of implementation does not vary then at least over-time comparison would be safeguarded. This is however a questionable assumption as the extent and timing of seasonal variations may not be the same from one year to the next. Also, any unforeseen implementation issues that would delay the onset of the fieldwork from its planned schedule would invalidate over-time comparability of the data. These are very serious concerns that point towards the need to abandon such practice for one that takes seasonality fully into account.

A hybrid approach that could be experimented which at least partly limits the more serious shortcomings of the 'one visit over a few months' approach, would be that of complementing that one visit with a second visit on a subsample of households. This additional subset visit could provide the information required to correctly annualize the data from the first visit. This is however only a hypothetical survey design that would need to be carefully tested before being applied at scale, and as such it represents at best an indication for further research.[14]

[14] An example of this design, using a 20% sub-sample who were revisited approximately five months after the initial visit, is provided by Gibson (2001). The revisit was estimated to add about 10% to total survey costs, and enabled a partitioning of poverty estimates into chronic and transient components.

2.3. Acquisition vs consumption[15]

Early on, food consumption data were mainly collected in HCES (in particular in Household Budget Surveys) to construct the consumer price indices or to inform national accounts. For such uses, the interest was mainly in collecting data on food items acquired through expenditures. Obtaining food through expenditures is now widespread throughout the world and is the prominent form of food acquisition in many locations, especially urban areas. However, in many countries a considerable share of households obtain some food from their own production, whether from crop cultivation, livestock rearing, fishing and aquaculture, and/or hunting and gathering. It is also quite common for households to obtain some of their food in-kind, whether in the form of gifts from other households, payments from an employer, or public or private assistance (school feeding, food assistance programs, social or private transfers in-kind).

As HCES became increasingly used for poverty and food security analysis, the emphasis has shifted towards also collecting data on food items obtained through not only expenditures but also through these other sources of acquisition. Accessibility is one of the dimensions of food and nutrition security as defined by the FAO, and that includes access to food from all possible sources. For poverty analysis, too, all sources of food acquisition enter the consumption aggregate, not only those that imply an outlay of cash. In nutritional assessments what matters is what is actually ingested. Again, that implies a focus on consumption (or more specifically, intake) of food regardless of how it was acquired. Understanding how food systems work and evolve, what share of foods households in different socio-economic groups acquire from different channels, what the relative prices are for households in different locations, and what the nutrition and welfare implications are, requires having access to food data.

For national accounts, food produced for own consumption is part of the household final consumption expenditure. Getting information on own-account production and consumption of food (as well as other goods) by households is thus critical, even though agriculture surveys or censuses may also provide this information. Such food should be valued at “basic prices” of similar goods, which can be approximated by the price of similar goods sold on a local market, or the price declared by the household producer if he/she had sold the food rather than consumed it. Information on food/ meals acquired through in-kind transfers is also important. Valuation should be based on actual cost if actually purchased by the provider or production cost otherwise, both being unknown and difficult to evaluate by the beneficiary.

Not all surveys, however, are designed to capture information on all the food that is consumed or available for consumption in a household from all the sources of acquisition. Three different approaches to collecting food data can be identified, following Conforti et al. (2017):

[15] This section is based on and reproduces parts of Conforti et al. (2017).

- *Acquisition*. Households report on food they acquired through purchases, own production and in-kind transfers. Actual consumption of the same food is not reported.

- *Combination of acquisition and consumption*. Households report on acquisition for food they incurred an expenditure for (purchases), without specifying the amount of food consumed. Food consumption derived from own-production or received from transfers is reported.

- *Consumption*. Households report on food consumed, and on whether that same food was purchased, own-produced or received as a transfer.

Differences in food measurement between surveys focused on acquisition and consumption are not always clear. In principle, the difference between the two measures is essentially a change in stocks, and food wasted in households. For surveys based on efficient representative samples, homogeneously spread across time and space, changes in stocks should on average be close to zero. In any given reference period some households may build stocks while others may consume food from stocks. However, surveys with less effective timing of household visits may show significant differences between acquisition and consumption (e.g. if the survey is implemented in one visit when most households are stocking, or de-stocking).

Smith et al. (2006) provide a general discussion about the difference between estimates of consumption and acquisition. Depending on the length of survey coverage and reference period, the distribution of acquired food is expected to have a higher variance and a higher mean than the distribution of consumption. The variance of acquisition surveys is higher because daily food consumption is smoother than acquisition. This difference is expected to decrease to zero as the length of the survey period increases. During the reference period, households can either consume from stocks (under-estimating household consumption using an acquisition survey) or build stocks (over-estimating household consumption using an acquisition survey). As a consequence, households can have zero expenditure during a given reference period, albeit consuming from stocks (Gibson and Kim, 2012). Acquisition surveys should be used to approximate aggregated consumption of population groups, rather than habitual consumption of individual households. Acquisition data are assumed to have a higher mean than consumption, because food waste, rotten stocks, or food given to pets is already detracted in consumption estimates (Smith et al., 2006). However, empirical studies suggest that the difference between averages of food acquisition and consumption is not always positive, but sometimes close to zero or even negative (Kaara and Ramasawmy, 2008; Martirosova, 2008; Smith et al., 2006; Bouis et al., 1992; Bouis, 1994). Conforti et al (2017), analysed 81 HCES[16] conducted between 1988 and 2014 and found that the average dietary energy consumption from surveys focusing on acquisitions was only slightly higher than that from surveys focusing on consumption, but the variability was in turn much higher (an average coefficient of variation of 76 compared to 52).

[16] Surveys analyzed by the FAO food security and nutrition statistics team from 2006 to 2014, using the ADePT-FSM software developed jointly by FAO and the World Bank (Moltedo et al., 2014).

Though the difference in the aggregate measure is not that significant, the difference in the CV is of real concern for FAO which is using the coefficient of variation derived from food data collected in HCES to estimate the prevalence of undernourishment. FAO has developed a methodology to overcome the issue of excess variability encountered in the food consumption measurement (Wanner et al., 2014) . Troubat and Grünberger (2017) applied this methodology to the 2007/08 Socio economic Survey of Mongolia that collects both food consumption, and acquisition[17] and they found that the difference in variability that exists between DEC from acquisition and DEC from consumption disappears after both distributions are corrected for excess variability (CVs decreased from 63 to 31 percent for food consumption measurement based on acquisition type data and from 52 to 30 percent for food consumption measurement based on consumption type data).

The decision to collect acquisition or consumption data is not expected to have a large impact on the estimate of the prevalence of undernourishment, nor poverty, as there is no significant effect on the mean and the impact on variability can be reduced using the control for excess variability. However, for countries with a large population and low average DEC, a small difference in kilocalories per capita per day can still affect food security and nutrition assessments.

Besides these general questions, there is a more specific – but not less important- set of risks associated with survey design not taking explicitly into account the difference between consumption and acquisition. According to the review performed by Smith et al. (2014), in surveys based on recall interviews it is not uncommon for questionnaires to include poorly worded leading questions or other forms of design ambiguity that can lead to incomplete enumeration of foods consumed. Such issues arise from the survey design failing to properly account for the fact that not all the food acquired by a household is consumed during the survey reference period, and that food can be consumed during the reference period that was acquired earlier. Their findings are reproduced with minor changes in the reminder of this section, and summarized in in Table 1.

[17] The latter survey measures a household's food acquisition, and food stocks at the beginning and the end of the reference period. Combining the information of acquisition and stock variation, the household's food consumption can also be derived from food acquisition.

Table 1. Completeness of enumeration of food acquisition and/or food consumption

Table 1: Completeness of enumeration of food acquisition and/or food consumption			
	Interview	Diary	All
	(percent)		
Whether acquisition or consumption data are collected			
Acquisition	36.1	48.7	41.0
Consumption	36.1	10.3	26.0
Both	27.9	41.0	33.0
Problems of incomplete enumeration			
Rule-out leading question on consumption	13.1	0.0	8.0
Rule-out, short-recall-period leading question on acquisition	3.3	0.0	2.0
Rule-out leading question on food purchases	1.6	0.0	1.0
Own production question refers to food harvested rather than consumed	3.3	0.0	2.0
Ambiguity whether to report on acquisition or consumption	6.6	5.1	6.0
"Usual month" surveys: Ambiguity whether to report consumption in any month or months with positive consumption	13.1	0.0	8.0
Percent of surveys with problems of incomplete enumeration	37.7	5.1	25.0

Note: N= 100 surveys.

Source: Smith et al. (2014)

For the food data in HCES to be reliably collected, a full accounting of either all acquired food intended for consumption or all food that was consumed over the recall period must take place. Additionally, only the food intended for consumption (when acquisition focused) or consumed (when consumption focused) during the reference period must be included, not any additional food. The following exclusion and inclusion accounting errors can plague the collection of HCES food data:

(1) Acquisition surveys: Rule-out leading question on consumption. If a leading or "filter" question on consumption of each food item over the recall period is answered "no," it can rule out collection of further data on the acquisition of the food. In this case, respondents are first asked whether they consumed each food item in the food list for a recall period of up to one year before the time of the survey. Then they are asked how much was purchased, consumed from own production, and/or received in kind over the survey recall period for food data collection. However, if the respondent answers "no" to the leading question, the instructions skip to the next item, then the household receives a zero for acquisitions of the food item regardless of whether or not it was acquired. This leads to systematic underestimation of the quantities and/or expenditures on food acquired. A rule-out leading question on consumption is considered to be a problem when the two recall periods are less than or equal to two months apart. Note that this issue does not afflict diary surveys because there is no pre-listing of foods to rule out.

(2) Acquisition surveys: Rule-out, short-recall-period leading question on acquisition. Here, if answered "no", a short-recall-period leading question on acquisition of each food item rules out further data collection on the acquisition of the food over the (longer) survey recall period. In this

case respondents are first asked whether or not they acquired each food item over the short recall period (e.g., two weeks). Further information is collected on the acquisitions of the food for the longer recall period for food data collection only for those food items that were acquired over the shorter recall period. This leads to underestimation of mean food acquisition for the population.

(3) Acquisition surveys: Rule-out leading question on food purchases. In this case if a respondent reports that the household did not purchase any of one food item, then no further information is collected on other forms of acquisition of that food item. Since home-produced or in-kind receipts are left out, this problem also leads to underestimation of mean food acquisition for the population.

(4) Data collected on food harvested rather than food consumed from home production. When interviewees declare food harvested instead of food consumed from own production or food from own production for consumption, the quantities and/or expenditures on food acquired include those entering into the households' production stocks – not the household pantry for immediate consumption – and are systematic overestimates of food consumed from home production. A similar situation occurs when there are household animals, like poultry and pigs, who may eat some of the food that was harvested from household food gardens (e.g., undersized tubers, and food that is deemed as otherwise unfit for human consumption given the food availability at the time).

(5) Ambiguity about whether to report on acquisition or consumption. The question asked to respondents does not make it clear whether they are expected to report on their acquisitions or consumption of each food item over the recall period. This problem leads to inaccuracies in calculation of the mean acquisition or consumption for the population as well as measures of inequality.

(6) Routine month surveys: Ambiguity about whether respondents should report on the routine month in the recall period or only on those months in which the food item is actually consumed. In many routine-month surveys, respondents are first asked to report on the number of months in the last year in which each food item was consumed. Immediately following, they are asked about the usual or average amount per month. Some questionnaires, however, fail to specify whether or not the average should be for those months in which it was consumed or for any month in the last year. When this type of accounting error occurs, some households may report on the former and some the latter, leading to over- or under-estimation of their consumption of any food item for which a positive number of months was reported for the initial question.

As can be seen in Table 1 above, 11 percent of the assessment surveys suffer from the use of the three types of rule-out leading questions. The collection of data on food harvested rather than food consumed from home production is a relatively rare problem from which only two percent of the surveys suffer. A full 14 percent of the surveys had problems of ambiguity in what is to be reported, which likely leads to incomplete enumeration for some households. The problem of ambiguity in expected reporting for routine month surveys was identified in eight percent of the surveys. Overall,

25 percent of the surveys did not meet the reliability criterion for completeness of enumeration, that is, they were afflicted by some of the identified problems of incomplete enumeration. Note that the large majority of the surveys that have these types of accounting problems are interview surveys.

2.4. Meal participation

The size of a household is only a proxy for the number of food consumers in the household during the reference period. Per-capita measures of food consumption should be based on the number of people who actually take part in the meals (food partakers) in the household (Smith et al., 2014; Weisell and Dop, 2012). People other than household members who could take part in the household's meals include employees who had their meals in the household, guests, and visitors. The number of food partakers should exclude household members who were not present in the household during the reference period.

Adjustment for food partakers might not be an issue for poverty measures, but collecting this information becomes crucial for the analysis of habitual per capita food consumption and food security estimates. Indeed, nutrient inadequacy is assessed with reference to requirements that are expressed in per person basis. Household surveys collect information on total amount of food consumed by households over a certain reference period. To convert this information to a per capita basis, it is important to account for meal participation in the household. The most common way to do this is to consider the number of people who consumed the total amount of food reported by the household.

Box 3. Estimating average per capita DEC

Based on household size and number of partakers, per capita DEC of a household *i* can be calculated in two ways. First by dividing the total number of daily calories consumed in a household by the exact number of people who participated in the meals

$$DEC_i^P = \frac{TotalDEC_i}{Partakers_i}$$

or, if the above is not available, by dividing total household calories by the number of household members

$$DEC_i^{HH} = \frac{TotalDEC_i}{HHsize_i}$$

In the latter case, food consumption will be underestimated if mean consumption is calculated on the basis of household size. When food is provided also to non-household member partakers, the total food consumption in the household increases. The household's mean consumption should be correctly calculated by dividing total

household food consumption by household size plus additional partakers minus absent household members. In omitting the additional partakers from the calculation, the denominator would be smaller and the household's mean consumption overestimated. If absent household members are not subtracted from household size, the denominator would be higher and household's mean consumption underestimated.

But there is no standard approach to capturing information on meal participation in households, and many surveys fail to collect that information. Smith et al. (2014), in their survey assessment, find that only 15 of 100 surveys ask whether non-household members were present or consumed meals in the household during the recall period. Even within this small number there is variability on the additional information collected, with 11 asking about the number of visitors in the household, 10 about the number of meals they consumed, and a handful asking also information on length of stay, type of meals they consumed and age and sex of the guests. Fiedler and Mwangi (2016) provide a thorough description of approaches to collecting partakers. Looking at 17 recent surveys they found 16 different possible questions combined in various ways, with no two countries collecting the same information on partakers (Table 2).

Table 2. Approaches to collecting information on partakers in 17 HCES (from Fiedler and Mwangi., 2016)

HCES Collecting Data for Developing more Detailed Estimates of Meal Attendance By Number and Type of Meal, and Number and Level of Participation

Topic		Data Item Collected	HCES Collecting	No. of HCES Collecting
A. Meals				
	1	Usual number of meals eaten daily	6,9,11,14,15,16,17	7
	2	Type of meal eaten (breakfast, lunch, dinner, snack)	10	1
	3	Type of meal eaten away from home (breakfast, lunch, dinner, snack)	9,15,16	3
	4	Total number of meals served	9,12	2
B. Person-specific data				
B.1. Household Members				
	5	Present during reference period? (Yes/No)	1,2,5,9,10,11,15,17	8
	6	At least 1 meal eaten at home during recall period?	1,2,5,6,7,8,9,11,16,17	10
	7	Number of days ate in the household during recall period?	11,16	2
	8	Meals eaten away from home? (Yes/No)	9,16	2
	9	Number of meals away from home	9,16	2
	10	Number of days away from home	9	1
B.2. Non-household Members / Guests				
	11	Were any guests present during the reference period? (Yes/No)	6,9,10,11,13,14,15,17	8
	12	Number of guests present	6,10,11,13,14,15,17	7
	13	Number of days guests were present	9,11,13,15	4
	14	Number of meals served to guests	1,11,15,16	4
	15	Type of meal served to guests	15	1
	16	Characteristics of the guests (age, gender)	3,4,7,8,11,16,17	7

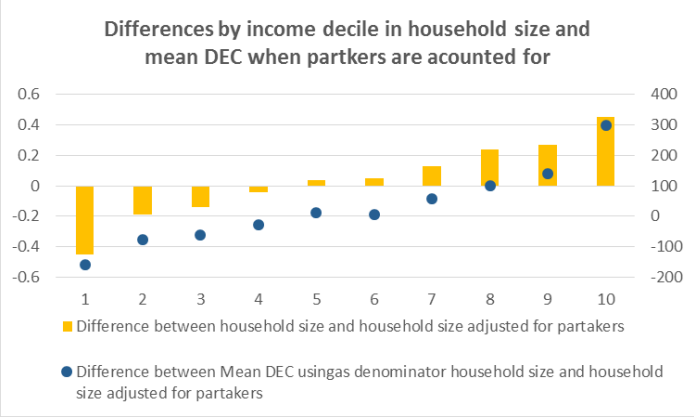
Source: Fiedler and Mwangi (2016)

Partaker correction should in theory have no impact on the overall sample mean of per capita food consumption because positive and negative deviations from the household size will balance out[18]. On average, the household size should be equal to the number of partakers. The multivariate analysis of 81 HCES[19] conducted between 1988 and 2014 (Conforti et al., 2017) found no significant difference between the mean and the CV of DEC per capita accounting for partakers or not, when controlling for other survey characteristics. But there is empirical evidence to believe that not accounting for partakers distorts the distribution of per capita DEC.

Bouis et al. (1992) and Bouis (1994), in analyzing surveys from Kenya and Philippines that both collect partakers, show that the relative difference between mean DEC of the first and fourth quartile is much lower when partakers are accounted for. Likewise, using an urban survey Gibson and Rozelle (2012) show how using the roster of meal partakers lowers the apparent calorie availability of the richest quartile by 7% and raises the calories of the poorest, where this pattern results from a coping strategy of the poor, which is to visit their richer kinfolk at meal times. These studies deliver therefore an evidence that the variability of DEC conditional on household income is lower if data are adjusted for partakers.

Results in line with those of Bouis et al (1992), are confirmed by a similar analysis conducted by FAO on food data collected in the 2010 Bangladesh survey (Grünberger, 2017b). In that survey, information is collected daily on the number of people partaking in the meal, by gender and age groups. The variability of DEC conditional on income is much lower once the DEC is adjusted for partakers, and the difference between household size and the number of partakers increases monotonically with income (Figure 5). A clear upward trend in the difference between per capita and partakers-adjusted mean DEC can be observed between the bottom and the top decile.

Figure 5. Differences in household size and mean DEC per capita when adjusting for partakers



Source: Grünberger (2017b)

[18] If meals consumed in another household have a corresponding entry as meals given to another household.

[19] Surveys analyzed by the FAO food security analysis team from 2006 to 2014, using the ADePT-FSM software developed jointly by FAO and the World Bank (Moltedo et al., 2014).

Analyzing five surveys that collect information on partakers, researchers at FAO's Statistics Division also found that the CV of DEC is systematically lower when household size is adjusted for partakers, even though the five surveys use different approaches to collect data on partakers (Table 3). In the 2010 Bangladesh HIES, the respondent was asked to report daily on the number of people present in the household as well as their demographic characteristics. In the 2007 SES of Mongolia, information on the number of visitors and the number of days they stayed in the household was collected. In the 2007 Afghanistan NRVA and 2010 survey of Niger, respondents were asked to report on number of meals and number of days that visitors stayed in the house. And finally in the 2010/11 Uganda NHS, information was collected on the number of people present in the household during the reference period. The reference period is sometimes different than that of the food module (Mongolia 2007/08).

Despite the difference in approaches, in all five countries the CV of DEC is systematically lower when household size is adjusted for partakers and the method designed by FAO to correct for excess variability fails to correct for excess variability due to omission of partakers[20] In terms of overall impact on the estimate of the prevalence of undernourishment, the effect of not-corrected partakers may lead to an over- or under-estimation as higher per capita DEC may counterbalance effects of the higher variability.

Table 3. Comparisons of the CV of DEC when adjusting for partakers

	CV of DEC per capita without correcting for excess variability		**CV of DEC per capita after correcting for excess variability**	
	Not accounting for partakers	Accounting for partakers	Not accounting for partakers	Accounting for partakers
Afghanistan 2007/08 NRVA	0.36	0.35	0.25	0.24
Bangladesh 2010 – HIES	0.32	0.24	0.26	0.22
Mongolia 2007/08 HSES	0.48	0.46	0.32	0.30
Niger 2011 ECVMA	0.63	0.58	0.35	0.34
Uganda 2010/11 - UNPS	0.63	0.61	0.38	0.35

One must be careful not to conclude from this evidence that meal participation is not an important issue; indeed, it still needs to be addressed[21]. The HCES that currently attempt to make these adjustments are few and highly diverse. In several countries, the questionnaires appear to capture only a portion of the requisite information and the results are likely subject to considerable measurement error.

[20] The FAO methodology should be able to correct for the excess variability due to the non-adjustment for partakers: for two (Bangladesh and Uganda) of the five countries analyzed the CV corrected based on DEC using household size and the CV corrected based on DEC using partakers were found to be different.
[21] This discussion is based on Fiedler and Mwangi, 2016.

Furthermore, there are several reasons why we would expect the importance of, and need for, these adjustments to be increasing; foremost the secular, seemingly universal, trend of the growing practice of consuming FAFH. It is noteworthy that, by implication, these studies may not provide an accurate portrayal of the actual situation in several of the studied countries: i.e., the findings may be false negatives regarding the importance of making adjustments for meal participation. This is especially likely to be the case in countries where there is greater travel away from home and where, more generally, there is more widespread practice of eating away from home. It is therefore difficult to make any definitive assessments about the value of making adjustments for meals, or about the feasibility or best practices of collecting the requisite information to make the adjustments.

A review of select questionnaires of more recent HCES revealed that over the course of the past few years a number of countries—India, Malawi, Mali, Niger, Nigeria, Uganda and Tanzania, among others—have introduced new questions to identify who among the household's members were meal partakers and which meals were eaten away from home, and by whom. Several of the surveys have also asked about meals that households provide to non-household members. These modifications appear to have been motivated by concerns that food away from home is increasing in frequency, is significantly under-reported, and is distorting the precision of food security analyses.

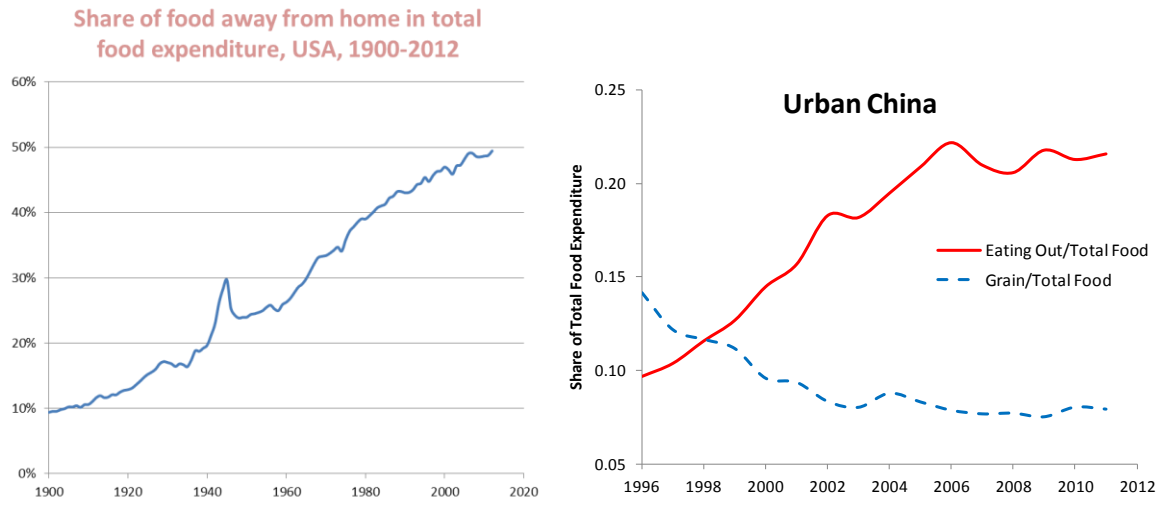
2.5. Food away from home

Consumption patterns are rapidly changing across the developing world, with prepared and packaged meals and meals consumed outside the home taking an ever-growing share of the households' food budget. With rising incomes, urbanization, women entering the labor force, and children eating at schools, among various reasons, this trend is expected to persist as economies transition to middle-income status (Maxwell and Slater, 2003; FAO, 2006; Popkin, 2008; Smith, 2013; USDA, 2011). As FAFH gains importance, failing to appropriately measure this component of food consumption and expenditure will make overtime comparisons of consumption patterns and poverty less and less meaningful.

In the US, the share of FAFH in total food expenditure increased from 10 to 50 percent during the 20th century. In urban China total expenditure on food away from home increased by 63 percent between 1995 and 2001 (Ma et al. 2006; see

Figure 6). Household per capita expenditure on FAFH rose at an average annual rate of 9.5 percent in China from 2002 to 2011, while the share of FAFH in total food expenditures increased from 18.2 percent to 21.5 percent (You, 2014). In both India and Egypt the prevalence of meals eaten away from home about doubled in less than 20 years.

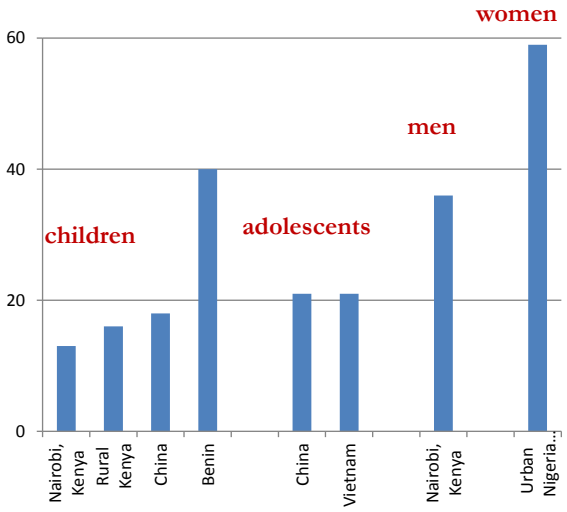
Figure 6 - The rapid rise of FAFH in the US and China



Source: USDA-ERS (left panel); Gibson, 2016 (right panel).

Taking into account FAFH consumption is particularly important for measuring calorie consumption (Figure 7), since food consumed outside the home tends to be more calorie-dense than food consumed at home (Poti and Popkin 2011; Mancino, Todd and Lin 2009) and the amount of food consumed away tends to increase faster with increases in income. FAFH has been found to contribute to as much as 36 percent of the daily energy intake among men in urban Kenya, and 59 percent among market women in urban Nigeria (Oguntona and Tella 1999; van't Riet et al. 2003). Among the younger population, FAFH contributes, for example, to 18 percent and 40 percent of daily energy intake among Chinese children and school-going adolescents in Benin, respectively (Liu et al., 2015; Nago et al., 2010).

Figure 7 - Energy intake from FAFH (%)



Source: R. Vakis, Improving measurement of Food Away from Home (FAFH) – presentation

Most nationally representative household surveys have not kept up with the pace of change in food pathways and collect very limited information on food away from home (FAFH). Smith et al (2014), when assessing the relevance and reliability in their sample of 100 surveys found that 90 percent of the surveys do consider FAFH in some form but that most of the approaches are "ad hoc and unsatisfactory". For example, a quarter of the surveys aim to capture all related household consumption from FAFH using just one question; one in five surveys considers multiple places of consumption; only 35 percent take snacks explicitly into account (when most snacking is expected to take place out of the home); and close to half of the surveys do not include FAFH received in kind.

While it is widely recognized that FAFH is subject to considerable measurement error, just how much it contributes to under-estimating consumption is unknown. However, since FAFH is expected to grow as a proportion of both total food consumed and total food expenditures, absent change in how information is collected, the magnitude of that under-estimation can be expected to increase. As it does, it will exacerbate the instability of HCES-based estimates of food insecurity and under-nutrition as currently measured, (Tandon & Landes 2012; D'Souza & Tandon 2014; Smith 2015), obfuscate trends and prompt more to question whether even the general order of magnitude of our estimates of global under-nutrition should be accepted (Banerjee & Duflo, 2011). The inadequate collection of FAFH data urgently needs to be better understood and systematically improved.

Few studies analyze the implications that failing to account for FAFH can have on food security analysis.[22,23] In a study from India, Smith (2013) argues that the great *Indian calorie* debate, originated by an apparent increase in undernourishment at the time of falling poverty rates, can be partly explained by inaccurate data on calorie intake due to the lack of measurement of FAFH. Similarly, Borlizzi et al. (2017) in Brazil show how the distribution of food consumption by income strata changes once food consumed at school is taken into account. In particular, they show that proper account for food received through a school feeding program targeted at the poorer strata of the population results in a more equal distribution of food consumption than previously thought. Capturing FAFH increases mean DEC, as it is an important food source, especially in urban areas. Smith (2015) shows that FAFH is positively correlated with the estimated mean dietary energy consumption. In many HCES, FAFH is only measured in terms of monetary value. But since meals eaten outside home are different than meals at home (Rimmer, 2001) the conversion of monetary

[22] With obesity increasingly becoming a pressing health issue in some middle-income countries, the link between eating out and obesity is also drawing attention in the developing world (Bezerra and Sichieri, 2009).

[23] The literature on FAFH in the developed world has a longer history, where a main focus has been on health and nutrition issues. There is widespread interest in studying the differences in the caloric and nutritional composition of the food provided by commercial outlets relative to home-made food, with the objective of understanding the health consequences of eating out (Vandevijvere et al., 2009). In particular, high calorie concentration found in certain meals raised particular concern, giving rise to a body of research devoted to understanding the link between obesity and eating out, among other health outcomes (Burns et al., 2002; Guthrie et al., 2002; Kant and Graubard, 2004; Binkley et al., 2004). There is also interest in establishing food-based dietary guidelines to prevent obesity and related chronic diseases developed later in life (Phillips et al. 2013).

value into calories can be misleading if home food consumption is used as a benchmark to calculate calories from FAFH.

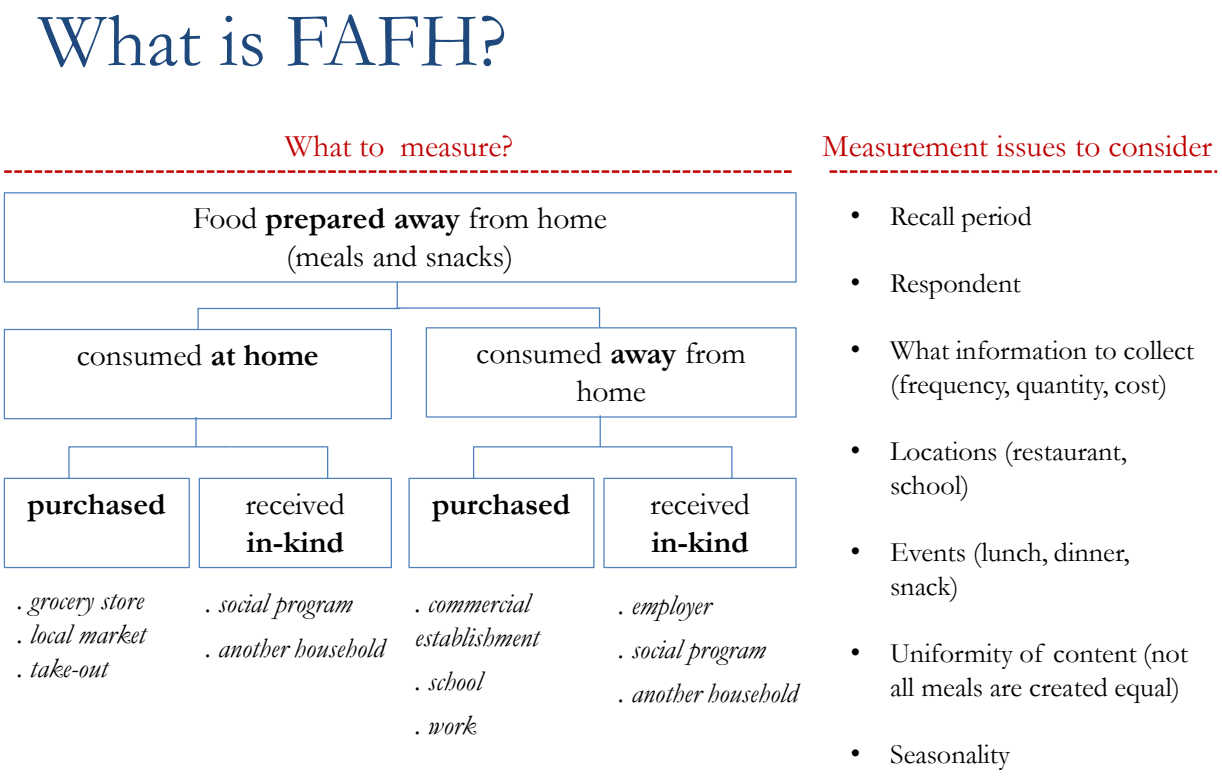
Using data for Peru, Farfan et al. (2017) evaluated the impact of accounting for FAFH on poverty and consumption inequality estimates. They show that from a theoretical point of view the direction of the effect on poverty or inequality cannot be predicted ex-ante. Empirically, they demonstrate that failure to adequately capture FAFH may generate serious biases in estimates of households' expenditure patterns and welfare measures and may change the underlying profile of the poor.

There is considerable evidence about a large variety of socio-economic factors associated with eating away from home. Measurement error from neglecting FAFH may therefore bias welfare analyses along these lines. In several countries, geography, household size and composition have been shown to be systematically related to the incidence and level of FAFH consumption and expenditure (Meenakshi & Ray, 1999; Mihalopoulos & Demoussis 2001; Yen & Jones, 1997; Mutlu & Gracia 2006; Meng et al., 2012). Households composed mainly by elder people have also been found to have lower probabilities of relying on, and lower expenditures on, FAFH (Redman 1980; Meng et al., 2012; Liu et al., 2015).

Conceptual and practical challenges make integrating FAFH in household surveys a complex exercise. First of all, a clear definition of what is meant by FAFH is needed.

Figure 8 outlines a useful way to conceptualize and measure FAFH, as well as some key measurement issues to consider. FAFH can refer to food *produced* outside, regardless of whether the food is consumed outside or inside the home. In this case, take-out meals would be considered FAFH. Alternatively, it can refer to food *consumed* outside irrespective of the origin of the food. Under this scenario, home-made meals consumed at work or school would be a component of FAFH. While there is a general preference towards defining FAFH based on the place of preparation of the food, a clear protocol that takes into account all these different pieces is required to be well defined regardless of the concept that is adopted. A second element to consider when collecting information on FAFH is snacks, which in modern eating habits are more likely to be consumed outside the home. Finally, there can be different modes of acquisition of the food, including purchased food or food received in kind, each of which can further originate from multiple sources such as commercial establishments, social programs, and other households, among others. While a great deal of attention has been paid to food that household member purchases and, to a lesser extent, food (or meals) that household members receives free as part of a social intervention (most commonly a school meal), there is evidence from China and India that "hosted" meals provided free to friends or relatives, are also an important, distinct category (Bai et al., 2010; Fiedler 2015). In China, "hosted" meals were found to account for nearly half of all FAFH and to be disproportionately important for lower income groups. In India, they accounted for 29 percent of all meals away from home, and 36 percent of all persons with at least one meal away from home reported having at least one hosted meal provided by another household.

Figure 8 - Defining Food Away From Home



Source: adapted from Smith and Frankenberger 2012

Of the issues covered in these guidelines, this is probably the area where it is hardest to trace one set of agreed upon international practices. The discussion that follows is poised on a FAFH definition based on where the food is prepared. Including food prepared at home but eaten outside would most likely result in double counting as the ingredients would already be accounted with the food available in the household from different sources. In addition, while for food-at-home the main food preparer is likely to be adequately informed about the food consumed by all household members, no one individual will be in such privileged position to report about the food consumption patterns of other household members away from home. FAFH may therefore need to be captured at the individual level, interviewing different respondents when possible.[24]

[24] In a small-scale study in an urban slum in India Sujatha et al. (1997) interview husbands and wives about the men's dietary intake, and find that women are not aware of the foods consumed by their spouses outside of their home. Similarly, Gewa et al. (2007) find that mothers of rural school-aged Kenyan children missed 77 percent and 41 percent of the energy intake originated in FAFH in the food shortage and harvest seasons respectively (where FAFH contributes to 13 percent and 19 percent of daily energy intake in each season). Collecting data from children is particularly challenging and not commonly done in household surveys in low- and middle-income countries. The report recommends a proxy respondent for children mainly because this is a widely accepted approach in practice and there

An additional question relates to what information to collect from respondents. For nutrition and food security analysis it will be important to have information (from within the survey or from other sources that can be integrated in the survey) on what is eaten, which is a challenge for FAFH as meals content is likely unknown to the consumer[25]. Options include: differentiating between meal, drink and snacks; asking by eating occasion (breakfast, midday snack, lunch, afternoon snack, dinner); reporting source of preparation (commercial, government or social program, employer, other household); differentiating type of establishment (fine dining, fast-food restaurant, street vendor); reporting day of the week (or weekday / weekend). Note that snacks can be as if not more important than some meals in terms of both calories and expenditures, and it is therefore very important that they are adequately captured in the data collection instrument.

Recognizing these challenges, several low- and middle-income countries have in recent years adopted innovative approaches to FAFH data collection. Details of experiences from India, Peru and the West Africa Economic and Monetary Union (UEMOA) are detailed in the online Annex 2[26]. In the US, the recent FoodAPS-1 survey, collects data at the individual level via food books, and attention to FAFH is ensured by including a reminder in the food books to record meals, snacks and drinks consumed at a number of different outlets (school, work, relatives, recreational sites etc.). Kirlin and Denbaly (2017) provide a detailed account of data collection in the context of this survey.

With all their differences, these approaches to collecting FAFH in recent HCES in such different settings have a number of aspects in common that can be useful in developing international standard for data collection directed at low-income countries: all surveys collect data at the individual level, all surveys differentiate meal types, and make explicit reference to snacks.

2.6. List of food items

The design of the list of food items included in a survey is known to affect how food consumption is reported by respondents (Deaton and Grosh, 2000; Gibson, 2005). Since a list with more items is likely to aid respondents in more accurately remembering their consumption or acquisition events, a longer and more disaggregated food list results in higher reported consumption (Beegle et al., 2012; Carrol et al., 2015). However, very long food lists can quickly lead to greater respondent and

is no established viable alternative implemented at scale. This is a topic however that warrants further research given the growing importance of food away from home, and particularly publicly-financed feeding programs, on children's diets.

[25] The caloric and nutritional content of meals consumed away from home can be sourced via complementary data sources, either purposely run surveys (as the survey of food establishments undertaken in Lima and describe in Farfan et al., 2017) or via administrative data (e.g. in the case of school meals).

[26] Online Annex 2 can be downloaded on one drive at: https://unfao-my.sharepoint.com/personal/nathalie_troubat_fao_org/_layouts/15/guestaccess.aspx?docid=1b9bc72a1073846588e90481b32a1f4dc&authkey=AdEoHB7wpnzz3Hg6dGClhlw&e=170c063d338e4c50ad979c2b8cc4803b

interviewer fatigue (Gibson, 2005). The choice of items to be included needs also to be country specific, to reflect differences in diets. An 'optimal' food list length should balance the lower memory lapses and the lower costs and interview time associated with short lists, with the better recall and more comprehensive reporting associated with the longer list.

Interview-based surveys, in which food items are pre-defined and listed, should include a sufficient number of food items to help respondents accurately remember what has been acquired and/or consumed. The list should fit all households in a given population, from poorest rural to richest urban households consuming a very wide variety of foods. Diary-based surveys can use a pre-defined list or be open-ended. The sample of surveys analyzed in Smith et al. (2014)[27] reveals substantial variability in the number of food items in HCES globally. The mean number of food items in diary-based surveys in their review is far higher than in interview-based surveys (229 and 102, respectively). The number of food items spans greatly across surveys, from a low of 19 to a high of 677.

When the number of food items recorded by respondents is not pre-determined, the aggregated list of items is usually far longer than the pre-defined list, which in turn impact the value of reported consumption. Tucker and Bennett (1988) report that diaries of an experimental US survey with preprinted items lead to higher total expenditure estimates than blank fill-in diaries, especially for older persons (see also Tucker, 1992).

Several studies in low- and middle-income countries indicate that longer disaggregated lists result in higher food consumption reporting than shorter aggregated lists. In the experimental study of Beegle et al. (2012), 7-day recall interviews with lists of 58, 17 and 11 food items were tested and the list with 58 items returned the highest consumption estimate. Joliffe (2001) reports that for a survey in El Salvador, a longer food item list (72 items) resulted in 20 percent higher food consumption compared to questionnaires with only 18 aggregated food categories. Food data collected in a Jamaican experimental survey shows 26 percent higher consumption with 119 items compared to 37 items (Statistical Institute and Planning Institute of Jamaica, 1996).

In Indonesia's SUSENAS survey, questionnaires with extensive item lists (218 items) led to about 7 percent higher food consumption than questionnaires with only 15 aggregated food categories (Pradhan, 2009) with underestimation increasing with consumption levels, as the variety of diets is positively correlated with consumption. The longer list required a much smaller than proportionate increase in time relative to the increase in the number of items, although published figures from the study combined the food and non-food consumption lists (52 minutes for 23 items and 82 minutes for 320 items).

Similarly in Tanzania (Beegle et al 2012), reducing the list length by as much as 80 percent resulted in reducing interview times by only 17 percent (49 and 41 minutes for the 58- and 17-item lists, respectively). Additionally, Bradburn (2010) has noted that grouping questions (and food types) can

[27] The numbers reported herein exclude the Brazil diary survey, which is a significant outlier at 5,407 items. Many of these are however simply similar items named or spelled differently.

help to minimize the cognitive effort for respondents to recall the requested information, thus leading to lower recall error. This implies, for example, that FAFH questions should be reported in a separate group.

Friedman et al. (2017) decompose the error response in food consumption measurement into two different components: the omission of any consumption item and the error in reporting the value. The shorter food list has lower consumption incidence as compared to a diary survey benchmark for the large majority of food groups, while for consumption value (conditional on positive consumption) the subset list module is the module that most over-reports consumption in comparison to the diary benchmark.

Analyses of the Food Frequency Questionnaires (FFQs) often used by nutritionists also support the use of longer lists. Wakai (2009) finds that long FFQs (97+ foods) display higher correlation with weighed food records (r=0.42 to 0.52), than short FFQs (< 70 food items, r=0.31 to 0.45). Similarly, Henriquez-Sanchez (2009) finds that FFQs with >100 food items correlated more strongly with weighed food records (r=0.52) than FFQs with less than 100 items (r=0.47). A longer, more qualitatively differentiated food list is preferable.

In sum, given that the incremental interview time required for additional items is relatively small, having a relatively long list is recommended. On the other hand, one should not underplay the possibility that a longer interview time may lead enumerators to take shortcuts in interviewing (Finn and Ranchhod, 2015) and respondents to refuse to participate or terminate the interview ahead of required time (Deaton and Grosh, 2000). This is particularly the case when a questionnaire has a cascading structure, as respondents are more likely to not report some expenditure to skip other following questions (Kreuter et al., 2011). In what follows some criteria that can help practitioners inform this complex balancing act are discussed.

One concern in drawing a food list is that for many users the food list will have to conform with standardized classification systems. The UN system for Classification of Individual Consumption According to Purpose (COICOP[28]) has become the basis of classification for compiling CPIs in a comparable way, in line with the requirement of System of National Accounts (SNA) 2008. This is for all practical purposes the standard for surveys designed to inform CPI compilation, and also the basis for the System of National Accounts. Another source for standard classification is FoodEx2, a comprehensive food classification carried out by the European Food Safety Authority and recently updated to cover global needs (European Food Safety Authority, 2014). FoodEx2 is an additional reference classification of particular relevance when it is expected that a survey will be used to conduct nutrition analysis.

An important external reference to consider when drawing a survey food list is food composition tables. Included items in the food list must match unambiguously to an entry in the food composition

[28] An updated version of the COICOP will be released in 2018. The new COICOP classification better reflects the food products consumed in low and middle income countries.

table (FCT) in order to convert food quantities into nutrient quantities (FAO/INFOODS, 2012) [29]. This is extremely important if the data are to be used for nutrition and food security analysis, as calorie and nutrient composition can only be derived if foods can be matched at the analysis stage to a corresponding relevant FCT. The conversion table should ideally be finalized before the survey goes to the field to ensure that with the breakdown of food items is compatible and that the availability of nutrient content data can inform the design of the survey food list[30].

The length and composition of the food list should be formulated bearing in mind how data are supposed to be used. From a welfare perspective, it is crucial that those items representing the large majority of food expenditures are included. From a nutrition perspective, those foods which are important sources of nutrients in individual diets must be included, while items that contribute little to the understanding of individual nutrient intake are less important. As a result, welfare and nutrition requirements do not necessarily correspond. Thus, choices regarding the food list tend to be "topic oriented", even if surveys should be designed for a wide range of users. A nutrition-oriented food list is likely to include more items than a welfare-oriented list. One common solution in HCES is to list the most common food items consumed by the population, and include "other, specify" items in each category where the acquisition/consumption of additional food items can be recorded. But this entails additional challenges if the intention is to estimate nutrient contents, as the matching with a FCT becomes uncertain[31].

Finally, when the objective is to collect data in order to evaluate the impact of a nutrient fortification programs, the food list should include all foods directly fortified with such nutrients and their products. For example, if we are interested in assessing the nutritional impact of fortified wheat flour, the list of foods should include wheat flour fortified and the products made with such type of flour (e.g., bread, biscuits pies).

As diets evolve, food lists must be regularly updated to reflect dietary changes. This is particularly relevant in urban areas where a wider variety of foods are typically eaten, and processed foods[32] and prepared foods form a larger share of diets (Popkin, Adair and Ng, 2012) and budgets. Again, a nutrition perspective entails a higher specificity of food list, including and distinguishing different levels of food processing, from minimally processed foods such as yogurt and cheese, through bread and frozen vegetables, to highly-processed and ultra-processed foods rich in sugar, fat and salt (that have been shown to be associated with obesity and other diet related diseases)[33]. Using data from

[29] It is also recommended that countries invest in the development of good reference food composition tables and in keeping them up to date.

[30] If a FCT is available, it can also be used to pre-program Computer Assisted Personal Interviewing software to perform built-in checks for excessive consumption and speed up data analysis and cleaning.

[31] See Fiedler and Mwangi (2016) for a discussion on this issue.

[32] Processed food is defined as "Any food other than a raw agricultural commodity and includes any raw agricultural commodity that has been subject to processing, such as canning, cooking, freezing, dehydration, or milling" (USDA 1946).

[33] The Brazilian dietary guidelines include a classification of foods by level of processing. See, for example, the "Dietary Guidelines for the Brazilian Population" (Ministry of Health of Brazil, 2014) http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/brazil/en/).

Brazil, for instance, a recent paper by Louzada et al. (2017) concludes that HCES hold potential in reporting consumption of ultra-processed foods, as there is substantial convergence between the data collected in an individual dietary intake survey and a HCES data in terms of relative energy consumption from ultra-processed foods.

2.7. Non-standard units of measurement

To use food quantities collected through HCES in a meaningful way, these quantities need to be standardized: this allows for the aggregation and comparison of consumption across food items and geographical areas, or for translation into nutrient content[34]. Many surveys fail in the standardization of units, especially in Sub-Saharan Africa where non-standard units are commonly used in daily life (Deaton and Dupriez, 2011).

Smith and Subandoro (2007) identified seven primary methods for collecting information about consumption quantities, and advocate using a combination of these methods, since one method may be the optimal solution for certain items, but it may not be appropriate for others.

1. Metric (i.e. standard) units: Respondents report quantities in metric units such as kilograms or liters.
2. Monetary value: Respondents report the monetary value of the quantity consumed.
3. Local (i.e. non-standard) units such as piles, baskets, or bunches.
4. A count of pieces: a type of non-standard unit.
5. Volumetric equivalents: respondents demonstrate how much space the food they consumed would take up.[35]
6. Linear dimensions: respondents provide linear measurements (length and width or circumference) for the amount of food consumed. As Smith and Subandoro (2007) point out, this method likely takes more time to complete as it requires physical measurement rather than a simple vocal response.
7. Food models: respondents choose a two- or three-dimensional depiction of a food item that best corresponds to their consumption. Three-dimensional models can provide very accurate estimates, but it can be costly to prepare the models and calculate their weights.

[34] Estimating the quantities of foods consumed *away* from home poses its own set of issues, mainly concerning the paucity of data broken down by individual food items. Anyway, quantities can be estimated if respondents report on the foods and dishes that were consumed rather than only their total expenditures (see Smith and Subandoro (2007) for the detailed methodology to be used when this information is available).

[35] An advantage of bounded recall is that the initial visit to begin the recall period also allows survey teams to distributed standardized volumetric containers, such as an empty sack. This can be especially helpful where bulky root crops or plantains are dietary staples because a rural household might fill a sack several times over in the course of a week, with root crop consumption of 50 kg or more. The PNG survey used by Gibson (2001) distributed empty 25 kg sacks and these were the preferred NSU for all root crops and vegetables, with local weighing trials for converting sacks (and partial sacks) into kilograms.

For all but the first method, additional data are required to convert the reported information into standardized, comparable (metric) quantities. Collecting quantities in non-standard units (NSUs) or restricting respondents to only reporting in standard units involves trade-offs in accuracy and feasibility. In the sample of HCES analyzed by Smith et al. (2014), the most common method employed is requiring respondents to report in a metric unit of measure. This method is usually the easiest to administer with the lowest budget and time cost; it also requires the least amount of post-data processing as the units are already comparable across items.

However, restricting respondents to using only standard units may result in inaccurate estimates, especially when respondents are not accustomed to use them for all of the items they consume. Reporting quantities in non-standard units (NSUs) is quite common for a wide range of commodities, especially in SSA. For example, in the second wave of the Ethiopia Socioeconomic Survey from 2013/2014 where NSUs were allowed, nearly 50 percent of farmers chose to report their harvests in NSUs. In the Malawi National Panel Survey, respondents chose NSUs about 73 percent of the time. This provides a strong indication that many respondents are more comfortable reporting quantities in NSUs. So, allowing for their use eases the burden on these respondents in terms of memory recall and conversion calculations reducing the accuracy of the resulting data.

Recent studies show that asking respondents to combine memory recall with cognitive tasks, such as abstracting consumption to a “typical week or month”, leads to less accurate self-reporting (Beegle et al, 2010). The forced conversion from non-standard to standard unit similarly combines cognitive and memory recall: respondents must (1) have a clear understanding of what a standard unit of the food item is (e.g. how much is a kilogram of rice), (2) estimate how many standard units correspond to the NSU they know, and finally (3) using the conversion from step 2, convert the quantity consumed into standard units. All three stages place a cognitive burden on the respondent and can lead to sizable measurement error. It is also common practice for such calculations to be conducted in-situ, often on-the-fly (as the respondent makes the calculations in their head, perhaps prompted or assisted by an interviewer) further increasing the likelihood for error.

Thus, the most important and overriding benefit of allowing respondents to report quantities using NSUs is that they will likely give a better report of the consumed quantities. Results from a recent methodological study of land area measurement further support the preference for NSUs. Carletto et al. (2016) found that when respondents are allowed to report land area in non-standard units – instead of being forced to convert area to standard units – the self-reported estimates are much more accurate. The additional costs and challenges associated with this method broadly fall into two categories: (1) those associated with preparation and implementation of a survey with NSUs, and (2) ensuring NSU measurements can be converted into comparable standard units.

Before NSUs can be used, country-specific information on common NSUs is needed to determine the list of allowable item-unit combinations. NSUs must also be accompanied by standard unit conversion factors. In order to directly compare and aggregate quantities, the data user must convert all quantities into a common standard unit such as kilograms. Capéau (1995) and Capéau and Dercon (2006) suggested using econometric techniques to compare unit values and estimate conversion

factors, which are relatively simple to implement. The challenge with this approach is that unit values can vary as a result of quality differences (Deaton 1997), or from price discounts on larger units (Attanasio and Frayne, 2006). Moreover, unit values vary over time and space due to the impact of transport and storage costs which alter the relative price of quality, and therefore will alter the composition of local demand, so the unit value is not a consistent measure, unlike for market prices (Gibson and Kim, 2015). These sources of variability in unit values unrelated to mass or volume can result in distorted or imprecisely estimated conversion factors. (Oseni et al., 2017).

An alternative method is to collect weights for each allowable item-unit combination and use these to create conversion factors. This is a relatively straightforward concept, but proper implementation entails its own set of challenges. The standard weight for the same item-unit combination can vary, even within a country. For example, Casley and Kumar (1988) found that in Nigeria an average "bundle" of sorghum weight between 26 and 49 kilograms depending on the area. Thus, region-specific conversion factors need to be considered as well. Complicating matters further, different levels of processing (fresh, dried, powered) lead to different conversion factors for the same food item[36]. This method has also been recently implemented in the context of a project of survey harmonization in the countries belonging to the West African economic and Monetary Union (UEMOA) following the recent guidelines published by the World Bank (Oseni et al., 2017; See box for details).

Box 4. Collection of conversion factors in context of the UEMOA project

As part of the Regional Program to Harmonize and Modernize Living Conditions Surveys in the West African Monetary Union, participating National Statistical Offices (NSO) will collect consumption data that allows reporting in non-standard units of measurement (NSUs). To prepare for this, market surveys were conducted to establish conversion factors from non-standard (bowl, heaps, etc.) to standard units (kg, liter) for commonly consumed food items. Following guidance from Oseni *et al* (2017), NSOs implemented market surveys ahead of the main household surveys. In the spirit of harmonization, all countries involved in the program followed the same procedure.

(1) *Preparation phase*: included a review of previous surveys and preliminary markets visits to identify products and commonly-used NSUs. The resulting product-unit list – and options to report additional found units – were programmed into Survey Solutions, a Computer Assisted Personal Interview (CAPI) software for data collection.
(2) *Field work:* was conducted in the post-harvest season to ensure wider availability of products. Enumerators visited 6 markets (3 rural + 3 urban) in each region, using tablets to collect measurements from up to three different vendors for each product-unit combination. They also took photographs of each measurement which, thanks to the CAPI application, were automatically linked to the measurement data.
(3) *Data analysis:* includes:
    - Data cleaning to detect outliers and re-arrange unit sizes according to the actual weights;

[36] This is similar to food crops harvested under different conditions – threshed, shelled, fresh, dried, etc. – which are proven to have a large impact on reported harvest quantities (Fermont and Benson, 2011; Diskin 1999; Murphy et al, 1991).

- Calculation of national- and regional-level conversion factors using the median measured weight for each product-unit combination; and
- Identification of the best photographs with actual weights closest to each conversion factor.

The resulting library of NSU materials will support the main household survey. Enumerators will use the photo albums to guide respondents in reporting food consumption quantities and the conversion factors will be used to flag unreasonable quantities for further verification, all ensuring greater data accuracy.

Unit conversion factors are often incomplete and lacking documentation, which decreases the number of usable observations and makes it difficult to cross-reference quantities or apply conversions across different datasets. Smith et al (2014), pointed out that calculating metric food quantities is feasible for only 53 percent of the surveys reviewed. Most of the difficulties associated with this method can be addressed by ensuring that unit conversion factor data are thoroughly collected and properly documented. When information on most common NSUs used and unit conversion factors is limited or not available, the survey team must take on the effort of collecting these data. This is most effectively done by consulting with local experts and conducting a market survey prior to the start of the regular data collection.

3. Conclusions and recommendations

This section represents the core of these guidelines, providing a set of recommended practices for data collection. The recommendations are based on the literature, empirical evidence and considerations discussed in Section 2. The objective is to promote the adoption of good practices, and encourage the abandonment of some bad practices that are still employed in some surveys. Some of these recommendations are straight-forward and easy to implement and some are grounded in firm empirical evidence, but some are based on balancing incomplete pieces of evidence with practical considerations. Additional research will be useful to reinforce the evidence base behind the entire set of recommendations, but surveys are being designed and implemented, and advice is being given: A set of guidelines can therefore be extremely useful in informing design decisions and fostering cross-country comparability in approaches.

Such guidelines can only be seen as temporary as best practice will evolve and depend on circumstances. As food consumption patterns evolve, statistical systems change and new technologies become available, survey design will have to adapt to stay relevant and cost- effective. It is anticipated that as more survey methodological work is performed and new lessons are learnt from survey implementation, these guidelines will have to be periodically revised to incorporate the new knowledge being generated and to respond to additional or different data needs that may emerge. Furthermore, consumption patterns change with income growth but also with changes in food technology and the modernization of food systems, response rates tend to decline in higher income economies, and technology is already proceeding at a fast pace with new technologies

relevant to survey operations entering the market every day. In such a rapidly evolving environment, the shelf-life of these guidelines is inevitably going to be limited, and it would be desirable for them to be revised and updated at least every 10 years.

The overarching concern in drawing these recommendations has been to ensure the most appropriate balance between accuracy and cost effectiveness, having in mind the typical constraints facing low- and middle-income country statistical offices. In some cases, the recommendations offered might entail a costlier option than what is currently practiced, in some cases it might imply a cost saving. The benefits should be assessed not only in terms of greater accuracy and comparability of the data being collected, but also with respect to the fact that if data can be made relevant for a wider user base (e.g. nutritionists and food security analysts, as well as statisticians and economists) their usefulness is enhanced accordingly.

While professionals from these four fields contributed to the preparation of these guidelines, future revisions would benefit from the involvement of an even broader range of scholars and experts. Food is part of our cultures and societies, and the way in which society (and survey respondents) relate to food is mediated by social and cultural constructs. Involving anthropologists may help devising questionnaires that are better able to incorporate social and cultural aspects in data collection processes and outcomes. Similarly, psychologists can help designing more effective survey approaches by providing insights on the cognitive process behind how respondents answer questions, when and how they shift from enumeration to estimation strategies, and how survey design should take that into account when thinking about details such the recall period, the length of a list, and the sequencing of survey modules.

One notable gap in these guidelines is the discussion of price data collection. Prices are clearly an important element of any analysis of poverty, and essential for welfare comparisons across households, regions and time. Price data collection is also a major goal of HCES when they are required to inform CPI calculations. But even just from a perspective of analyzing food consumption, food prices are needed to properly value consumption quantities, and having information on prices is invaluable in cross-checking the plausibility of reported values and quantities, especially when non-standard units are used. And as food policy often relies heavily on price interventions, measuring food prices is an important item for analysis in its own right. Price data collection was however left out of these guidelines firstly because it has implications well beyond food consumption, and secondly because being such an overarching topic it was felt it would probably be best served by its own set of guidelines. It would have been difficult to discuss and make recommendations about price data collection with reference to food alone, abstracting from all the other demands on price data from other uses and users.

It is also important to remind users that any change in method should have a controlled comparison to ‘bridge’ the effects of the old and new methodologies on the resulting data. Concerns about losing comparability over time, and the difficulty of explaining to the public the difference in estimates that come with changes in methodology often make statistical offices shy away from changes in survey design, even when it is clear that there would be gains in accuracy and cost savings in doing so. Building such controlled comparison in survey planning would allow easing those

concerns somewhat, and it is therefore something statistical offices in low- and middle-income countries, and the donors and international agencies assisting them, should consider whenever the implementation of a methodological change is being considered. One idea that was put forward as part of the international discussion leading to these guidelines was that of creating a global fund to support such controlled comparisons, as they are a source of invaluable learning and hence might be considered a global public good.

Box 5. A research agenda for food consumption data collection

The Guidelines put forward clear recommendations based on existing evidence and experience accumulated in survey practices over the last decades, but also recognize that in several areas there is little sound methodological work to base recommendations on. A list of priority areas based on the gaps that were identified during the preparation of the document is provided below.

Food Away from Home (FAFH) is an area where different approaches are emerging, but where little or experimental methodological studies are currently available. Given the increasing importance of this component in both calories and expenditures this is probably the area where methodological research would be the largest. Several National Statistical Offices have signaled their interest in participating in methodological studies that would help improve the quality of the data they collect on FAFH.

There are a number of different aspects of the choice between diary and recall, and the length of the recall period that would benefit from more methodological studies. Only a handful of experimental studies have been conducted on the topic in low- and middle-income countries and a larger evidence base would be required to make the extrapolation of results more robust. Specific questions in need of investigation in this domain are:

- *Bounding of recall*. One concern with short recall period (such as 7-days) is telescoping. 'Bounding' the recall period for a household with another visit to mark the beginning of the recall period could in principle help reduce telescoping and improve the quality of the recall. While this idea has been around for many years, it has not been formally tested and compared to unbounded recall in a low- and middle-income country setting.
- *Telephone interview aids*. Telephones are increasingly used in surveys also in low- and middle-income countries as the coverage of the mobile phone network increases. One way in which phones could help in person interviews is that of using follow-up phone calls to aid the filling of a diary, or the collection of a second set of recall data.
- *Multiple visits*. One issue with 7-day recall discussed in the Guidelines is that the data are affected by 'excess variability'. One way to reduce that (Gibson, 2016) is to perform a second non-consecutive visit to the same households (in person or possibly via telephone). This option is potentially attractive but has not been tested at scale.
- *Expand the evidence-base*. In general, many of the conclusion in the Guidelines come from small number of studies, with the SHWALITA dataset from Tanzania exerting probably an excessive influence on the current consensus simply because of its uniqueness. Replicating more studies, in different settings and regions with a similar set-up would help expand the knowledge base and provide more confidence when extrapolating results across countries.

There are other topics that might benefit from more research that the Guidelines have not touched upon. Some of these, that were also brought up during the global consultation, include the measurement of food waste, the measurement of individual consumption of specific population sub-groups, such as children and women of reproductive age, and the integration of different data sources. The IAEG calls on countries to consider setting up a global fund that could finance the implementation of methodological studies and experiments to test and validate survey design options in these domains.

Finally, as with all survey design choices, countries implementing these recommendations should carefully evaluate the extent to which their adoption might increase the burden on the respondent, how that risk can be mitigated, and whether the return in terms of data quality is large enough to justify a possible increase in non-response. As for the evaluation of survey costs, it is impossible to evaluate in principle these trade-offs with any level of accuracy and hence to be prescriptive about how to handle these survey design choices. When implementing the Guidelines in practice, however, care must be taken in finding the right balance between keeping the overall length of the survey manageable so as not to compromise, rather than improve, the quality of the information collected.

This section is developed as follows: for each domain discussed in Section 2, a summary of main findings is provided, followed by a set of recommendations.

3.1. Recall vs diary and length of reference period

Summary

The trade-offs between diary and recall and between shorter and longer recall periods have been highlighted in various survey experiments and analysis of diary/recall approaches. In low-income economies, evidence suggests that recall interviews are generally preferable to diary methods for capturing food consumption when balancing implementation costs and reliability of resulting estimates. The majority of the studies found that food consumption or monetary value data collected with recall interviews provides estimates that are similar or higher than those recorded in diaries. However, depending on implementation methods, diaries often show patterns of rapidly declining consumption (and data quality) over the reference period. Lower and decreasing consumption recorded in diaries is frequently attributed to respondents' fatigue and illiteracy in combination with poor supervision. Under close supervision, diaries have shown to be reliable in several contexts and are sometimes considered the "gold standard", but when implemented with appropriate levels of supervision (daily visits to households, etc.), they are generally costlier than recall surveys (the detailed cost calculations by Beegle et al, 2012, suggest that diaries are from 6-10 times more expensive than recall, once account is taken not only of fieldwork but also their time-consuming coding and data entry requirements).

Recall surveys suffer from memory decay (memory loss) as the recall period increases, and telescoping error (reporting of consumption outside of the recall period) for shorter periods. The experimental evidence suggests that a 7-day recall can perform as well as a diary in capturing food expenditures and their variability. Recall periods of longer than 14 days suffer from significant memory decay, while diary fatigue already appears to be significant after the first week. Regardless of how accurately they capture mean consumption, however, surveys based on short recall periods ('snapshots') will always over-estimate the variability in habitual consumption.

For individual food items, short recall periods (such as 7-day recall) are affected on the one hand by an underestimate of the incidence of consumption, particularly for infrequently consumed items; and on the other by an overestimate of the value of consumption (conditional on positive consumption) due to telescoping error. The recall error appears to be larger for less-frequently consumed items and on short recall periods. The "usual month" consumption was designed to deal with the conflict between a long reference period (to get 'typical' living standards) and a short recall period for feasible interviewing. But it does not work as expected as it results instead in an overestimate of the incidence of consumption, particularly for infrequently consumed items, and an underestimate of consumption values for staples. Importantly, the 'usual month' approach also imposes a significantly higher burden on the respondent and results in longer interviews.

Bounding the recall period with an earlier visit and asking household to recall their consumption since the last visit of the enumerator is a possible option for improving the quality of recall data. The evidence on the effectiveness of bounding is limited, and since this approach requires an additional visit to the household and it is therefore costlier the method is not recommended until more research is performed to evaluate its benefits. Another approach gaining ground recently is to complement 7-day consumption recall with data on the last purchase in the past 30 days, aimed to better capture the unit values of purchased food items. More methodological work is needed to assess the performance of these approaches.

Recommendations

The following recommendations are provided for the choice of methods in capturing food consumption in HCES surveys:

- Low-income countries are advised to adopt recall interviews and a 7-day recall period, as this method provides the best balance between accuracy and cost-effectiveness.
- Any survey using diary methods must be closely supervised to ensure proper and consistent completion, especially where illiteracy rates are high. The reference period should not exceed 14 days. Detailed metadata on how the diary was administered and supervised should be distributed with the primary survey data.
- The 'usual month' approach should be abandoned.
- Any change in the recall period or method (recall vs diary) should be accompanied by an experimental component aimed at assessing the impact of the change on survey estimates, thus enabling the reconciliation of estimates before and after the change in methodology. The studies by Backiny-Yetna et al. (2017) and Beegle et al. (2012) provide examples for how this can be done in practice.

3.2. Seasonality, number of visits

Summary

Food consumption and expenditure can show systematic variation related to time of the year, month, or week, as well as for agricultural seasons, holidays, and festivals. Such seasonal patterns

need to be considered in survey design and analysis, as they are important possible sources of bias and measurement error.

A survey that only captures food consumption or expenditure data in one period of the year will miss any seasonal variation and will likely not be representative of habitual consumption throughout the year. Many surveys currently collect data using one visit per household, concentrated over 3-4 months of fieldwork (48 percent of the surveys assessed by Smith at al., 2014). This approach cannot capture seasonality effects and is therefore not recommended. If it is adopted, to ensure that the timing of each round does not affect comparability in the estimates from one year to the other, it is critical to at least maintain consistency in the timing of the fieldwork. It should be noted however that even this provision may not be enough to ensure over-time comparability as seasonal weather patterns may change year-on-year, as do dates of some important festivities affecting consumption events (e.g. Ramadan). One possibility to capture seasonality might be to implement multiple visits on a sub-sample of households, but that is not an approach that has been tested widely and is therefore not offered as a recommendation.

The only way to accurately capture habitual consumption for each household is to survey them multiple times over the year, but this is also the most expensive option and in practice is difficult to implement. Data collection spread over the year, but with only one interview per household (and using a short recall period) will result in an accurate estimate of average consumption for the population, but in excess variability around the mean (Deaton and Zaidi, 2002; Deaton and Grosh, 2000).

Recommendations

When interest is in analyzing the habitual consumption of a population through an extended period of time (usually one year), it is recommended that seasonality is taken into account in survey design. The two options to consider, in order of preference are:

- Conduct one visit per household, spreading the sample over 12 months of fieldwork. The overall sample should be stratified quarterly (e.g., split the overall sample into 12 monthly subsamples in a manner that allows them to be aggregated into quarterly, nationally representative subsamples).
- Conduct two visits per household, where the timing of the visits is scheduled to capture seasonal variations (for instance the first visit could be during the lean period and the second after the main harvest).

Countries should carefully consider using more than two visits, because of the higher cost and the difficulty to manage field teams that is associated with more visits. Implementing more than two visits is not impossible, but several instances where that has been attempted have run into implementation problems. Respondents' burden will also increase more than proportionally with each additional visit.

Regardless of the approach chosen, care should be exercised so to ensure enumeration is equally spread throughout the days of the week and the month and change in timing in holidays, festival and harvest needs to be considered.

3.3. Acquisition vs consumption

Summary

Surveys differ in whether and how they capture food consumption and acquisition. Typically, household budget surveys focus on collecting data to construct consumer price indices, therefore recording food items acquired through market purchases. However, as HCES are increasingly used for poverty and food security analysis, emphasis has shifted towards also collecting data on food items procured through own-production, barter, gifts, and payments in-kind, which are particularly common in rural areas. Information on own-production, barter, gifts, and payments in-kind is also important for national accounts statistics since food acquired through these channels are included in the household final consumption.

Conforti et al. (2017) classified food data collection practices in three approaches: 1) *Acquisition*. Households report food acquired through purchases, own production, and in-kind transfers, while consumption of food is not reported. 2) *Combination of acquisition and consumption*. Households report food acquired through purchases without specifying the amount consumed; and report food consumption from own-production or in-kind transfers. 3) *Consumption*. Households report food actually consumed, and indicate whether that same food was purchased, own-produced, or received as a transfer.

Irrespective of which of the three approaches is adopted by a survey, it is important that information is collected on the food that becomes available to the household through all the possible means of acquisition. It is also important that the survey objectives be clear on whether the information being collected is on acquisition, consumption or both. Regarding consumption, clarity is needed on whether surveys refer to food intended for consumption (i.e. including food waste) or food actually consumed (which excludes food waste). In the review by Smith et al. (2014) issues emerged in current practices on the neglect of food received in-kind, which were not collected in 14 percent of the surveys (4 percent did not collect information on own-produced food).

When combining information on sources of acquisition and consumption, care should be taken to ensure that question wording does not lead to incompleteness or ambiguity in enumeration. Smith et al. (2014) found that 38 percent of HCESs have issues with the wording of rule out questions, ambiguity on whether acquisition or consumption is being asked.

Recommendations

- All surveys should collect data on all main modes of food acquisition, namely:
    - Purchases.
    - Household's own production.
    - Received in-kind. Surveys need to prompt explicitly for in-kind sources that are otherwise likely to be missed, such as payments for labor, participation in social programs, and more. These in-kind sources can be aggregated, and care should be taken to avoid duplicating information captured in other sections of the survey (e.g. employment, social assistance). If public social assistance transfers are not captured elsewhere then it would be important to disaggregate.
- Surveys should be designed so that it is clear to respondents, enumerators, and data users exactly what information is requested and reported, whether it is acquisition, consumption, or both.
    - In the case of consumption, it should be clear whether the questions concern food intended for consumption (including food waste) or food actually consumed (net of food waste).
    - If total amount of food purchased over the recall period is the variable of interest, it is then recommended to add an additional question on the amount consumed out of these purchases to avoid mixing acquisitions from purchases with consumption from own production and in-kind.
- Surveys should exercise care to avoid possible sources of incomplete or ambiguous enumeration commonly found in current survey practice.
    - When using a filter question (30 percent of surveys assessed by Smith at al. have a leading/filter question):
        - Avoid leading/filter questions where respondents are asked first if they consumed a food over a certain recall period then details about consumption. A negative response to the first question will result in skipping questions on quantities acquired but not consumed during the recall period. This leads to systematic underestimation of the quantities and/or expenditures of food acquired.
        - Avoid filter questions focused on food purchases. This leads to underestimation of mean food acquisition for the population by failing to account for food acquired through own-production or in-kind transfers.
    - For consumption from own-production, the question must be worded to clearly indicate food *consumed* from own production rather than all food *harvested*. When this distinction is lost, the quantities and/or expenditures reported may include food entering the households' production stocks – not for immediate consumption – and thus systematically overestimate food consumed from home production.

3.4. Meal participation

Summary

Household surveys collect information on the total amount of food to be consumed by a household over a certain reference period. To convert this information to a per capita basis, and to perform analysis of the adequacy of food consumption and nutrient intake, it is important to know exactly how many people (partakers) consumed the total amount of food reported by the household. From the point of view of food and nutrition security analysis it is also useful to collect data on the physiological status of the partakers (e.g. pregnant or lactating status for women) as that affects nutrition requirements. Furthermore, food may be shared with non-household members, or household members might not be in the household when the food is consumed. Neglecting these occurrences adds measurement error to the distribution of per capita food consumption.

There are two main approaches to adjust household per capita consumption for the number of partakers. The first approach consists in counting the number of people who shared the household's meals and divide the total household consumption by this number. Such approach however is not very precise as it is not easy to account for situations in which people participate only in some meals per day, e.g. employees. The second approach consists in counting the number of meals taken by each household member and non-household members over the reference period for which food data is collected. Such approach is more precise but also more difficult to implement. There is very little methodological work that has formally tested the cost and benefits of adopting competing options for accounting for partakers, so this is an area to focus on future research. As a result, the recommendations provided below are also somewhat more generic that recommendations provided for other areas of survey design.

Recommendations

It is recommended that all HCES should consider adding an individual household member-based meal module. Since collecting information on individuals is expensive and difficult, this could be implemented as part of the module collecting information on food away from home (see below). On the other hand practitioners should realize that adding an individual household member-based meal module would enable eliminating other questions that are commonly used in surveys[37].

As a less preferred alternative, if an individual member-based meal module cannot be adopted, it is recommended that the following information is collected at the household level via proxy respondent:

- How many meals does [NAME] usually take in a day?
- How many days in the past X days was [NAME] present in the household?
- How many meals during the past X days, did [NAME] purchase or receive, and eat away from home?
- How many meals during the past X days, did [NAME] eat at home?

[37] This includes questions such as "How many meals are usually taken per day in your household?", "How many days in the past X days was [NAME] present in the household?", "Did [NAME] eat meals in this household in the last X days?", "Does [NAME] get meals at school?", "Did [NAME] consume any meals/snacks/drinks outside the household in the past X days?". The information collected in these questions would now be captured in an individual level module.

- Does [NAME] get meals at school?
- How many meals were served to non-household members during the last X days?
- Did the household host a ceremony, party or festival in the past X days, during which a large number of meals (not just snacks) were served to non-household members? If “yes”: How many attended?
- During the last X days, were there non-household members who stayed one or more nights in the household as a guest? If “yes”: How many nights did they stay? How many meals were they (summed together) served during their stay? How many of the guests were children <15 years old? How many adults 16 and older?
- During the last X days, were any meals served to non-household members? (Other than those served guests who stayed overnight.)

3.5. Food away from home

Summary

Rapid urbanization and economic growth are typically associated with an increase in the consumption of food away from home (FAFH) in absolute terms as well as a share of calories and food expenditures. Implementing traditional HCES questionnaires focused on household food consumption at home has the risk of underestimating FAFH by missing the increasing effect on the proportion of calories and expenditure through food systems changes. FAFH consumption is particularly important since food consumed outside the home tends to be more calorie-dense and less nutrient-dense than food consumed at home. The increase in the amount of food consumed away tends to be faster with increases in income.

Failing to account for FAFH has been shown to affect measures of poverty and inequality, including inequality in the distribution of dietary energy consumption. FAFH has a variety of sources, restaurants, school, place of work, street vendors. Survey design needs to be able to account for all of them as they can be of different importance to different groups in the population (so that failing to account for FAFH affects not only the mean but also the distribution of the indicators of interest).

An additional challenge is that while the ‘main food preparer’ can be expected to be reasonably informed about food at home, it is much more difficult for her to respond to FAFH questions as these relate mostly to meal events taking place away from her sight. Proxy respondents may be able to report on which household members consumed which meals away from the households, but they are unlikely to be informed about the cost or content of those meals. Such information can only be reliably collected via individual interviews.

Recommendations

Data collection on FAFH should preferably be done at the individual level, asking the questions separately for each individual. For all individuals reporting having had meals outside the home,

information should be solicited at a minimum regarding the value of meals by meal event (breakfast, lunch, dinner, snacks). While more research on this topic is urgently needed, based on the current state of knowledge the following guidelines are suggested for the design and implementation of a survey module for the measurement of consumption of food away from home in recall surveys:

- The practice of collecting FAFH information with just one question should be discontinued.
- The importance of FAFH warrants the design of a separate module, based on a clear definition of FAFH. In particular, surveys should be clear in identifying how to collect information on potentially ambiguous categories of food: "food prepared at home and consumed outside" and "food prepared outside and consumed at home". The latter can be integrated in the food at home module (e.g. take-out food) provided there is clarity to enumerators, respondents and data users that this is the case.
- Data collection should be organized around meal events, including snacks and drinks. At a minimum, surveys should collect information on the value of all meals consumed by meal event away from home (breakfast, lunch, dinner, solid snacks and drinks). The meal events list should be adapted to the local context
- Considerations regarding the feasibility, costs, and accuracy should inform the determination of which option to choose between individual module and the proxy respondent. FAFH is best collected via individual level interviews of adults. A proxy respondent can be used to report on children's meals away from home and other adults. Possible variations include:
    1. Proxy respondents (i.e. a household level module) can be used to report on the number of individuals who consumed meals away (as in block 4 of the 68th round of the India Survey). Detailed information on the meals (such as cost and meal content) should be collected directly from the relevant household member (including possibly on a targeted, carefully designed subsample).
    2. Total expenditures on FAFH can be collected at the household level using a daily FAFH record sheet provided to a trained proxy respondent.
- Surveys should identify the most frequent place of consumption for each meal event (restaurant, street vendor, work, school etc.), adapting the place of consumption categories to the local context.
- Surveys should use for FAFH the same reference period used in the food consumed at home module.
- The data to estimate FAFH related nutrient content, when feasible, will have to come from other data sources integrated to the HCES, such as a survey of food establishments (Farfan et al., 2017) or administrative data on the content of public meals (e.g. schools, social programs).

3.6. List of food items

Summary

The following basic principles should inform the design of the HCES list of food items:

- The number of food items should balance the lower memory lapses, costs and interview time associated with short lists, with the better recall and more comprehensive reporting associated with the longer list.
- Food items representing the large majority of food expenditures, but also nutrient rich food items should all be included.
- The description of the food items must be explicit enough to match only one entry in the reference food composition table.
- Each food item must be exclusive (each food group should be represented and each food item should belong to only one food group).

Adoption of a food classification system may help meeting these criteria. Although the adoption of a standard classification system may involve some challenges given the country specificity of the diet, such a classification is recommended as it allows survey harmonization both in terms of methodology and analysis/findings. For many of the basic purposes of HCES, such as the computation of CPI or the input into national account systems by far the most widely adopted standard classification is the COICOP. Harmonization eases cross-survey comparability, ensures international comparability of inflation and purchasing power parities, and allows for the assessment of the comprehensiveness of the food lists.

For surveys not aiming to inform CPI calculation or national accounts, the COICOP list may be overly extensive. There is widely acknowledge trade-off in the number of items to be included in a food list. Aggregated item lists (usually around 15 items) provide lower estimates than more detailed item lists. On the other hand, a too detailed list of items might have a negative effect, increasing enumerator and respondent fatigue. A universally valid solution does not exist because the optimal quantity of items strongly depends on regional food consumption habits. Thus a food list must be country specific, must represent the dietary/consumption habits of all segments of a population, and must capture evolving trend in dietary patterns. Useful information about the frequency and importance of each food item's dietary/expenditure patterns can be drawn from previous HCES or dietary survey data carried out in a given country.

As noted, food lists will inevitably be country-specific. Even so, some rules-of-thumb or general guiding principles can be identified to help survey designers drawing food lists to capture food consumption and expenditure providing information that is disaggregated in a way that can be useful for dietary quality analysis. Involving nutritionists in the design of the food lists can ensure their data needs are properly taken into account. Fiedler and Mwangi (2016) suggest that satisfying all these requirements will in most cases result in a list of 100-125 items, and many experts would agree with that, but this can only be seen as an indicative rule of thumb.

<u>Recommendations</u>

Different classifications can be used to harmonize data and foster comparability across surveys. These guidelines encourage survey designers to use the COICOP system, since nowadays it is the basis of classification used in a wide number of datasets, in line with the requirement of SNA 2008. Food COICOP classifications are mainly structured using the Basic Food Groups described in Table 4[38]. Additional regional or global food lists that can inform the development of a survey food list include the FoodEx2[39] classification now being tested in Asian countries (European Food Safety Authority, 2014).

Table 4 - Basic Food Groups for COICOP classifications

Basic Food Group	**Food Items in Each Group**
Bread and cereals	Rice Other cereals, flour and other products Bread Other bakery products Pasta products
Meat	Beef and veal Pork Lamb, mutton, and goat Poultry Other meats and meat preparations
Fish and seafood	Fresh, chilled, or frozen fish and seafood Preserved or processed fish and seafood
Milk, cheese, and eggs	Fresh milk Preserved milk and other milk products Cheeses Eggs and egg-based products
Oils and fats	Butter and margarine Other edible oils and fats
Fruit	Fresh or chilled fruits Frozen, preserved, or processed fruit and fruit-based products
Vegetables	Fresh or chilled vegetables other than potatoes Fresh or chilled potatoes and other tubers Frozen, preserved or processed vegetables and vegetable-based products
Sugar, jam, honey, chocolate and confectionery	Sugars Jams, marmalades, and honey Confectionery, chocolate, and ice cream
Food products not elsewhere classified	Salt, spices, condiments… Dessert preparations, soups, broths… Baby food, dietary preparations

[38] See http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5 for the detailed list of all COICOP categories. See also the revisions being discussed at the 49th Session of the UNSC: https://unstats.un.org/unsd/statcom/49th-session/documents/BG-Item3l-Classification-2-E.pdf.

[39] See https://www.efsa.europa.eu/en/supporting/pub/en-804 for more details on FoddEx2 classification.

Non-alcoholic beverages	Coffee, tea, cocoa Mineral waters, soft drinks, fruit and vegetable juices
Alcoholic beverages	Spirits Wine Beer Stimulants
Catering services	Restaurants, cafés and the like Canteens

Data should be collected on all of the types of foods and beverages that make up the country's human diet. Lists should be kept up to date to take into account changing dietary habits, and should be drawn having in mind that products that account for minimal budget shares can have particular nutritional values. A list of general principles that can guide the design of a food list includes the following criteria:

- The presence of foods from all the main food groups (e.g. the 16 food group classification on which the Household Dietary Diversity Score is based (Kennedy et al., 2011).
- An adequate representation of processed foods (with all the degree of processing from highly processed to moderately processed).
- The inclusion of only foods and no other commodities (principle of "food exclusivity"; e.g., the food list should not include an item composed by food and non-food items like "alcohol and tobacco").
- The list needs to include a reasonable number of individual items (the most common ones) for each of the main food groups. An "other" category should be added when relevant (e.g. "Other fruits", "Other vegetables", etc.) to record the acquisition or consumption of additional food items. It is important though that such categories remain marginal as quantities cannot be collected under these categories and food matching is imperfect.
- Food items (other than prepared dishes) should not span multiple food groups (e.g. avoid "eggs or milk products" as one group). Only group food with similar nutritional properties in one question (e.g. avoid "Mineral water or soft drinks"). Avoid grouping different status of the same food item with different nutritional properties (e.g. fresh or dried fish, fresh or dried milk).
- Avoid broad categories that do not allow identifying the type of food, such as "Snacks", "Canned foods", "Baby food", etc..
- Food items that are the object of product-specific government-subsidized programs should be listed individually in the food list.

- Foods that are fortified or have the potential to be the vehicle of food fortification programs (e.g. iodized salt, fortified flour or cooking oil) should be listed individually in the food list.
- Micronutrient (e.g. vitamin-A, iron) rich foods such as sweet potatoes and liver should be listed individually.
- Food lists can be built from national food composition tables or databases to ease later food matching. Because of the importance of having good and updated food composition tables or databases, it is also recommended that countries invest in the update or development of country or regional food composition tables.

3.7. Non-standard units of measurement

Summary

At present, there is no standard methodology applied to the collection of food quantities in cases where respondents (and enumerators) are less familiar with standard measurement units. Common practice has been either to require households to report everything in standard units, or for enumerators to estimate the standard unit conversions on an ad-hoc basis. Both approaches are problematic and lead to inaccuracies and inconsistencies in the data reported. Several NSOs maintain library of conversion factors for local units, but these are often incomplete, not updated, and implemented less than consistently. Units of measurement are critical for many aspects of data collection and quality assurance. Qualitative feedback from field practitioners, and extensive feedback from initial piloting of these methods, suggest that allowing non-standard units (NSUs) of measurement can increase the accuracy of reported quantities, primarily by reducing respondent burden. To properly benefit from allowing NSU options, reporting must be paired with a framework for consistently converting NSU into standard units, based on reliably documented conversion factors. Since conversion factors typically involve weighing and measuring, which are some of the same activities carried out in price surveys – at least for unpackaged foods like root crops and vegetables – the treatment of NSUs should also be covered in any guidelines for price surveys.

Recommendations

- The decision on whether to allow the use of non-standard units (NSU) of measurement should be addressed during the design phase. Doing so will reduce the tendency for units of measurement to be determined on an ad hoc or inconsistent basis during fieldwork.
- Though NSUs are used throughout the world (in countries of all income levels) the cost-benefit ratio of incorporating them into each survey should be evaluated, focusing particularly on their prevalence of use. If needed, conduct a pilot survey to determine the extent to which respondents need NSUs– extensively, minimally, or not at all.

- When feasible, allow households to report in both standard and non-standard units (NSU) of measure, according to what they are most familiar with for each item reported. If avoiding formal use of NSUs will still lead to ad hoc field conversions, then even a rather limited set of NSUs, with less of an implementation burden, may still be worthwhile.

- It is critical to establish (define or collect) conversion factors for all NSUs that will be used. Additional features to improve the accuracy of reported NSU quantities – such as market surveys to establish accurate NSUs and conversion factors, photo reference aides, and on-the-spot value verification using CAPI, may also be considered.

- NSOs and implementation partners should work together to establish NSU databases that can be used across surveys, effectively increasing the standardization of the units while also limiting the cost of their implementation. To this end, survey implementers should thoroughly document all NSU protocols and related conversion factors and make them publicly available.

4. Bibliography

Ahmed, N., Brzozowski, M. and Crossley, T.F. (2006). Measurement errors in recall food consumption data. IFS Working Papers, 06(21), Institute for Fiscal Studies.

Alderman, H. (1996). Saving and economic shocks. *Journal of Development Economics*, 51(2): 343–65.

Alkire, S. and Foster, J. 2011. Counting and Multidimensional Poverty Measurement', *Journal of Public Economics*, 95(7-8): 476- 487.

Attanasio, O. and Frayne, C. (2006). Do the poor pay more? Presented at: Eighth BREAD Conference on Development Economics. Ithaca, New York.

Backiny-Yetna, P., Steele, D. and Yacoubou Djima, I. (2017). The Impact of Household Food Consumption Data Collection Methods on Poverty and Inequality Measures in Niger. *Food Policy*, 72: 7-19.

Bai, J., Wahl, T.I., Lohmar, B.T. and Huang, J. (2010). Food Away From Home In Beijing: Effects Of Wealth, Time And "Free" Meals. *China Economic Review*, 21: 432-441.

Bailey, R.L., West Jr. K.P. and Black, R.E. (2015). The Epidemiology of Global Micronutrient Deficiencies. *Annals of Nutrition and Metabolism*. 66 (suppl. 2): 22-33.

Banerjee, A.V. and Duflo, E. (2011). *A Radical Rethink of the Way to Fight Global Poverty*. New York: Public Affairs, pp.303.

Barsky, R.B. and Miron, J.A. (1988). The seasonal cycle and the business cycle. NBER Working Paper No. 2688.

Bee, A., Meyer, B.D. and Sullivan, J.X. (2012). The Validity of Consumption Data: Are the Consumer Expenditure Interview and Diary Surveys Informative? NBER Working Paper Series, w18308.

Beegle, K., de Weerdt, J., Friedman, J. and Gibson, J. (2012). Methods of household survey consumption measurement through surveys: experimental results from Tanzania. *Journal of Development Economics*, 98(1): 3-18.

Bezerra, I.N. and Sichieri, R. (2009). Eating out of home and obesity: a Brazilian nationwide survey. *Public health nutrition*, 12(11): 2037-2043.

Binkley J.K., Eales, J. and Jekanowski, M. (2004) The relation between dietary change and rising US obesity. *Int J Obes Relat Metab Disord*, 24: 1032–1039.

Black, R.E., Allen, L.H. and Bhutta, Z.A. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet*, published online Jan 17. DOI:10.1016/S0140-6736(07)61690-0

Borlizzi, A., del Grossi, M.E. and Cafiero, C. (2017). National food security assessment through the analysis of food consumption data from household budget and expenditure surveys: the case of Brazil's Pesquisa de Orçamento Familiares 2008/09. *Food Policy*, 72: 20-26.

Bouis, H., Haddad, L. and Kennedy, E. (1992). Does it matter how we survey demand for food? Evidence from Kenya and the Philippines. *Food Policy*, 17(5): 349-360.

Bouis, H.E. (1994). The effect of income on demand for food in poor countries: Are our food consumption databases giving us reliable estimates? *Journal of Development Economics*, 44(1): 199-226.

Bradburn, N. (2010). Recall period in consumer expenditure surveys program. Unpublished. National opinion Research Center (NORC) University of Chicago. Available at: http://www.bls.gov/cex/methwrkshp_pap_bradburn.pdf.

Browning, M., Crossley, T.F. and Winter, J., 2014. The measurement of household consumption expenditures. *Annu. Rev. Econ.*, 6(1): 475-501.

Burns, C., Jackson, M. and Gibbons, C. (2002). Foods prepared outside the home: association with selected nutrients and body mass index in adult Australians. *Public Health Nutr*, 5: 441–448.

Capéau, B. (1995). Measurement error and functional form: a proposal to estimate prices and conversion rates from the ERHS1994, mimeo.

Capéau, B. and Dercon, S. (2006). Prices, Unit Values and Local Measurement Units in Rural Surveys: an Econometric Approach with an Application to Poverty Measurement in Ethiopia, *Journal of African Economies*, 15(2): 181-211.

Carletto, C., Sydney, G. Murray, S. and Zezza, A. (2016). “Cheaper, faster, and more than good enough: is GPS the new gold standard in land area measurement ?," Policy Research Working Paper Series 7759, The World Bank.

Carroll, C.D., Crossley, T.F. and Sabelhaus, J. (2015). Improving the measurement of consumer expenditures. National Bureau of Economic Research, *Studies in Income and Wealth*, 74.

Casley, D. and Kumar, K. (1988) *The Collection, Analysis and Use of Monitoring and Evaluation Data*. Baltimore, MD: Johns Hopkins University Press for World Bank, 1988.

Chen, S. and Ravallion, M. (2010). The Developing World is Poorer than We Thought, But No Less Successful in the Fight Against Poverty, *The Quarterly Journal of Economics*, 125(4): 1577–1625.

Coates, J., Colaiezzi, B., Fiedler, J.L., Wirth, J., Lividini, K. and Rogers, B. (2012). A program needs-driven approach to selecting dietary assessment methods for decision-making in food fortification programs. *Food & Nutrition Bulletin*, 33(2): 146-156.

Conforti, P., Grünberger, K. and Troubat, N. (2017). The impact of household budget surveys’ design on the measurement of food consumption. *Food Policy*, 72: 43-52.

Coudouel, A., Hentschel, J.S. and Wodon, Q.D. (2002). Poverty Measurement and Analysis, in J. Klugman (Ed.), *A Sourcebook for Poverty Reduction Strategies*, pp. 29–69. Washington, DC: World Bank.

D’Souza, A. and Jolliffe, D. (2012). Rising Food Prices and Coping Strategies: Household-Level Evidence from Afghanistan. *Journal of Development Studies*, 48(2): 282–99.

D'Souza, A. and Tandon, S. (2014). How well do household level data characterize food security? Evidence from the Bangladesh Integrated Household Survey. Mimeo. Paper presented at the Inter-Agency Working Group Workshop on Strengthening HCES for Food and Nutrition Analysis. Rome, November 5–6, 2014.

Deaton, A. (1997). *The Analysis of Household Surveys: a Microeconometric Approach to Development Policy*, Washington D.C. and Baltimore: The World Bank and Johns Hopkins University Press.

Deaton, A. and Grosh, M. (2000). Consumption, in M. Grosh and P. Glewwe (Eds.), *Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Surveys,* Washington, D.C.: World Bank.

Deaton, A, and *Zaidi, S.* (2002). "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." Washington, DC: World Bank.

Deaton, A, and Dupriez, O. (2011). Purchasing power parity exchange rates for the global poor. *American Economic Journal*, 3: 137-166.

De Weerdt, J., Beegle, K., Friedman, J. and Gibson, J. (2016). The challenge of measuring hunger through Survey. *Economic Development and Cultural Change*, 64(4): 727-758.

Diskin, P. (1999). Agricultural productivity indicators measurement guide. Food and Nutrition Technical Assistance Project (FANTA), Academy for Educational Development.

European Food Safety Authority, 2014. Guidance on the EU Menu methodology. EFSA Journal 2014; 12(12):3944, 77 pp.

FAO/WHO (1994) Fats and oils in human nutrition: report of a joint FAO/WHO expert consultation. FAO technical papers 57. Rome, FAO, 1994.

FAO (2006). World Food Summit. Rome, Italy: FAO.

FAO/INFOODS. 2012. Guidelines for food matching. Version1.2. Rome, Italy: FAO.

FAO, IFAD, WFP, UNICEF, WHO (2017). The State of Food Security and Nutrition in the World. Building resilience for peace and food security. Rome, Italy: FAO.

Farfán, G., Genoni, M.E. and Vakis, R. (2017). You Are What (and Where) You Eat: Capturing Food Away from Home in Welfare Measures. Food Policy, 72: 146-156.

Fermont, A. and Benson, T. (2011). Estimating yield of food crops grown by smallholder farmers. Washington D.C.: IFPRI.

Ferreira, F.H.G., Chen, S., Narayan, A., Sangraula, P., Dabalen, A.L., Serajuddin, U. and Yoshida, N. (2015). A Global Count of the Extreme Poor in 2012 : Data Issues, Methodology and Initial Results. Policy Research Working Paper. Washington, DC: The World Bank, October 3, 2015.

Fiedler, J.L., Sanghvi, T.G. and Saunders, M.K. (2008). A review of the micronutrient intervention cost literature: program design and policy lessons. *Int J Health Plann Manage,* 23: 373–397.

Fiedler, J.L. and Macdonald, B. (2009). A strategic approach to the unfinished fortification agenda: feasibility, costs, and cost-effectiveness analysis of fortification programs in 48 countries. *Food Nutr. Bull*., 30: 283–316.

Fiedler, J.L., Carletto, C. and Dupriez, O. (2012). Still waiting for Godot? Improving Household Consumption and Expenditures Surveys (HCES) to enable more evidence-based nutrition policies. *Food & Nutrition Bulletin*, 33(2): 242-251.

Fiedler, J.L. (2013). Towards Overcoming the Food Consumption Information Gap: Strengthening Household Consumption and Expenditures Surveys for Food and Nutrition Policymaking. *Global Food Security*, 2(1): 56-63.

Fiedler, J.L., Martin-Prével, Y. and Moursi, M. (2013). Relative Costs of 24-hour Recall and Household Consumption and Expenditures Surveys for Nutrition Analysis. *Food Nutr Bull*, 34(3): 318–330.

Fiedler, J.L. 2015. Household Consumption and Expenditures Surveys: Comparing 7- and 30- Day Recall Estimates of Food Consumption from India's 2011/12 National Survey Sample. Paper prepared for the International Statistical Institute's World Congress of Statistics Panel Measurement Issues for Food Insecurity, Poverty and Malnutrition: State of Research and Methods, Rio de Janeiro, July 27–31.

Fiedler, J.L. and Mwangi, D.M. (2016). Improving Household Consumption and Expenditure Surveys' Food Consumption Metrics: Developing a Strategic Approach to the Unfinished Agenda. IFPRI. Unpublished.

Fiedler, J.L. and Yadav, S. (2017). How Can We Better Capture Food Away from Home? Lessons from India's Linking Person-Level Meal and Household-Level Food Data. *Food Policy*, 72: 81-93.

Fiedler, J.L., Lividini, K., Bermudez, O.I. and Smitz, M.F. (2012). Household Consumption and Expenditures Surveys (HCES): a primer for food and nutrition analysts in low-and middle-income countries. *Food & Nutrition Bulletin*, 33(2): 170-184.

Finn, A. and Ranchhod, V. (2015). Genuine fakes: The prevalence and implications of data fabrication in a large South Africa survey. *The World Bank Economic Review*.

James, Suman Seth, Michael Lokshin, and Zurab Sajaia. 2013. A Unified Approach to Measuring Poverty and Inequality: Theory and Practice. Washington, DC: World Bank.

Friedman, J., Beegle, K., De Weerdt, J. and Gibson, J. (2017). Decomposing response error in food consumption measurement: implications for survey design from a randomized survey experiment in Tanzania. *Food Policy*, 72: 94-111.

Gewa, C.A., Murphy, S.P. and Neumann, C.G. (2007). Out-of-home food intake is often omitted from mothers' recalls of school children's intake in rural Kenya. *The Journal of Nutrition*, 137(9): 2154-2159.

Gibson, J. (2001). Measuring chronic poverty without a panel. *Journal of Development Economics*, 65(2): 243-266.

Gibson, J., and Rozelle, S. (2002). How elastic is calorie demand? Parametric, nonparametric, and semiparametric results for urban Papua New Guinea. Journal of Development Studies, 38(6): 23-46.

Gibson, J., Rozelle, S. and Huang, J. (2003), Improving estimates of inequality and poverty from Urban China's Household Income and Expenditure Survey, *Review of Income and Wealth*, 49(1): 53-68.

Gibson, R.S. (2005). *Principles of Nutritional Assessment*. 2nd edition. New York: Oxford University Press.

Gibson, J. (2007). A guide to using prices in poverty analysis. Mimeo The World Bank http://siteresources. worldbank. org/INTPA/Resources/429966-1092778639630/Prices_in_Poverty_Analysis_FINAL. pdf.

Gibson, J. and Kim, B. (2012). Testing the infrequent purchases model using direct measurement of hidden consumption from food stocks. *American journal of agricultural economics*, 94(1): 257-270.

Gibson, J. (2013). Two decades of poverty in Papua New Guinea. Presentation at Crawford School, ANU, Canberra, 5.

Gibson, J. and Kim, B. (2015). Hicksian separability does not hold over space: Implications for the design of household surveys and price questionnaires, *Journal of Development Economics*, 114(C): 34-40.

Gibson, J., Beegle, K., De Weerdt, J. and Friedman, J. (2015). What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxford Bulletin of Economics and Statistics*, 77(3): 466-474.

Gibson, J. (2016). Poverty Measurement: We Know Less than Policy Makers Realize. *Asia & the Pacific Policy Studies*, 3: 430–442.

Gilbert, C.L., Christiaensen, L. and Kaminski, J. (2016). Price seasonality in Africa: measurement and extent. World Bank Policy Research Working Paper, 7539. Washington, D.C.: World Bank Group.

Grosh, M. and Glewwe, P. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*. Washington, DC: World Bank. © World Bank.

Grünberger, K. (2017a). Estimating habitual food consumption with Household Consumption and Expenditure Surveys. How long households' food consumption should be observed? Mimeo, 2017.

Grünberger, K. (2017b). The impact of meal participation on food security indicators derived from Household Consumption and Expenditure Surveys. Mimeo, 2017.

Guthrie, J.F., Lin, B.H. and Frazao, E. (2002). Role of food prepared away from home in the American diet, 1977–78 versus 1994–96: changes and consequences. *J Nutr Educ Behav*, 34: 140–150.

Henríquez-Sánchez, P., Sánchez-Villegas, A., Doreste-Alonso, J., Ortiz-Andrellucchi, A., Pfrimer, K. and Serra-Maje, L. (2009). Dietary assessment methods for micronutrient intake: a systematic review on vitamins. *British Journal of Nutrition*, 102: S10---S37

Hurd, M. and Rohwedder, S. (2009). Methodological Innovations in Collecting Spending Data: The HRS Consumption and Activities Mail Survey. *Fiscal Studies*, 30(3-4), 435-459.

ILO (International Labour Organization) (2003). Household Income and Expenditure Statistics, Report II, Seventeenth International Conference of Labour Statisticians, Geneva, 24 November-3 December 2003.

Jolliffe, D. (2001). Measuring Absolute and Relative Poverty: The Sensitivity of Estimated Household Consumption to Survey Design, *Journal of Economic and Social Measurement*, 27(1): 1–23.

Jolliffe, D., Lanjouw, P., Chen, S., Kraay, A., Meyer, C., Negre, M., Prydz, E., Vakis, R., and Wethli, R. (2014). *A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals*. The World Bank, 2014.

Jolliffe, D. and Serajuddin (2015). Estimating Poverty with Panel Data, Comparably: an Example from Jordan, Policy Research Working Paper Series, 7373, The World Bank. Washington, D.C.: World Bank Group.

Kaara, J. and Ramasawmy, S. (2008). Food data collected using acquisition and consumption approaches with a seven-day recall method in Kenya's KIHBS 2005/2006, in FAO: *Deriving food security information from national household budget survey: experiences, achievements, challenges*, 69-79. Rome, Italy: FAO.

Kant, A.K. and Graubard, B.I. (2004). Eating out in America, 1987–2000: trends and nutritional correlates. *Preventive medicine*, 38(2): 243-249.

Kapteyn, A. (1994). The measurement of household cost functions: Revealed preference versus subjective measures, No 1994-3, Discussion Paper, Tilburg University, Center for Economic Research.

Kemsley, W.F.F. (1961). The Household Expenditure Enquiry of the Ministry of Labour: Variability in the 1953-54 Enquiry. *Applied Statistics*, 10(3), 117-135.

Kennedy, G., Ballard, T. and Dop, M. (2011). Guidelines for Measuring Household and Individual Dietary Diversity. Nutrition and Consumer Protection Division, FAO.

Kirlin, J.A. and Denbaly, M. (2017). Lessons Learned from the National Household Food Purchase and Acquisition Survey in the United States. *Food Policy*, 72: 62-71.

Kreuter, F., McCulloch, S., Presser, S. & Tourngeau, R. (2011). The effects of asking filter questions in interleafed versus grouped format. *Sociological Methods and Research*, 40(1): 88-104.

Liu, H., Wahl, T.I., Seale, J.L. and Bai, J. (2015). Household Composition, Income and Food-Away-from-Home Expenditure in Urban China. *Food Policy*, 51: 97–103.

Louzada, M.L., Levy, R.B., Martins, A.P.B., Claro, R.M., Steelel, E.M., Junior E.V., Cafiero, C. and Monteirol, C.A. (2017). Validating the usage of household food acquisition surveys to assess the consumption of ultra-processed foods: evidence from Brazil. *Food Policy*, 72: 112-120.

Ma, H., Huang, J., Fuller, F. and Rozelle, S. (2006). Getting rich and eating out: consumption of food away from home in urban China. *Canadian Journal of Agricultural Economics/Revue Canadienne d'Agroeconomie*, 54(1): 101-119.

Mahalanobis, P.C. and Sen, S.B. (1954). On Some Aspects of the Indian National Sample Survey. *Bulletin of the International Statistical Institute*, 34.

Mancino, L., Todd., J. Lin, B.H. (2009). Separating What We Eat From Where: Measuring the Effect of Food Away From Home on Diet Quality, *Food Policy*, 34(6): 557- 562.

Martirosova, D. (2008). Food data collected using acquisition and consumption approaches with daily diaries in Armenia's ILCS 2004, in FAO, *Deriving food security information from national household budget survey: experiences, achievements, challenges*. Rome, Italy: FAO, 59-68.

Maxwell, S. and Slater, R. (2003). Food Policy Old and New. *Development Policy Review*, 21: 531-553.

McWhinney, I. and Champion, H. (1974). The Canadian experience with recall and diary methods in consumer expenditure surveys, in S. Berg, *Annals of Economic and Social Measurement*, 3(2): 411-437. NBER.

Meenakshi, J.V. and Ray, R. (1999). Regional differences in India's food expenditure pattern: a complete demand systems approach, *Journal of International Development*, 11(1): 47-74.

Meng, T., Florkowski, W.J. and Kolavalii, S. (2012). Food expenditures and income in rural households in the northern region of Ghana, Agricultural & Applied Economics Association, Annual Meeting. Seattle, Washington, 2012.

Mihalopoulos, V.G. and Demoussis, M.P. (2001). Greek household consumption of food away from home: a microeconometric approach, *European Review of Agricultural Economics*, 28(4): 421–432.

Moltedo, A., Troubat, N., Lokshin, M. and Sajaia, Z. (2014). Analyzing Food Security Using Household Survey Data, FAO and World Bank.

Murphy, J., Casley, D.J. and Curry, J.J. (1991). Farmers' Estimations as a Source of Production Data. World Bank Technical Paper, 132. Washington, DC: World Bank.

Mutlu, S. and Gracia, A. (2006). Spanish food expenditure away from home (FAFH): by type of meal, *Applied Economics*, 38: 1037-47.

Nago, E.S., Lachat, C.K., Huybregts, L., Roberfroid, D., Dossa, R.A. and Kolsteren, P.W. (2010). Food, energy and macronutrient contribution of out-of-home foods in school-going adolescents in Cotonou, Benin. *British Journal of Nutrition*, 103(2): 281-288.

Neter, J. and Waksberg, J. (1964). A study of response errors in expenditures data from household interviews, *Journal of the American Statistical Association*, 59(305): 18-55.

NSSO (National Sample Survey Organization of India) (2003). Results of a Pilot Survey on Suitability of Different Reference Periods for Measuring Household Consumption, 475. Ministry of Statistics Y Programme Implementation, Government of India, New Delhi.

NSSO (National Sample Survey Organization of India) (2012). Instructions to Field Staff, Volume I, NSS 68th Round. Ministry of Statistics Y Programme Implementation, Government of India, New Delhi.

NSSO (National Sample Survey Organization of India) (2014). Level and Pattern of Consumer Expenditure 2011–12, NSS 68th Round, July 2011–July 2012. NSS Report 555(68/1.0/1). Ministry of Statistics Y Programme Implementation, Government of India, New Delhi.

Oguntona, C.R. and Tella, T.O. (1999). Street foods and dietary intakes of Nigerian urban market women. *Int J Food Sci Nutr*, 50(6):383-390.

Oseni, G., Durazo, J. and McGee, K. (2017). The Use of Non-standard Units for the Collection of Food Quantity: A Guidebook for Improving the Measurement of Food Consumption and Agricultural Production in Living Standards Surveys. Washington DC: World Bank.

Paxson, C.H. (1992). Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand, *American Economic Review*, 82(March): 15–33.

Paxson, C.H. (1993). Consumption and income seasonality in Thailand. *Journal of political Economy*, 101(1): 39-72.

Phillips, C.M., Dillon, C., Harrington, J.M., McCarthy, V.J.C., Kearney, P.M. and Fitzgerald, A.P. (2013). Defining Metabolically Healthy Obesity: Role of Dietary and Lifestyle Factors. PLoS ONE8(10): e76188.

Popkin, B.M. (2008). *The World Is Fat--The Fads, Trends, Policies, and Products That Are Fattening the Human Race*. New York: Avery-Penguin Group.

Popkin, B.M., Adair, L.S. and Ng, S.W. (2012). Global nutrition transition and the pandemic of obesity in developing countries, *Nutrition Reviews*, *70*(1): 3-21.

Poti, J.M., Duffey, K.J. and Popkin, B.M. (2014). The association of fast food consumption with poor dietary outcomes and obesity among children: is it the fast food or the remainder of the diet?, *The American Journal of Clinical Nutrition*, 99(1): 162–171.

Pradhan, M.P. (2001). *Welfare Analysis with a Proxy Consumption Measure*. TI Discussion Paper, 01-92/2. Amsterdam: Tinbergen Institute.

Pradhan, M.P. (2009). Welfare analysis with a proxy consumption measure: Evidence from a repeated experiment in Indonesia, *Fiscal Studies*, *30*(3/4): 391-417.

Redman, B. (1980). The Impact of Women's Time Allocation on Expenditure for Meals Away from Home and Prepared Foods, *American Journal of Agricultural Economics*, 62: 234–37.

Ravallion, M. (1998). Poverty Lines in Theory and Practice, Living Standards Measurement Study Working Paper No. 133, World Bank.

Ravallion, M. and Bidani, B. (1994). How robust is a poverty profile?, *World Bank Economic Review*, 8(1): 75–102.

Rimmer, D.J. (2001). An overview of food eaten outside the home in the United Kingdom National Food Survey and the new Expenditure and Food Survey, *Public health nutrition*, 4(5b): 1173-1175.

Schündeln, M. (2017). Multiple Visits and Data Quality in Household Surveys, *Oxford Bulletin of Economics and Statistics*, DOI: 10.1111/obes.12196.

Scott, C. and Amenuvegbe, B. (1990). Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana. Social Dimensions of Adjustment in Sub-Saharan Africa, Working Paper, 6, Washington, DC: World Bank.

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N. and Dabalen, A.L. (2015). Data deprivation: another

deprivation to end. Policy Research working paper; no. WPS 7252. Washington, D.C.: World Bank Group.

Silberstein, A.R. and Scott, S. (1991). Expenditure diary surveys and their associated errors, in P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz and S. Sudman (Eds.), *Measurement errors in surveys*, 303-326. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Smith, L.C., Alderman, H. and Aduayom, D. (2006). Food insecurity in sub-Saharan Africa: new estimates from household expenditure surveys. Washington DC: IFPRI.

Smith, L.C. and Subandoro, A. (2007). Measuring food security using household expenditure surveys, in IFPRI, Food Security in Practice. Washington DC: IFPRI.

Smith L., Dupriez, O. and Troubat, N. (2014). Assessment of the Reliability and Relevance of the Food Data Collected in National Household Consumption and Expenditure Surveys. International Household Survey Network. FAO, World Bank.

Smith, L.C. (2013). The Great Indian Calorie Debate: Explaining Rising Undernourishment during India's Rapid Economic Growth. IDS Working Paper 430.

Smith, L.C. (2015). The great Indian calorie debate: Explaining rising undernourishment during India's rapid economic growth, *Food Policy*, 50: 53-67.

Statistical Institute and Planning Institute of Jamaica (1996). Jamaica Survey of Living Conditions 1994. Kingston.

Stephens, M. (2003). "3rd of the Month": Do Social Security Recipients Smooth Consumption Between Checks?, *American Economic Review*, 93(1): 406-422.

Sudman, S. and Ferber, R. (1971). Experiments in obtaining consumer expenditures by diary methods. *Journal of the American statistical Association*, 66(336): 725-735.

Sujatha, T., Shatrugna, V. and Narasimha, G.V. (1997). Street food: an important source of energy for the urban worker. *Food Nutr Bull*, 18(4).

Tandon, S. and Landes, M.R. (2011). Counting Indians Food Insecure Is Complicated, Amber Waves, United States Department of Agriculture, Economic Research Service.

Troubat, N. and Grünberger, K. (2017). Impact of survey design in the estimation of habitual food consumption. The case of the 2007/08 Socio Economic Survey of Mongolia applied to urban households, *Food Policy*, 72: 132-145.

Tucker, C. and Bennett, C. (1988). Procedural effects in the collection of consumer expenditure information: the diary operation test, in *Proceedings of the Section on Survey Research Methods*, 256-261 *Washington, DC: American Statistical Association.*

Tucker, C. (1992). The estimation of instrument effects on data quality in the Consumer Expenditure Diary Survey, *Journal of Official Statistics*, 8(1): 41.

Turner, R. (1961). Inter-Week Variations in Expenditure Recorded During a Two-Week Survey of Family Expenditure, *Journal of the Royal Statistical Society*, Series C, 10(3): 136-146.

United Nations. (2014). Report of the World Bank on Improving Household Surveys in the Post-2015 Development Era: Issues and Recommendations towards a Shared Agenda. UN Economic and Social Council E/CN.3/2015/10, New York.

United Nations Statistics Division (2005). Handbook on Poverty Statistics: Concepts, Methods and Policy Use. Special Project on Poverty Statistics, United Nations. New York: 2005. Available: http://unstats.un.org/unsd/methods/poverty/pdf/UN_Book%20FINAL%2030%20Dec%2005.pdf

United States Department of Agriculture (1946). National Food Guide. AIS-53, 1946.

United States Department of Agriculture, Economic Research Service (2011). International Food Consumption Patterns, Washington, DC: USDA/ERS.

Vandevijvere, S., Lachat, C., Kolsteren, P. and Van Oyen, H. (2009). Eating out of home in Belgium: current situation and policy implications, *British Journal of Nutrition*, 102: 921-928.

van't Riet., H, den Hartog, A.P. and Hooftman, D.A. (2003). Determinants of non-home-prepared food consumption in two low-income areas in Nairobi. *Nutrition*, 19: 1006-1012.

Wanner, N., Cafiero, C., Troubat, N. and Conforti, P. (2014). Refinement to the FAO methodology for estimating the prevalence of undernourishment indicator. FAO Statistic Division. ESS working paper 14-05. Rome.

Weisell, R. and Dop, M.C. (2012). The adult male equivalent concept and its application to Household Consumption and Expenditures Surveys (HCES), *Food & Nutrition Bulletin*, 33(2): 157-162.

Yen, S. and Jones, A.M. (1997). Household Consumption of Cheese: An Inverse Hyperbolic Sine Double-Hurdle Model with Dependent Errors, *American Journal of Agricultural Economics*, 79(1): 246–251.

You, J. (2014). Dietary change, nutrient transition and food security in fast-growing China, in R. Jha, R. Gaiha and A.B. Deolalika (Eds.), *Handbook on Food*, chapter 9: 204-245, Edward Elgar Publishing.

Zezza, A., Carletto, G., Fiedler, J.L., Gennari, P. and Jolliffe, D. (2017). Food counts. Measuring food consumption and expenditures in household consumption and expenditure surveys (HCES). *Food Policy*, 72: 1-6.

5. Glossary

Coefficient of Variation (CV) – it is a measure of dispersion of a (frequency or probability) distribution, and it is defined as the ratio of the distribution's standard deviation to the mean.

Cost of Basic Needs - cost of acquiring enough food for adequate dietary energy – usually 2,100 Calories per person per day –with the addition of the cost of other essentials such as clothing and shelter.

Diary – method of data collection. One or more individuals in the household are asked to record at daily level respectively the household's or individual's food acquired and/or consumed during the reference period of food data collection.

Dietary energy consumption (DEC) – Measure of calories consumed by a population group. It is expressed in kilocalories per capita per day. It is estimated from households' food quantities collected in the survey after being adjusted for non-edible portions (e.g. bones and peels). The DEC is calculated by converting consumed quantities into calories using Food Composition Table.

Food Acquisition – food (in terms of quantity and/or monetary values) acquired by households during the reference period of food data collection for the purpose of proper consumption. It includes food from purchases, own-production and from other sources such as food received as gift, aid or as payment.

Food-away-from-home – (from US bureau of Labor Statistics) Food away from home includes all meals (breakfast and brunch, lunch, dinner and snacks and nonalcoholic beverages) including tips at fast food, take-out, delivery, concession stands, buffet and cafeteria, at full-service restaurants, and at vending machines and mobile vendors. Also included are board (including at school), meals as pay, special catered affairs, such as weddings, bar mitzvahs, and confirmations, school lunches, and meals away from home on trips.

Food Composition Table – Table with information on nutritional properties of foods usually consumed in a country (national) or a group of countries (regional). The table includes information on the content of selected macro and micronutrients and the fraction of edible portions. It is generated according to international guidelines to be comparable and reliable.

Food Consumption – food (in terms of quantity and/or monetary values) consumed by a household over the reference period of food data collection. The consumed food can either come from food acquired during the same reference period or from household stocks.

Food expenditure – monetary value that was paid to purchase a specific amount of food.

Food exclusivity - the food list must include only foods and no other commodities such as tobacco or chewed stimulants.

Food from own production – food acquired by one or more household members from the household's own production for household's at-home consumption.

Food monetary value – it includes food expenditure for purchases and the amount of money that respondent would have spent if he/she would have paid for the food which was acquired from own production or received in-kind.

Food purchased – food acquired by one or more household members. It includes food consumed at-home or away from home.

Food from other sources – food acquired or consumed at-home from sources different from purchases and own production (e.g. gift, charity, as part of payment, government programs excluding those were food is consumed away from home such as school feeding programs, etc.)

Fortified food – food to which the content of a vitamin or mineral was deliberately increased irrespective of whether the nutrient was in the food before processing it. Aims to improve the nutritional status of individuals within a population by improving the nutritional quality of a food consumed by the population.

Household – usually defined as an individual or group of individuals, related or unrelated, who live together in the same dwelling unit, who acknowledge one adult male or female as the head of household, who share the same living arrangements, who pool some, or all, of their income and wealth and who consume certain types of goods and services collectively, mainly housing and food, and are considered as one unit.

List of food items – foods listed in a pre-filled food consumption module.

Memory decay – it refers to the process of forgetting details (e.g. quantities, monetary values and/or foods) linked to consumption/acquisition.

Non-standard units – measurement units for food consumption (or acquired for consumption) different from those included in the metric system (e.g., kilogram, liter). Quantities expressed in non-standard units needs to be converted into standard (metric) units to derive nutritional properties from Food Composition Tables. Non-standard units can vary between and within countries.

Partakers – number of individuals who actually consumed the total amount of food reported by the household during the reference period of food data collection. The household size might be different from the number of food partakers because foods may be shared with non-household members such as guest, employees and relatives and/or household members might have been absent in the household during the reference period.

Percentile – within a probability distribution identifies the value of the observed variable below which a percentage (equal to the percentile) of the observations falls. For example, the 20^{th} percentile corresponding to 940 Kcal/caput/day means that 20 per cent of the population consumed less than 940 Kcal/caput/day.

Prevalence of Undernourishment (PoU) – Measure of the proportion of individuals in a population suffering from chronic hunger (a state, lasting for at least one year, of inability to acquire enough food, defined as a level of food intake insufficient to meet dietary energy requirements). Within a probability distribution framework is the probability that a randomly selected individual from a population has an inadequate habitual access to food to satisfy his/her dietary energy requirements.

Processed foods - The term 'processed food' applies to any food that has been altered from its natural state in some way, either for safety reasons or convenience

Recall interviews – method of data collection. One or all individuals in the household are asked to recall the household's or individual's food acquired and/or consumed during the reference period of food data collection (recall period).

Recall period - reference period of data collection of a food recall interview. Respondents report the amount of food consumed or acquired during the recall period.

Reference period – time period for which respondents are asked to report food acquisitions/consumption.

Relevance (as defined in Smith et al 2014) – whether the data provide the information or indicators needed by different users.

Reliability (as defined in Smith et al 2014) - whether the survey design and method complies with good practice, on the basis of a number of criteria.

Respondent burden - effort required by the respondent to answer a questionnaire. A longer questionnaire usually increases the respondent burden and may decrease the response rate and the accuracy of the response.

Round – the total period (usually 12 months) over which a survey is carried out.

Seasonality – effect on variation of food acquired/consumed and related expenditures over a long period (e.g., 6 months, one year). Seasonality is usually linked to agricultural production season. Other cyclical events such as floods and droughts may cause such variation affecting both food availability and prices.

Snack

Telescoping – the action of reporting food consumption/acquisition that actually occurred before or after the recall period. The most common type of telescoping is including events that occurred before a short recall period, leading to an over-estimation of consumption/acquisition.

Usual month – a typical month (during the previous year or other reference period such as six months) over which respondents are asked to remember their food acquisition/consumption.

Visit – it refers to the visit of enumerators conducting the interviews. Over a survey round multiple visits may occur.