Improving Household Consumption and Expenditure Surveys’ Food Consumption Metrics

Developing a Strategic Approach to the Unfinished Agenda

John L. Fiedler

Dena M. Mwangi

Poverty, Health, and Nutrition Division
INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), established in 1975, provides evidence-based policy solutions to sustainably end hunger and malnutrition, and reduce poverty. The institute conducts research, communicates results, optimizes partnerships, and builds capacity to ensure sustainable food production, promote healthy food systems, improve markets and trade, transform agriculture, build resilience, and strengthen institutions and governance. Gender is considered in all of the institute’s work. IFPRI collaborates with partners around the world, including development implementers, public institutions, the private sector, and farmers’ organizations, to ensure that local, national, regional, and global food policies are based on evidence.

AUTHORS

John L. (Jack) Fiedler was a senior research fellow in the Poverty, Health, and Nutrition Division of the International Food Policy Research Institute (IFPRI), Washington, DC when this paper was written. He is now retired.

Dena M. Mwangi was a research analyst in the Poverty, Health, and Nutrition Division of IFPRI, when this paper was written. She is currently a consultant for the Development Impact Evaluation Team in the Research Group at the World Bank.

Notices

1. IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI’s Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by the International Food Policy Research Institute.

2. The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

3. This publication is available under the Creative Commons Attribution 4.0 International License (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

Copyright 2016 International Food Policy Research Institute. All rights reserved. Sections of this material may be reproduced for personal and not-for-profit use without the express written permission of but with acknowledgment to IFPRI. To reproduce the material contained herein for profit or commercial use requires express written permission. To obtain permission, contact ifpri-copyright@cgiar.org.
# Contents

Abstract v
Acknowledgments vi

1. Introduction 1

3. Study Topic 1: Assessing the Adult Male Consumption Equivalent as a Proxy for Intrahousehold Distribution 13

4. Study Topic 2: Optimizing the Recall Period 16

5. Study Topic 3: Modifying the Food List to Better Capture Important Food Sources of Essential Nutrients 30


7. Study Topic 5: Better Capturing the Consumption of Food Consumed Away from Home 40

8. Study Topic 6: Adjusting Food Consumption Estimates for Meal Participation 47

9. Prioritizing the Six Key Studies 54

Appendix A: Supplementary Table 61

Appendix B: Considerations for Developing Household Consumption and Expenditure Survey Country Typologies 62

References 64
Tables

2.1 Potential types and sources of error in HCESs and those targeted by the unfinished agenda’s six priority studies and the 100-country HCES review’s eight key recommendations 10
4.1 Number of replications required per individual for 95 percent of observed values to lie within a specified percentage of the true mean in a sample of 184 Malawian women 20
4.2 Changes in individuals’ intake inadequacy status by reference period, Bangladesh Household Income and Expenditure Survey, 2010 28
5.1 Micronutrient content of different types and preparations of beans, Uganda 33
5.2 Key parameters of the Indonesian National Socioeconomic Survey experiment 35
5.3 Example of survey subsection on specific food items consumed in the past week 36
8.1 Household consumption and expenditure surveys (HCESs) collecting information on meal participation, as identified in a review of 77 countries’ HCESs 50
8.2 Types of information collected on food given to non–household members 50
8.3 Average number of meals consumed per household by type of meal in the past 30 days, India, 1983 51
9.1 A prioritized, global agenda for the key household consumption and expenditure survey food consumption studies 56
9.2 Food consumption and nutrient intake information needs during nutrition program cycles 59
A.1 More recent food insecurity and nutrition: Related applications of household consumption and expenditure surveys 61

Figures

4.1 Estimates of the prevalence of inadequate and excessive intakes in the same population 19
4.2 Prevalence of inadequate individual intakes with seven different recall periods 27
4.3 Percentage of total changes in inadequate intake prevalence captured with increasing length of recall period, Bangladesh Household Income and Expenditure Survey, 2010 28
7.1 The diversity of household consumption and expenditure survey approaches to capturing food away from home 41
7.2 The Indian household consumption and expenditure survey questionnaire’s sections on meals and meal participation 45

Boxes

2.1 The subset of the 100-country review: Identified key challenges that we defer to the non–nutrition analysts 11
4.1 Bounded recall 17
4.2 How interviewees formulate responses: Episodic enumeration versus estimation 21
4.3 Purchase data, food stocks, and recall period 23
4.4 “Usual week” or “usual month” as an alternative to a longer, fixed recall period 25
5.1 Some food list–specific fortification and biofortification considerations 32
ABSTRACT

As the nature of global malnutrition changes, there is a growing need and increasing urgency for more and better information about food consumption and dietary patterns. The past two decades have seen a dramatic increase in the number, availability, and analysis of the food consumption data collected in a variety of multipurpose household surveys, referred to collectively as household consumption and expenditure surveys (HCESs). These surveys are heterogeneous, and their quality varies substantially by country. Still, they share some common shortcomings in their measurement of food consumption, nutrient intakes, and nutrition status that undermine their relevance and reliability for purposes of designing and implementing food policies and programs. This review crafts a strategic approach to the unfinished global agenda of improving HCESs’ collection of food consumption data. Starting with the priority studies recommended by a 100-country HCES review (Smith, Dupriez, and Troubat 2014), it focuses on a strategic subset of those studies that deal most directly and exclusively with the measurement of food, and that are of fundamental importance to all HCES stakeholders in low- and middle-income countries.

Drawing from the literature, this study provides a more detailed, more circumspect justification as to why these particular studies are needed, while identifying key hypotheses, explaining why these studies are of growing urgency, and demonstrating why now is a propitious time for undertaking them. The review also identifies important study design considerations while pointing out potential challenges to successful implementation stemming from technical capacity, economic, administrative, and political considerations. Six key studies are rank ordered from a global perspective as follows, taking into account (1) the likely shared consensus that a topic is an important source of measurement error in estimating consumption; (2) the perceived urgency of the need for addressing a particular source of measurement error; (3) the perceived likelihood of success—that is, that the efforts will improve the accuracy of measurement; (4) whether or not the study entails modifying the questionnaire; (5) the ease with which a study may begin; and (6) the extent to which the study is independent of necessary negotiations with existing HCES stakeholders because of the types of changes it is likely to entail (in either the questionnaire or the way the data have traditionally been processed). The following topics are discussed in detail:

- Assessing the adult male consumption equivalent as a proxy for the intrahousehold distribution of food
- Optimizing the recall period
- Modifying the food list to better capture important food sources of essential nutrients
- Improving nutrient imputation methods for foods with only expenditure (not quantity) data
- Better capturing the consumption of food away from home
- Adjusting food consumption estimates for meal participation

Given the heterogeneity of HCESs and country contexts, country-specific prioritizations are likely to vary and require taking into account country-specific weightings of criteria, technical and resource constraints, political will, and other considerations.

**Keywords:** household surveys, household consumption and expenditures surveys (HCES), dietary assessment, food consumption, nutrient intakes, nutrition, nutrition policy
ACKNOWLEDGMENTS

This paper was prepared for the International Dietary Data Expansion (INDDEX) Project of the Friedman School of Nutrition Science and Policy, Tufts University with financial support from the Bill & Melinda Gates Foundation. IFPRI is a partner agency on the INDDEX Project.

The authors thank Terra Carter, IFPRI Discussion Paper Coordinator, for her assistance in preparing the paper for publication.
Preview

There is a long-standing food consumption information gap that has encouraged the use of existing routinely conducted household consumption and expenditure surveys to address food and nutrition policies and programs.

Food constitutes about half of total household expenditures in low-income countries and is at the core of basic human welfare, figuring prominently in food security, nutrition, health, and poverty (USDA Economic Research Service 2011). Thus its accurate and reliable measurement is critical to understanding and monitoring a population’s welfare, and is of keen interest to a host of national and international organizations. This review seeks to craft a strategic approach to the unfinished global agenda of improving the collection of food consumption data in household consumption and expenditure surveys (HCESs). Starting with the six priority studies recommended by a 100-country HCES review (Smith, Dupriez, and Troubat 2014), it focuses on a strategic subset of those studies that are of fundamental importance to all HCES stakeholders in low- and middle-income countries, namely, those addressing shortcomings that deal most directly and exclusively with the measurement of food. By maximizing the potential value-added of strengthened and standardized collection and use of HCESs to inform food, nutrition, and agriculture programs and policies, it is hoped that this approach will benefit from the support of all key existing HCES stakeholders—central bankers; poverty analysts; and ministries of finance, economy, and agriculture—at the same time that it helps to accelerate the process of change.

The nature of global malnutrition is changing. Since 1980, the prevalence of obesity has nearly doubled while that of stunting has fallen—since 1990, by more than one-third. Globally there remain 165 million people who are stunted and 1 billion people who lack adequate food, but twice the latter number now suffer from micronutrient deficiencies. Half a billion persons are now obese, and fully one-third of the world’s population is overweight or obese (Black et al. 2013; FAO 2013; Ng et al. 2014; UNICEF 2013; WHO 2013). The morphing nature of malnutrition, and the speed with which it is occurring, has shifted attention from a narrow focus on calories, quantity of food, and food security to an increasing concern about nutrients, the quality of food and diets, and nutrition security (Acharya et al. 2014). While there is now a greater appreciation of the importance of diet quality, we remain ill equipped to do much about it in the short term because—at other than a national level—globally, and particularly in low- and middle-income countries, we know so little about current food consumption and dietary patterns or how they are changing. Our ignorance stems from the long-standing food consumption information gap.

Most countries have undertaken one or more individual-level dietary intake surveys. With the notable exceptions of Colombia, Indonesia, and the Philippines, however, nearly all have done so haphazardly and generally only once or twice over the past three decades. In a recent multiple-year undertaking, Imamura and colleagues (2015) collected and analyzed food consumption surveys to assess global trends in dietary quality. They were able to obtain 325 individual- and household-level food consumption surveys covering 187 countries over the past 25 years. Of these, 171 were conducted in the post-2000 era, providing an average of less than one survey per country over the past decade and a half. The lack of a more routine and systematic approach has been discouraged because the surveys—and in particular those employing the two most widely accepted scientific methods of quantifying food and nutrient intakes, individual-level observer-weighed food records and 24-hour recall (24HR)—are expensive as well as difficult to conduct and analyze (R. Gibson 2005; Neufeld and Tolentino 2013; Fiedler, Martin-Prével, and Moursi 2013).

As a result, the vast majority of countries—low- and high-income countries alike—continue to rely on national food balance sheets (FBSs) to diagnose, plan, program, evaluate, and monitor food consumption and nutrition. FBSs, however, contain only information about national food availability; they do not provide information on access to available food, about how food is distributed within the country, or about how much of it is consumed or by whom. Nor are these data produced regularly; only 38 percent of the production data used for compiling FBSs are official data from countries (Keita and
Mane 2015). In short, FBSs do not contain any information about household- or individual-level food consumption patterns, yet that is precisely the kind of information that is essential to understanding the changing face of malnutrition. This situation has given rise to new proposals and strategies to enable “bridging the evidence gap” (United States Institute of Medicine 2010) and to “update the way ‘food’ is monitored” (Herforth 2015, 155). The long-standing food consumption information gap has prompted turning to alternative data sources, such as HCESs. Despite acknowledged shortcomings, HCESs are increasingly being used to analyze food and nutrition issues and inform food, nutrition, and agriculture programs and policies.

Stakeholders have increasingly turned to HCESs because they are done routinely, once every three to five years in more than 125 countries, and are done with increasing frequency in a number of countries (Ravallion 2011). Furthermore, these multipurpose surveys have large samples (about 10,000 households), are statistically representative both nationally and subnationally, are much less costly than other dietary assessment data sources, and are already being conducted and paid for by government agencies (Fiedler, Carletto, and Dupriez 2012; Fiedler, Martin-Prével, and Moursi 2013).

An Overview of Household Consumption and Expenditure Surveys and Trends in Their Use

Preview

There is growing use of HCESs in increasingly diverse food and nutrition programming and policy applications.

HCESs are a collection of multipurpose surveys that are used to measure (among other things) food consumption and/or acquisition. HCESs include (1) household budget surveys, (2) household income and expenditure surveys, (3) surveys of the World Bank’s Living Standards Measurement Study (LSMS), (4) integrated household surveys, (5) surveys using the Core Welfare Indicators Questionnaire, and (6) welfare monitoring surveys. These surveys have diverse objectives, including providing input into consumer price indexes and national accounts, understanding poverty and other dimensions of welfare (including food and nutrition security), and monitoring selected socioeconomic indicators.

Over the last two decades there has been a dramatic increase in the number and availability of HCESs. In 1990, the World Bank’s World Development Report was based on 22 datasets from 22 countries. Today 125 countries have consumption or expenditure information, and there are more than 850 surveys, an average of seven per country (Ravallion 2011). For decades, countries did not make their HCES data generally available. That practice has changed markedly in the past few years, spurred by the Millennium Development Goals (MDGs), growing demands for increased accountability and access to government documents, and the growing transparency requirements of partner organizations (Serajuddin et al. 2015; Demombynes and Sandefur 2014). The increased availability of HCESs has sparked the use of these surveys in ways not conceived of when they were first designed. Moreover, as the data are being increasingly repurposed by new stakeholders with different interests, these actors have brought with them demands for additional types of data and new approaches to the surveys. HCESs are now being used for a host of food consumption and nutrition-related analyses that they were never intended to support. Among these applications are the use of HCESs to

1 General introductions to HCESs and key issues in their use for food and nutrition analysis include an HCES primer (Fiedler, Lividini, et al. 2012); a guide to using HCESs to measure food security (Smith and Subandoro 2007); a guide to and criteria for selecting among the HCES, FBS, 24HR, and food frequency methods for data to design and monitor fortification programs (Coates et al. 2012); descriptions of how to use an HCES together with food composition table data to develop estimates of nutrient availability (Bermudez et al. 2012; Dary and Imhoff-Kunsch 2012; Imhoff-Kunsch et al. 2012); a comparative analysis of the costs of HCES and 24HR (Fiedler, Martin-Prével, and Moursi 2013); general reviews of HCESs that also lay out a global strategy for strengthening HCESs for undertaking food and nutrition analyses (Fiedler, Carletto, and Dupriez 2012; Fiedler 2013); and an assessment of the relevance and reliability of HCESs based on a review of 100 countries’ questionnaires (Smith, Dupriez, and Troubat 2014).
• conduct subnational food security analysis,
• assess diet quality and dietary change,
• assess dietary diversity and the diversity of food supplies,
• proxy nutrient intakes and the prevalence of inadequate nutrient intakes, and
• design and model the impact of nutrition programs and nutrition program portfolios.

For a more complete list of new applications and citations of the papers in which they were reported, see Appendix A.

Putting the Cart before the Horse? Assessing the Quality of Household Consumption and Expenditure Surveys’ Food Consumption Data

Preview

Even though the use of HCESs is growing, there remain questions about their quality and the appropriateness of some of their uses, in part because we don’t know “true consumption” and in part because HCESs are heterogeneous.

How good are HCES food consumption data for food consumption and nutrition–related analyses? Despite the considerable amount of analytical work that has already been done with HCESs, and despite the fact that they have already been used to design programs and guide policy, there is considerable controversy about the quality of HCES data. There are several reasons why there is no definitive judgment about the quality of food consumption data in HCESs. The major reason is that we simply do not know what “true consumption” is. The best available benchmark is data from surveys using what are considered the most scientifically rigorous approaches, namely, observed-weighed food records (OWFR) or 24HR, although they too have shortcomings. How well do HCES estimates compare with those of OWFR or 24HR? Surprisingly, there is considerable uncertainty about that as well, owing to the lack of directly comparable HCES and 24HR surveys. Although comparative analyses of selected indicators in HCESs and 24HR surveys have been undertaken (as discussed below), the lack of directly comparable surveys precludes a more definitive judgment, which has prompted some to assert that those who have been using HCES-based measures of food consumption have put the cart before the horse (Murphy, Carriquiry, and Ruel 2012).

Why are there so few comparable HCES and 24HR or OWFR surveys? Can we expect this situation—the dearth of comparable surveys—to change in the near future? How different are the surveys and why? Several factors contribute to this impasse. Owing primarily to cost considerations and budgetary constraints, most 24HR surveys cover only a small geographic area, such as a limited number of districts or states. When they are nationwide studies, they generally have covered a relatively small number of households (persons) and usually have not taken into account the diversity of regional dietary patterns in the development of their samples (that is, in their sample size calculations). As a consequence, the national average portrays a diet that may not be characteristic of any of the inhabitants of the country; that is, it may result in an ecological fallacy.

Another cost- and survey size–related constraint that has made most 24HR surveys less than ideally comparable to HCES surveys is the generally more restricted population included in 24HR surveys; that is, they are not demographically representative either. These individual-based studies most often focus on vulnerable populations, such as children less than five years of age and their mothers. Usually there is no information about other members of the household, precluding the calculation of household-level estimates to compare with the household-based HCES.

Developing individual-level estimates of nutrient intakes using an HCES requires making an assumption about how food is distributed within the household. Most commonly it is assumed that food is distributed among household members in direct proportion to their age- and sex-adjusted energy requirements as estimated using the adult male equivalent (AME), also known as the adult consumption
equivalent (FAO 2004), and assuming a moderate physical activity level and basal metabolic rate. How well does the AME-based approach approximate the actual distribution of food within households? Here, too, the jury is still out, owing again to the fact that less-than-ideal data are available with which to make the assessment. The paucity of assessments of using the AME to proxy the intrahousehold distribution of food, and the questionable external validity of the few assessments that exist, are a direct reflection of the challenges of making direct empirical comparisons using data from these two types of surveys, which are done at such different scales; use different sampling frames, periodicity, and units of analysis; and contain data on different sets of household members.

In sum, to date there have been only opportunistic, piecemeal, and less-than-ideal comparisons of HCESs with 24HR or OWFR studies, and thus no definitive judgment rendered about the quality of HCESs in terms of how well they proxy 24HR or OWFR. To address this information gap, at minimum we need to encourage the collection of complete household roster data in 24HR and the unpacking of HCESs’ household-level responses, prioritizing children under three years of age. A first step in this direction in low-income countries is currently being implemented in Bangladesh, where the International Food Policy Research Institute (IFPRI) is fielding a second round of the Bangladesh Integrated Household Survey (BIHS) that includes for the first time individual-specific questions for children under five.2 The data from this one-time survey will be available in 2016.

A second reason for ambivalence about the quality of HCESs stems from their heterogeneity. Reflecting the diversity of their original objectives, HCESs vary in terms of data capture methods; recall periods; the length and composition of their food lists; whether they collect only information about food purchases or food consumption or some combination thereof; what food sources they collect data on—purchased food, food consumed from own production, or food received as payment in kind or free of charge; whether they collect information about the value of food purchases, the quantities of food purchased, or both; the extent to which they are subject to seasonality; how well they capture food eaten away from home; how well trained and conscientious their enumerators are; and how well the data are cleaned and processed. A particular HCES may do well in addressing some of these issues and criteria, but may not do well on all of them. Since all of these factors affect the quality of the measurement of food consumption, not doing well on any one of them undermines confidence in the accuracy of the data and leaves uncertainty about the quality of HCES food consumption data.

Without knowing “true consumption” or having comparable OWFR or 24HR data, we simply do not know the quality of HCES-based estimates of consumption (whether they are at the household or individual level): in short, a definitive characterization cannot be made. Definitive, general characterizations, however, are not the only assessments of HCESs that are of interest or relevance to food and nutrition analysts and policy makers. Clearly, what might be regarded as good or acceptable data depends on the particular application of the data, and for some measures and some applications there has been growing evidence that high-quality HCESs can help fill the food consumption information gap.

Starting 25 years ago with the Data Food Networking (DAFNE) project (Trichopoulou and Lagiou 1997; Lagiou, Trichopoulou, and DAFNE contributors 2001; National and Kapodistrian University of Athens 2010; DAFNE-ANEMOS 2011), a growing body of work has demonstrated that some of the food data collected in HCESs can provide a reasonable proxy for information about food consumption and contribute to making nutrition policy more evidence based.3 A small, but growing, 2 There is a precedent for this alternative approach. A special supplementary data collection module was added to the 1998 Danish Household Expenditure Survey (a continuous survey of 800 households per year using a 14-day diary) to collect data about for whom each good was bought (Bonke and Browning 2009). For each good that the household reported consuming, the respondent had to check one of five additional columns that were added to the diary form to indicate for whom the good was bought. Responses were “mostly for the husband,” “mostly for the wife,” “for the whole family,” “for the child(ren) in the household,” “for people outside the household,” and “do not know.” Eighty-six percent of all foods consumed at home and 55 percent of all “eating out” were reported to have been “for the whole family.” Although no quantitative information was provided on the additional time the questions added to the interview, it is promising that the module was reported to be “not very time-intensive.”

3 See Fiedler (2013) for a chronology of the major developments. There have been warnings about the impact of alternative data.
number of studies have juxtaposed 24HR and HCES data and found relatively high levels of consistency between the surveys in quantifying the proportion of households reported to be consuming most food items (Lambe et al. 1998; Friel et al. 2001; Naska, Paterakis, et al. 2001; Naska, Vasdekis, and Trichopoulou 2001; Lividini, Fiedler, and Bermudez 2013; Rambeloson et al. 2012); the proportion of households purchasing most food items (Lambe et al. 1998; Friel et al. 2001; Naska, Paterakis, et al. 2001; Naska, Vasdekis, and Trichopoulou 2001; Lividini et al. 2013 Rambeloson et al. 2012); total caloric intake (Sekula et al. 2005; Naska et al. 2007; Rambeloson et al. 2012), the nutrient density of most food items consumed (Rambeloson et al. 2012; Naska et al. 2007; Nelson et al. 1985; Sekula et al. 2005); the nutrient density of most food items purchased (Rambeloson et al. 2012; Naska et al. 2007; Nelson et al. 1985; Sekula et al. 2005); and estimated median intakes of specific macro- and micronutrients, which have found mixed results (Sekula et al. 2005; Rambeloson et al. 2012).

Although these results are generally encouraging about the use of HCES data to proxy consumption, in light of the heterogeneity of HCESs, the external validity of these findings is likely to be limited, leaving uncertainty and skepticism about the practice and thus about using HCESs to assess dietary intake at either a household or an individual level (Murphy, Carriquiry, and Ruel 2012; Fiedler, Carletto, and Dupriez 2012). More comparative analyses are warranted in order to better understand the sources and the nature of these differences. It is hoped that we can learn both the level of confidence that we might have in using HCESs for particular applications, nutrition programs, country characteristics, or target populations, and how we might improve HCESs to strengthen their use in nutrition policy making and programming.

The Materializing Global Agenda: Now or Never?

**Preview**

Given the number of recent and ongoing initiatives to strengthen development statistics and promote the use of empirical data to shape policies and programs, this is a propitious time for pursuing a global effort to improve the collection of food consumption data in HCESs by making such collection part of this larger strengthening effort, with which it has much in common.

Although there are international guidelines and recommendations for the design and implementation of each of the specific types of surveys within the HCES “family” (ILO 2003; Grosh and Glewwe 2000; UN 1984, 1989), they are specific to each type of survey and generally leave considerable flexibility to the implementing countries. Over time, the lack of adequate, uniform guidelines has contributed to the proliferation of variability in HCES questionnaire design, data collection methods, and data processing practices (Carletto, Zezza, and Banerjee 2013; De Weerdt et al. 2014; Serajuddin et al. 2015).

This is not meant to convey the impression that all of the changes in HCES design that have occurred over time have been negative. Whereas most HCESs used to collect data only on the value of food purchases, today, roughly 80 percent of HCESs also collect information about the quantity of food purchased (Fiedler, Lividini, et al. 2012). More countries are now distinguishing food consumption from food purchases and from food received in kind (Kaara and Ramasawmy 2008; Martirosova 2008). Also, although two decades ago few HCESs collected information on food consumed away from home, now 90 percent of countries collect it (Smith, Dupriez, and Troubat 2014). How well they do so, however, remains an important issue. Finally, although most HCESs still contain a mixture of food consumption and purchase data, there is a clear trend of increasing the number of countries focusing more closely and more exclusively on consumption.

The idiosyncratic nature of HCESs also reflects the absence of a body of conceptual, theoretical, and empirical literature on household survey design issues related specifically to food and nutrition data.
collection and use, including the lack of empirical estimates of the trade-offs involved in alternative survey design options.

Fifteen years ago, the World Bank published *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study* (Grosh and Glewwe 2000). This work remains the single most important effort to distill guidelines for implementing HCESs in low- and middle-income countries. In their chapter on consumption, Deaton and Grosh bemoaned what they viewed as the inadequate attention that had been paid to food consumption metrics and what they regarded as a missed opportunity for improving specifically the LSMS variety of HCES:

The LSMS program of surveys has contributed little to the methodology of measuring consumption. Systematic experiments have been conducted only recently and are still largely not written up in places easily available to the community of survey researchers. The emphasis on collecting data rather than on furthering methodology may be inevitable given that the countries and country departments (within the World Bank) that fund the surveys are more interested in increasing their understanding of development policies than in increasing survey knowhow. Nonetheless, it is lamentable that the LSMS program has not done more in this area and desirable that it should do more. Good survey practices include the continual evaluation of methods, just as good social policy making includes evaluation of the impact of government programs. The LSMS program has been remiss by not more strongly supporting investigations into survey methodology (2000, 101).

This monumental work pulled together disparate studies that were based primarily on theory, experience, and a handful of natural experiments. The editors’ introduction to this three-volume work offered this explanation:

The formal literature is scattered across disciplines—statistics, economics, sociology, psychology—and often contained in conference proceedings or government document series that are not widely indexed and are seldom available outside the country where they were written. An additional limitation is that a considerable amount of the formal literature pertains to surveys in industrialized countries. While much can be learned from such literature, it is still unclear how well the literature applies to settings with lower literacy rates, different income levels and employment and consumption patterns, and differing factors that affect the social interaction of the interview (Grosh and Glewwe 2000, 1:6).

Household survey work in the developing world has been mostly piecemeal and haphazard, with two noteworthy exceptions being this work by Grosh and Glewwe (2000), and the little-publicized, high-quality, and ongoing work of India’s long-standing Expert Group on National Survey Sample (Mahalanobis and Sen 1954; India, NSSO 2000, 2002, 2007, 2014; NSSO Expert Group 2003; Deaton and Kozel 2005). Given the conspicuous absence of a readily identifiable or readily accessible body of literature on household survey methodology, it is hardly surprising that there are highly diverse practices. The multipurpose nature of these surveys means they are subject to the demands of a wide variety of stakeholders with heterogeneous interests, information needs, and goals. With little methodological guidance, many countries have responded to the changing composition of stakeholders and their demands in an ad hoc manner, without assessing (either a priori or ex post) the implications of the changes, and thus without fully understanding them.

As a result, many countries have found themselves in a difficult, spiraling situation: there are no global guidelines because HCESs are so diverse that developing such guidelines is a difficult task that nobody has been willing to take on. At the same time, however, HCESs have become so diverse, in part, because there exist no global guidelines. In addition, most low- and middle-income countries’ national statistical organizations—that is, the HCES-implementing agencies—have experienced low levels of funding over the past two decades. This confluence of factors has “trapped many countries in a vicious
cycle of statistical underdevelopment and underperformance which manifest themselves in many ways” (Kiregyera 2013, 2).

This situation is not unique to HCESs: it is part of a broader problem with national statistics for development that has attracted increasing attention in recent years (UN and FAO 2010; World Bank and FAO 2011; Jerven 2013; Jerven and Johnston 2015) and has spawned growing recognition of the need for a global strategy for improving statistics. This trend is encouraging. It makes for an enabling environment for introducing modifications in HCESs, suggesting that the timing is good and the prospects for addressing the common roots of these problems are promising. To date, much of the focus has been on agricultural and rural statistics (UN and FAO 2010; World Bank and FAO 2011). Although it is clear that a great deal of overlap exists between food issues and agriculture issues (Carletto et al. 2015), there remains a long-standing and increasingly pressing need to directly address food consumption issues per se (Herforth and Hoberg 2014; Carletto, Zezza, and Banerjee 2013; Fiedler, Carletto, and Dupriez 2012) and to develop some minimal standardization and guidance to aid countries in extricating themselves from this destructive spiral. Clearly, international agencies have a role to play—indeed, a responsibility to fulfill—in helping to bring together a global partnership to stem the spiral by addressing this agenda and helping to retool countries’ HCESs to better deal with the changing nature of malnutrition.

Since 2009, however, this situation has started to change. That year, the World Bank and the Bill & Melinda Gates Foundation launched the Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS-ISA) project. The primary objective of the project was to encourage innovation and efficiency in statistical research on the interplay between agriculture and poverty reduction (World Bank 2016). The project targeted eight countries in Africa south of the Sahara and was focused foremost on improving agricultural production and yield data, but it also included particularly vexing topics that are of interest to all parties—members of the nutrition community and others—concerned about improving food consumption metrics, in particular, quantifying (in either monetary units or nutrients) agricultural production for home consumption and standardizing the units of measurement.

The nutrition community’s increasing recognition that the long-standing food consumption information gap has left it with an inadequate understanding of the changing nature of malnutrition has sparked renewed interest in food consumption metrics, leading to the formation of organizations such as the Inter-agency Working Group on Rural and Agricultural Statistics (IAWG-RAS) (IAWG-RAS 2014) and the International Network for Food and Obesity / Non-communicable Diseases Research, Monitoring and Action Support (INFORMAS) (INFORMAS 2015; Swinburn et al. 2013; Vandevijvere et al. 2013), and to an increase in the segment of the nutrition community working with HCESs. Over the past six years, recognition of the common agenda that the nutrition community and the household survey community share has grown slowly but steadily, and the two communities have increasingly tried to work together to identify and address that agenda as well as strengthen HCESs’ collection of food consumption data (A2Z 2011; World Bank 2016; India, NSSO 2014; Swinburn et al. 2013; Vandevijvere et al. 2013; Farfan, Genoni, and Vakis 2014, 2015; IAWG-RAS 2014; Statistics Indonesia and World Bank 2014; Jolliffe 2014; IHSN 2014, 2015; Carletto et al. 2015; SPRING 2015).

The IAWG-RAS has four goals: (1) to address some of these HCES weaknesses in the collection and use of data, (2) to demonstrate how to use some of the data to inform policy, (3) to improve the capacity of national statistical offices to collect and use the data to improve policy, and (4) to make HCES data more relevant and encourage their use among a larger audience of researchers and policy makers (Carletto et al. 2015). The potential of repurposing can best be achieved by harmonizing HCESs on key food and nutrition issues and metrics, so as to minimize threats to their relevance and reliability, and improve them in ways that benefit all users and stakeholders (for example, through information on poverty, prices, and food quantities, and through standardized units of measure).5

5 In light of the heterogeneity of HCESs and the fact that they are implemented by countries at virtually all levels of development, it may prove useful to identify a subset of countries or country prototypes for which particular types of changes are especially pertinent, or to prioritize countries that are particularly poor or have the most alarming nutrition problems. Such a prioritization would be similar to the LSMS-ISA’s focus on eight countries in Africa south of the Sahara: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda.
Several of these organizations are now making a concerted effort to coordinate their work, prioritize issues, and accelerate the pace of this endeavor. The approach adopted in this literature review—building on the 100-country HCES review (Smith, Dupriez, and Troubat 2014), which was itself jointly sponsored by the International Household Survey Network (IHSN), the World Bank, and the Food and Agriculture Organization of the United Nations (FAO)—exemplifies this coordination. The shared goal is to produce enough additional evidence of “what works” to catalyze the development of a consensus about minimal acceptable standards for key survey parameters and to produce a set of global guidelines that will, it is hoped, come to carry the imprimatur of the United Nations Statistical Commission.

These efforts have been spearheaded by the World Bank, with substantial support and participation from the FAO and IHSN. They have generated a growing body of methodological and applied studies (Jerven and Johnson 2015; Carletto et al. 2015; Carletto et al. 2016; World Bank 2015), including the 100-country HCES review (Smith, Dupriez, and Troubat 2014) and the development and publication in 2014 of the ADePT–Food Security Module (ADePT-FSM), an open-access, user-friendly software program intended to facilitate and standardize the analysis of HCESs. ADePT-FSM combines HCES data with food composition table data to produce a wide range of food consumption and nutrition-related indicators at national and subnational levels (Moltedo et al. 2014). With the recent release of ADePT-FSM Version 2, it is becoming apparent that the use of HCESs in food security and nutrition-related applications is becoming both easier and more commonplace. The FAO’s database (FAOSTAT) now includes HCES-based analyses of 43 countries (FAO 2015). In December 2015, the World Bank hosted an eight-country workshop with the West African Economic and Monetary Union in Ouagadougou, Burkina Faso. The purpose of the meeting was to review a harmonized HCES questionnaire that the World Bank had developed.

Further underscoring the propitious timing of this initiative and its promising potential is the confluence of a variety of other disparate forces:

- The continuing digital revolution, which has brought ever-increasing ability to process ever-larger quantities of data at ever-lower costs, encouraging growth in the number of HCESs (Serajuddin et al. 2015)
- The growing demand for better understanding about food prices and food systems, first triggered by the 2008 food crisis, which fomented unrest in several countries
- The coming of age of IHSN (IHSN 2015), the National Household Survey Capability Program, MECOVI, and other global and regional initiatives that have promoted increased availability of, access to, and quality of survey data, and strengthened national capacities to conduct household surveys (UN Economic and Social Council, Statistical Commission 2014)
- The growing demand for household- and individual-level socioeconomic data, driven by increased demand from citizens, governments, and international donors for accountability and evidence of program and policy effectiveness, as best exemplified by the MDGs (Over 2014)
- The increased need for indicators of food consumption and diet quality stemming from increased attention to nutrition-sensitive interventions—especially in agriculture (A4NH 2014; European Union et al. 2014; Herforth and Hoberg 2014; Herforth 2015)
- The awarding of the 2015 Nobel Prize in Economics to the dean of household survey analysis, Angus Deaton, demonstrating ongoing global recognition of the importance of these surveys in helping to guide social policy

---

6 MECOVI is the acronym for Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida (Improvement of Surveys and the Measurement of Living Conditions), a Latin America and Caribbean regional program.
7 Household surveys (but not exclusively HCESs) are the source for 27 of the MDGs (UN Economic and Social Council, Statistical Commission 2015).
Both the growing initiatives to reify and improve national statistics for development and the rapidly growing use of HCESs in food and nutrition analyses underscore the need for, and have added voice to, the call for a global effort to better harmonize HCESs and to improve them in fulfilling their repurposed role as a source of food and nutrition information. How should this work proceed?

**Key Types of Household Consumption and Expenditure Survey Shortcomings and Limitations: Survey Design Measurement Errors and Processing Errors**

**Preview**

A strategic global agenda that will benefit all HCES stakeholders consists of addressing six key shortcomings that stem from common survey design or processing errors.

HCES data are subject to a number of different errors that undermine their relevance and reliability in developing measures of food consumption, nutrient intakes, and nutrition status. Improving HCESs’ measurement of food consumption requires understanding the key causes of common errors and the significance of these errors so as to enable prioritizing and obviating them altogether, or minimizing them and ameliorating their impacts. These potential errors occur at different steps in the HCES cycle of design, planning, data collection (fieldwork), data cleaning, and analysis.

For purposes of designing a work plan for undertaking studies to strengthen HCESs, it is useful to consider two dimensions of the most common HCES errors: their sources and the nature of the expertise required to address them. The two general sources of HCES shortcomings are survey design–related measurement errors and data processing errors. Survey design–related measurement errors are data collection and reporting errors that may stem from respondent errors, interviewer errors, or problems attributable to the questionnaire. Data processing errors occur after data collection and encompass what is done with the data, that is, its cleaning and analysis (Biemer et al. 1991; Carroll, Crossley, and Sabelhaus 2015). Given the sequential nature of these activities, adherence to the adage “first things first” would suggest that the initial focus should be on attenuating the data collection and reporting errors stemming from the questionnaire. That is also the approach that was taken by Smith, Dupriez, and Troubat (2014) in their recent review—referred to throughout this paper as “the 100-country HCES review”—which focused exclusively on questionnaire shortcomings.

Table 2.1 identifies types and sources of common HCES shortcomings, including those identified in the 100-country HCES review’s eight key recommendations for strengthening HCES collection of food consumption data (Smith, Dupriez, and Troubat 2014). The table includes shortcomings that stem from the processing and analysis of topical areas that are particularly important and problematic for food policy and nutrition analysts. These are priorities for food policy and nutrition analysts and are thus the specific applications of interest to this subset of HCES users. It is noteworthy that, in general, the especially challenging topics are those associated with using HCESs to develop nutrient intake estimates. While there is no one-to-one mapping of the 100-country review’s key recommendations to food and nutrition analysts’ priority topics (owing to overlapping categories and differences in the perceived importance of particular subtopics), three of the eight key recommendations of the 100-country HCES review are topics that foremost involve expertise that is not unique to food and nutrition specialists and that might best be led by other disciplines. These are the collection of data on food from each potential source individually (purchases, own production, and gifts/in-kind), the collection of data enabling measurement of quantities, and data collection over an entire year to account for seasonality.

---

8 This discussion is not comprehensive: it focuses on only those types of errors that are thought to have particularly important implications for the measurement and use of food consumption data from HCESs.

9 While nutritionists have specific expertise in estimating the quantities of food consumed (in particular in their use of food models in 24-hour recall surveys), we regard their potential contribution to this issue as proportionately less than their potential contribution to the other priority topics identified here. This reflects our belief that a quantitatively more important and equally challenging aspect of the issue of quantifying households’ food consumption involves purchases and consumption from own
### Table 2.1 Potential types and sources of error in HCESs and those targeted by the unfinished agenda’s six priority studies and the 100-country HCES review’s eight key recommendations

<table>
<thead>
<tr>
<th>Type of Error</th>
<th>Source of Error</th>
<th>Unfinished Agenda’s Priority Studies</th>
<th>8 Key Recommendations of 100 Country HCES Review (Smith et al., 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire Design Errors</td>
<td>Questionnaire</td>
<td>#1</td>
<td>(1) Collect data on food given to non-household members</td>
</tr>
<tr>
<td>a. No data collected on individual level consumption (only household level)</td>
<td></td>
<td></td>
<td>(2) Improve capturing of FAFH</td>
</tr>
<tr>
<td>b. No data on number of meal partakers</td>
<td></td>
<td></td>
<td>(3) Modify food list to improve specificity to enable identifying food groups and capturing prepared dishes/processed foods consumed away from home</td>
</tr>
<tr>
<td>c. Inadequately capturing FAFH</td>
<td></td>
<td>#3 &amp; #5</td>
<td></td>
</tr>
<tr>
<td>d. Data only collected on acquisition and not consumption</td>
<td></td>
<td></td>
<td>(3) Modify food list to improve specificity to enable identifying food groups and capturing prepared dishes/processed foods consumed away from home</td>
</tr>
<tr>
<td>e. Not including all 3 food sources: purchases, home-produced, received in-kind</td>
<td></td>
<td></td>
<td>(4) Collect data individually on each source</td>
</tr>
<tr>
<td>f. No data collected on food stocks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Pertinent only for HCES-collecting purchase data)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Food list: Not comprehensive, lacks specificity; doesn’t include all food groups</td>
<td></td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>(5) Collect data individually on each source</td>
</tr>
<tr>
<td>g. Recall error due to recall period (too long or too short)</td>
<td></td>
<td>#2</td>
<td>(8) Recall Period maximum: 14 days</td>
</tr>
<tr>
<td>Reporting Error in Consumption</td>
<td>Data Collection</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>(6) Collect data enabling calculation of metric quantities</td>
</tr>
<tr>
<td>a. Recall error (Recall Period)</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>#2</td>
<td>(7) Nutrient intake estimation requires metric quantities and improved FAFH</td>
</tr>
<tr>
<td>b. Telescoping error (Recall Period)</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>#2</td>
<td></td>
</tr>
<tr>
<td>c. Rounding error (Recall Period, Food perishability)</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>#2</td>
<td></td>
</tr>
<tr>
<td>d. Cognitive error (e.g., in construction of 'usual month' estimate)</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>#2</td>
<td></td>
</tr>
<tr>
<td>e. Interviewer error</td>
<td>&amp; Qx (Length of Recall Period)</td>
<td>#2 &amp; #3</td>
<td></td>
</tr>
<tr>
<td>IV. Unit Conversion Error (Mismeasurement of Quantity)</td>
<td>Questionnaire</td>
<td>#6</td>
<td></td>
</tr>
<tr>
<td>v. Nutrient Intake Imputation</td>
<td>QUESTIONNAIRE</td>
<td>Processing</td>
<td></td>
</tr>
<tr>
<td>a. Precision of the match of the HCES food item with Food Composition Table</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Food items without quantities (i.e., with only expenditure data)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Processed, composite foods with unknown food and nutrient contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Meals received from other HHs/provided to non-household members</td>
<td></td>
<td></td>
<td>(Could be used to adjust apparent intake estimates)</td>
</tr>
<tr>
<td>(Shortcoming: There is no food or nutrient content data provided about 'meals')</td>
<td></td>
<td>&amp; Qx (Length of Recall Period)</td>
<td></td>
</tr>
<tr>
<td>e. Intra-household distribution assumed (use of adult consumption equivalent)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sources of Error in Estimating Nutrient Requirements, Due to Not Taking into Account</td>
<td>Questionnaire</td>
<td>Processing</td>
<td></td>
</tr>
<tr>
<td>a. Metabolism</td>
<td></td>
<td></td>
<td>(9) Served data collection over entire year</td>
</tr>
<tr>
<td>b. Weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Height</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Activity level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Breastfeeding status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Pregnancy status</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Biemer et al. (1991); De Weerdt et al. (2014); Smith, Dupriez, and Troubat (2014).

Note: The unfinished agenda’s key study topics are as follows: (1) Assess adult male consumption equivalent, (2) optimize the recall period, (3) get the food list “right,” (4) improve nutrient imputation methods, (5) capture food away from home, and (6) adjust for meal participation. FAFH = food away from home; HCES = household consumption and expenditure survey; Qx = questionnaire.

production—areas in which agricultural economists and other current HCES stakeholders have greater expertise and in which they are likely to be more interested. Accordingly, we anticipate they will insist on being involved in and probably on leading any initiatives to fine-tune or reform these methods. Therefore, although we recognize that nutritionists have relevant skills to contribute to any such efforts, because they have relatively more important contributions to make in leading other significant activities, we elected not to include this topic as a priority for the nutrition community.
Box 2.1 offers a more detailed discussion. This leaves five areas for food and nutrition specialists to focus on, which—together with a study assessing the appropriateness of the adult male consumption equivalent—we regard as the unfinished agenda’s six priority studies. The remainder of this paper is structured around these six sets of topics, specifically as follows:

- Assessing the adult male consumption equivalent as a proxy for intrahousehold distribution of food
- Optimizing the recall period
- Modifying the food list to better capture important food sources of essential nutrients
- Improving nutrient imputation methods for foods with only expenditure (not quantity) data
- Better capturing the consumption of food away from home
- Adjusting food consumption estimates for meal participation

Box 2.1 The subset of the 100-country review: Identified key challenges that we defer to the non–nutrition analysts

**1. Unstandardized quantity units.** One of the most important data elements collected in HCESs are the quantities of individual foods consumed or purchased. These measures constitute the foundation from which a large and diverse set of welfare indicators are constructed. The 100-country HCES review (Smith, Dupriez, and Troubat) found the following:

The lack of familiarity of respondents with standard units of measurement is one of the main challenges in accurately estimating food quantities. There are several approaches to solving the issue in survey practice, but a lack of hard evidence and clear guidelines on what method works best…. The most common method in survey practice is to rely on the respondents’ own report of quantities in standard units (2014, 31).

Most HCESs collect data on the volume or weight of food using a substantial number of different reporting units. Many of these units are standardized and universal—such as a grams, ounces, or liters. Others, however, may be regional or local measures (for example, Kenya’s *pakaacha* or *debe*), and still others may have common names but may not be standardized in terms of how they are measured (such as a “piece,” “heap,” “bunch,” or “stack”). This is a general problem plaguing most HCES users and stakeholders, both those interested foremost in their nutrition-related contents and those interested in other components (such as price). Most of the purposes for which HCES food data are collected—including poverty analysis, national accounts, and the construction of price indexes—would benefit from improved measurement of household food acquisition/consumption, which would simultaneously improve the accuracy of the food and nutrition–related indicators they use that are based on these measures.

This topic has also been identified as a priority issue to be addressed by the Inter-agency Working Group and the Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS-ISA) initiative in Africa. Carletto, Jolliffe, and Banerjee (2015) presented some lessons learned from recent LSMS-ISA surveys and suggested elements of a shared agenda and some systematic approaches for standardization going forward.

**2. Ensuring comprehensive collection of food data from each of the three potential food sources: purchased, consumed from own production, and received in kind.** The types of food data collected in HCESs vary from country to country. A recent paper analyzing variations in the design of 77 HCESs found that 32 of them (41 percent) collected exclusively consumption data, 28 (37 percent) collected only purchase data, and 17 (22 percent) collected a combination of purchase and consumption data (Conforti et al. 2015). Regardless of the types of food data collected, HCESs also collect information on the source of each food type and potentially identify three different sources: purchases, food consumed from the household’s own production, and food received as in-kind
Box 2.1 Continued

purchases, food consumed from the household’s own production, and food received as in-kind transfers (including wages paid in kind, social assistance provided in kind, and gifts). To best ensure that the food data being collected are comprehensive, it is essential that the survey inquire about the amount of each food item that is obtained from each of these common sources. According to the 100-country HCES review (Smith, Dupriez, and Troubat 2014), all HCESs collect purchase data, 96 percent collect data on food from own production, 86 percent collect data on food received in kind, and 85 percent collect data from all three sources; yet this is still one of the review’s eight key recommendations, reflecting the fact that countries that do not ask about all three sources are likely to have enormously underreported food consumption.

3. Seasonality and the temporal and spatial dimensions of HCES fieldwork. Moreover, a recent study of Mongolia’s HCES found evidence of cyclical variation in food consumption levels during a month, as well (Troubat and Grünberger 2016). The study found systematic variation in reported purchases and acquisition by week of the month, which the authors attributed to the timing of when most people are paid. While this “first-of-the-month effect”—that is, the finding that food expenditures vary by the timing of income payments—has been found in other studies as well (for example, Damon, King, and Leitbag 2013), it is likely to be context specific—that is, of questionable generalizability. United States–based studies have found that low-income households are particularly likely to follow a monthly cyclical pattern and have demonstrated that it results in both a reduction in expenditures and a reduction in food energy intake at the end of the month (Wilde and Ranney 2000; Stephens Jr. 2006; Hastings and Washington 2010). Even if such findings are of limited generalizability, they are important for sensitizing practitioners to the need to be aware of this possible source of measurement error so that it may be eschewed or attenuated.

Although seasonality may be attenuated by conducting the interviewing over the entire year (and perhaps at different times of month), the impact of variations in season may also vary across geographic areas, necessitating implementation of the survey throughout the country and concomitantly throughout the year, in order to capture both its potential spatial and temporal dimensions. The Bangladesh Household Income and Expenditure Survey interview sample, for example, is based on 34 primary sampling units (PSUs). The year is divided into 18 terms, and each PSU is visited 18 times during the yearlong survey.

Source: Authors.
3. STUDY TOPIC 1: ASSESSING THE ADULT MALE CONSUMPTION EQUIVALENT AS A PROXY FOR INTRAHOUSEHOLD DISTRIBUTION

Preview

Using household-level HCESs to make inferences about individual consumption requires assumptions about how the household’s food is distributed among its members. It is commonly assumed that household food distribution is in direct proportion to the individual household member’s share of the household’s total adult male consumption equivalents (AMEs). This section explains what the AME is and reviews empirical assessments of how well it proxies the intrahousehold distribution of food and selected essential nutrients.

As noted earlier, one of the most important and universal shortcomings of HCESs for conducting nutrition-related analysis is that they collect only household-level data. To make inferences about individual consumption, therefore, it is necessary to make assumptions about the intrahousehold distribution of food (apparently) consumed by the household. Three common approaches are taken to account for the number of persons consuming the household’s food. Most commonly, a simple per capita measure is calculated by dividing the food quantity by the number of household members (a method that assumes all food is equally distributed across individuals in the household). Another approach has been an offshoot of work of primarily economists analyzing the demand for food (which has been extensively studied and will not be reviewed here—see, for example, Deaton, Ruiz-Castillo, and Thomas 1989; Deaton 1997; Deaton and Paxson 1998; Gibson and Kim 2007; Rogers, Coates, and Blau 2012; Blow, Lechene, and Levell 2015). This approach does not focus specifically on the intrahousehold distribution issue but uses econometric techniques to investigate how households’ demand for different food items varies by characteristics of the household’s size and composition.

The third approach originated in a distinct but related body of work, in which Chesher (1997, 1998) sought to estimate the availability of energy, fat, calcium, and vitamin C from food purchase data in the United Kingdom. In developing his methodology, Chesher noted that the variations in household size and composition and in the incidence of purchasing for food stocks created a threat of heteroscedasticity; he therefore sought to add controls for the age and gender composition of households to attenuate the threat. He noted that high-order polynomials should theoretically be able to capture the highly nonlinear relationship between age and nutrient intakes and the distinct gender requirements for nutrients but, in practice, had not proven able to provide satisfactory results. His approach was to estimate a series of parametric and nonparametric age-gender splines, which could then be summed to provide an estimate for the entire household.

In four studies, members of the Data Food Networking (DAFNE) group proposed modifications on Chesher’s basic approach and applied it to individualize HCES data to enable comparison with individual nutrition survey data (Vasdekis and Trichopoulou 2000; Vasdekis, Stylianou, and Naska 2001; Naska, Paterakis, et al. 2001; Naska, Vasdekis, and Trichopoulou 2001). That work drew on data from Belgium, Greece, Norway, and the United Kingdom, and analyzed 14 food groups. After adjusting the HCES-based estimates for (1) weight changes during cooking, (2) edible proportion coefficients (to estimate the portion of the weight of a purchased food that was consumed), and (3) a 10 percent reduction factor to capture waste, spoilage, or food given to pets, the authors’ estimated Pearson correlation coefficients indicated that for 13 of the 14 food groups, the individualized HCES provided statistically significant “good” or “very good” estimates of the individual nutrition survey data estimates. The only poorly performing food group was fish and seafood. The correlation coefficients ranged from 0.42 to 0.96 and averaged 0.67. Naska, Vasdekis, and Trichopoulou gave the following warning:

---

10 These studies were the antecedents to, or among the work done by, DAFNE 15 years ago (discussed in Section 1) in its early efforts to develop a standardized approach to tracking changing dietary patterns across European countries with which to study nutrition epidemiology.
It is tempting to try to compare correlation coefficients between various food groups. However, confidence intervals for correlation coefficients relying on 24 points are generally large and significance levels refer only to differences from the null value of no correlation. Instead, the emphasis should be on the general overall satisfactory pattern of positive correlation between the (two survey types’) estimates (2001, 1164).

Yet another approach is that mentioned earlier in this paper in the discussion of why there have been so few direct comparisons of HCES and 24-hour recall (24HR) results. This approach, one that is becoming increasingly common, assumes that food is distributed to household members in direct proportion to each member’s share of the household’s total energy (kcal) requirements. This approach generally uses the Food and Agriculture Organization of the United Nations (FAO) concept of the AME to identify age- and sex-adjusted energy requirements, while usually assuming a moderate physical activity level and an average basal metabolic rate (FAO 2014). Claro and colleagues juxtaposed empirical estimates of per capita and adult consumption–equivalent measures for Brazil and found the following:

The adult-equivalent calorie availability levels were higher than the per capita levels, with the largest differences in rural and low-income households. Differences in household calorie availability varied from 22 kcal/day (households with adults and an adolescent) to 428 kcal/day (households with elderly individuals), thus showing that per capita measurements can underestimate the real calorie availability, since they overlook differences in household composition (2010, 2188).

All HCESs collect data on the household’s size and demographic composition, which can be used to construct AMEs. The growing share of HCES applications using the AME that have gone beyond the traditional and most common use of HCESs in analyzing nutrition status—namely, household-level analysis of food security—to undertake individual-level analyses (Fiedler et al. 2011; Bermudez et al. 2012; Lividini et al. 2013; Rambeloson et al. 2012; Fiedler et al. 2013; Engle-Stone et al. 2014; Fiedler et al. 2014; Fiedler et al. 2015; Coates et al. 2016) have modified the use of the AME and extended it in two ways: (1) to proxy the intrahousehold allocation of food (thereby enabling individual level–based nutrition analyses) and (2) going beyond energy and applying the same methodology to micronutrients. These extensions generally assume that the only criterion used to allocate food to household members is the food’s energy content; other nutrients are not taken into consideration.

Given the uncertainties about the accuracy of HCES data, it is not surprising that some observers have railed against the use of the AME. Others have urged eschewing the AME issue and the household versus individual debate altogether in favor of analyzing (apparent) consumption only at the household level. This is how the FAO has generally analyzed food security since 2008, when it first began using the HCES to estimate food security and nutrient adequacy (Sibrián, Seevalingum, and Mernies 2008). Three recent studies, however, based on uncommon databases that proffer an opportunity to more closely assess the reasonableness of the AME assumption, provide some “push-back” on simply abandoning the individual-cum-AME-based approach. The three studies, which all collected consumption data for all household members using 24HR analysis of two Bangladesh household surveys and one Ethiopia survey, found the AME assumption to be a good proxy (Coates et al. 2016). In an analysis of the Bangladesh Integrated Household Survey (BIHS), Sununtnasuk and Fiedler (2016) directly compared average estimates of intakes, intake gaps, and the prevalence of inadequacies as estimated by 24HR with those calculated using the AME approach for each of five nutrients: energy, vitamin A, iron, zinc, and calcium. At the population level, for iron, vitamin A, and calcium, more than 98 percent of all individuals had the same estimated prevalence of adequate or inadequate nutrient intakes, whether using 24HR or the AME-based estimates. The magnitude of inadequacies (the nutrient gaps) differed by less than four percentage points.

Two studies have adopted a different approach to using the AME to proxy intrahousehold food distribution. Instead of using the FAO standardized metric, these studies analyzed 24HR data and empirically estimated the consumption levels of different household members relative to that of adult males (Arsenault et al. 2014; Moursi et al. 2012). These studies have found substantially less concordance between AME-based estimates of individualized household consumption and 24HR-based estimates than those using the FAO’s AME-based approach.
points for all five nutrients. Disaggregating intakes and inadequacies by age group revealed that the vast majority of variance between the two methods was highly concentrated in the first few years of life. Of children less than three years old, 48 percent had at least one pair of inconsistent prevalence estimates. While there were important differences in the levels of estimated energy and nutrient intakes using the 24HR and the AME-based estimates for children under three, the results are remarkably comparable for the rest of the population.

To put this finding in proper perspective, it is important to recognize that 24HR surveys, too, generally do not do a good job of estimating the dietary intakes of much of this same subpopulation because of the difficulty of measuring breastfeeding. This was the reason why, for instance, a recent well-funded national 24HR survey in Uganda did not analyze the data for 6- to 23-month-olds despite the fact that they constituted nearly half of the entire sample of children 6–59 months of age for whom data was collected (Harvey, Rambeloson, and Dary 2010).

In another AME assessment, Coates and colleagues (2016) compared average intakes in Bangladesh and Ethiopia. They found in both countries that “the energy AME–predicted intakes fell within 10 percentage points of the individually reported AME for the majority of the groups across energy, protein, animal-source protein, and iron, suggesting that the AME-based predictions perform reasonably well overall” (Coates et al. 2016, n.d.). Like Sununtnasuk and Fiedler (2016), however, they too found that for young children (less than two years old in this case), the AME did a poor job, substantially overestimating the intakes of this important group. Coates and colleagues issued a caveat that is pertinent to both studies: “the question remains whether the results would hold when information on individual intake was compared with household-level consumption imputed from expenditure data of the kind collected in HCES” (2016, n.d.). It is also important to recognize that the external validity of these studies is uncertain and questionable given the small samples drawn from specific subpopulations—for example, focusing only on children 24 to 59 months of age and their mothers (Harvey, Rambeloson, and Dary 2010), or including only households in a geographically limited part of a country—for example, two districts in Bangladesh (Arsenault et al. 2010).

A third study that also encourages “push-back” to those decrying efforts to go beyond household-level analyses is a recent paper by D’Souza and Tandon (2016), which also analyzed the BIHS. These authors found that 28 percent of household members were miscategorized when their individual consumption of energy was compared with their household-level–based categorization. In food-secure households, 24 percent of household members were found to have inadequate energy intake, and in food-insecure households, 35 percent of members were found to have adequate energy intake because of the way in which food is distributed in the household. By implication, current food security measures that rely on household-level data may not be adequate: they are vulnerable to miscategorizing a significant proportion of households and an even larger proportion of the members of households. Although this study found that the AME did not provide an acceptable method by which to proxy intrahousehold distribution, it underscored the urgency of the need to get beyond household-level measures.12

**Takeaway Messages**

Studies have generally found that the AME provides a reasonably close approximation of how households distribute food and nutrients among their members. The studies, however, have been few in number, nonexperimental (and thus subject to threats to the validity of their conclusions), and of uncertain external validity. Moreover, when estimates have been found to diverge, the discrepancies generally have been among the most vulnerable populations, especially young children. More studies that are designed to avoid the cited shortcomings are warranted.

12 The inconsistency of these findings with those of Sununtnasuk and Fiedler (2016) were due to the two studies’ very different definitions of the AME. D’Souza and Tandon (2016) used a government of India–specific measure that has a higher level of required calories and much less variation by age, whereas the other study used the FAO’s standardized measure (see Sununtnasuk and Fiedler 2016 for more details). It is essential that researchers using AMEs explicitly explain how the AMEs have been constructed. We recommend adopting the FAO definition of AME as the standard so as to avoid this type of confusion.
4. STUDY TOPIC 2: OPTIMIZING THE RECALL PERIOD

Preview

This section documents the enormous variation in HCES recall periods and discusses determinants of recall, recall-related measurement error, and the relationship between the length of the recall period and the capturing of usual intakes of nutrients. It discusses two different ways in which people formulate their responses—related to the length of the recall period—and notes their distinct implications for measurement error.

Recall periods are one of the most important parameters in an HCES, yet there is no obvious or commonly agreed-upon recall period. In part, this is because the choice is likely to be conditioned by other characteristics of the survey—such as whether data are collected using a diary or an interview, whether or not there are multiple visits, and so on, as well as other contextual factors—such as the diversity of the diet, shopping routines, the rate of literacy, the degree of urbanization, and the level of income. Given the long history of HCESs—with many low- and middle-income countries having now conducted several rounds of them—it is striking how few empirical investigations of the optimal recall period exist. Surprisingly, it has never been a topic that has drawn wide and sustained attention.

HCES recall periods vary from one day to one year, resulting in a variety of recall biases (Grosh and Glewwe 2000; Deaton and Grosh 2000; J. Gibson 2005). The 100-country HCES review (Smith, Dupriez, and Troubat 2014) found an enormous variety of recall periods, varying from one day (in the case of a daily diary) to 360 days, with a mode of 7 days. Moreover, it identified 33 countries that used multiple recall periods within the same survey. Of those, 17 had different recall periods for different foods, although they did not report whether there were systematic ways in which the length of the recall period varied by type of food. The other 16 countries’ HCESs asked households about their consumption or acquisition (or both) of all foods during each of two or more recall periods. The 100-country HCES review (Smith, Dupriez, and Troubat 2014) recommended that the recall period for food consumed at home be no more than 14 days (independent of country- and food-specific characteristics), and found that roughly one-third of HCESs have recall periods longer than this recommended level.

Recall Period and Measurement Errors

Time and Memory

Two different potential sources of memory-related measurement error are associated with the recall period. The first is error of omission due to simply forgetting something. Although more of a concern with longer recall periods, it is an issue with any recall-based survey. It results in underestimating consumption or purchases. The second is telescoping, which occurs most commonly when recall periods are short and events that occurred earlier in time, before the recall period, are remembered as having occurred during it. Telescoping results in the overestimation of consumption or purchases.

How quickly does memory fade, and how is this fading likely to affect HCES data? Experimental and cognitive psychologists have studied memory extensively and tell us that short-term and intermediate memory both decay exponentially (Sudman and Bradburn 1973). Relatively little empirical work has been done on the length of recall periods, but what exists is consistent with this finding and suggests that shorter recall periods are better. An experiment conducted in Ghana more than 20 years ago examined different recall periods ranging from 1 to 14 days for 13 frequently purchased items. It found that for each additional day added to the recall period, reported spending on the items fell by 2.9 percent, and reporting reached its maximum decline of 20 percent after two weeks (Scott and Amenuvegbe 1991). In a study in Papua New Guinea, Gibson (2002) found that seven daily diaries of consumption yielded estimates of

While the 100-country review (Smith, Dupriez, and Troubat 2014) recommended that the recall period for food consumed at home be no more than 14 days, it proffered no evidence or other justification for having suggested this particular threshold.
average weekly food expenditures that were 26 percent greater than a single weekly expenditure. A more recent study in Ireland that randomly assigned respondents to three different recall periods found that relative to weekly food expenditures, monthly expenditures were 20 percent underreported and annual expenditures were 58 percent underreported (Comerford, Delaney, and Harmon 2009).

**Box 4.1 Bounded recall**

Telescoping can be reduced by *bounded recall*. Although single-interview methods for bounding recall have been developed (Sudman, Finn, and Lannom 1984), the way in which bounded recall is most commonly implemented requires two interviews. In the first interview, respondents are asked about household expenditures during a particular recall period, but the responses are not used to estimate consumption. Rather, they are used only to provide a process that requires respondents to establish the date of an event. During the second interview, the pre-established date serves as a reminder of when the recall period started, facilitating a more accurate memory of when other events happened in relation to that date, and thus reducing telescoping (Deaton and Grosh 2000). Silberstein (1990) demonstrated how the significance of telescoping bias is reduced by bounded recall, showing that the consumption rates from the first interview (which are discarded) are substantially higher than those from the second interview.

A number of HCESs use multiple visits to collect data (for example, the 2012/2013 Ghana Living Standards Survey, as well as surveys in Mongolia, Bangladesh, and all eight of the African countries participating in the Living Standards Measurement Study–Integrated Surveys on Agriculture panel surveys). In some instances, the multiple visits are designed to reduce the duration of each interviewing session by collecting data on different behaviors on different occasions. In other cases, they are intended to capture more data on the same behavior over a longer period of time while reducing recall error due to memory lapse. Depending upon how the data are used, the latter practice provides an opportunity to analyze the impact of bounded recall by comparing the responses from the first visit with those from subsequent visits. We are not aware of any studies that use such data to conduct this type of analysis.

**Salience and Memory**

Time is not the only determinant of how well we recall a behavior or an event. Our ability to recall is also a function of the significance or salience of the behavior or event (Silberstein and Jacob 1989; Bradburn, Sudman, and Wansink 2004). We are more likely to remember events that (1) are unusual, (2) have relatively greater economic or social costs or benefits, and (3) have more continuing consequences than other events. In general, the purchasing and consumption of food items is not likely to be long remembered. It is a common, routine event, not an unusual one. It does not have great economic or social costs or benefits, and it generally does not have consequences of lasting duration. This suggests that the length of the recall period for food should be relatively short.14 But just how short, and when is it too short?

A legitimate question, especially for countries considering shortening long recall periods (for example, 30 days or more), is, which is better, 7 or 14 days? Countries considering making such a fundamentally important modification to their HCESs will want to get it “right” the first time. The Indian National Survey Sample Organization (NSSO) has conducted several experiments to try to sort out this question. More than half a century ago, it set the preferred recall period at 30 days following a small-scale study in which it was reported that a 7-day recall (7DR) resulted in reported consumption of a list of

14 Another consideration for HCESs collecting purchase data is that the length of the recall period should take into account the bias that might be introduced by income level because wealthier households can purchase larger quantities at a time, less frequently. As a result, the likelihood of missing a purchase by wealthier households is relatively greater.
staples that was too high (Mahalanobis and Sen 1954). For years thereafter, however, the 30-day recall (30DR) was periodically questioned as being too long and thus resulting in an underestimation of consumption.

Eventually another experiment was introduced, in 1993/1994. The NSSO continued to use its standard 30DR period—called Schedule 1—and also introduced Schedule 2, a 7DR period that was used for a subset of items while the 30DR was retained for others. A random sample of households was assigned to each schedule. Households using the shorter recall period reported food expenditures that were 30 percent higher than those of households using 30DR. The results were disparaged because the sample size of these rounds was smaller and was not specifically designed to provide enough power to accurately estimate expenditures. Further complicating the matter, the instability of some of the estimates, coupled with the fact that the 7DR resulted in poverty estimates that were considerably lower than previous estimates, raised suspicions about the acceptability of the results. They were set aside.

Then in 1999/2000, in the 55th round of India’s HCES—which was one of quinquennial “thick rounds,” when a much larger sample is used—the experiment was revisited. Unfortunately, however, the design of the experiment was modified so that rather than having two independent samples, each being assigned a separate schedule, all households were assigned both the 7DR and the 30DR schedules. The 7DR results were found, again, to be higher than the 30DR results, but now only 6 percent higher, and the results were again not accepted. Critics charged that the flawed study design contaminated the experiment by sensitizing households—that is, asking the same households about both recall periods encouraged them to make their responses consistent with each other. The issue remained unresolved, and India thereafter returned to using only 30DR, until very recently. We will return to this story below to provide an update.

In another, more recent experiment—this one an extraordinary eight-armed undertaking in Tanzania led by the World Bank’s Living Standards Measurement Study team, referred to as the Survey of Household Welfare and Labour in Tanzania (SHWALITA)—Beegle and colleagues (2012) posited a single-day diary to be the closest available measure of “true consumption.” The authors found 7DR-generated quantities just 2 percent lower than those of the benchmark, whereas a 14DR provided estimates that were 17 percent lower than those of the benchmark. The authors concluded that the 7DR was the preferred approach.

What can we conclude about recall periods? Given the limited number of studies, the age of many of them, and the risk that their findings may be of out of date or of, at best, questionable relevance, however, one must be cautious about making definitive conclusions. Although it seems clear that lengthy recall periods—that is, those of more than 14 days—are suspect, not much more can be said with any degree of confidence. Clearly, the recall period in general, and in particular differentiating the pros and cons of 7- versus 14-day recall periods, are areas that require more empirical work.

Moreover, the “best” recall period is likely to vary by a variety of contextual factors—including the characteristics of the food and the characteristics of the respondents, as has been found in a number of studies (Friedman 2016; Deaton and Grosh 2000; Tarozzi 2007). To better understand the exact nature of the trade-offs involved in using different recall periods, we need a better understanding of dietary patterns. We need, for instance, to identify the key food sources of each essential micronutrient and to understand more about those foods. We need to know how concentrated the micronutrient intakes from each of the key food sources are, the nutrient density of these key foods, and details about their shelf life and seasonal availability, as well as what their common purchasing patterns are and how much consumption levels vary. Developing a better understanding of these issues will allow us to determine the trade-offs involved in longer and shorter recall periods.
Takeaway Messages

The ideal recall period balances the attractions of a shorter period (greater accuracy due to less memory lapse, less respondent fatigue, less survey time, and lower survey costs) with those of a longer period (greater accuracy due to less telescoping, greater likelihood of capturing usual intakes). While more empirical evidence of the trade-offs involved in striking these balances and a better understanding of the determinants of these trade-offs is needed, the recall period should be no more than 14 days.

Recall Period, Episodic Consumption, and Estimating “Usual Intake”

The food that people eat generally varies from day to day, resulting in what is referred to as “episodic consumption” or “episodic intakes.” When usual intakes are measured over a single day or a short recall period, measures of a central tendency (for example, the mean or median) are not affected by within-subject variability, but estimates of the prevalence of inadequate intakes are affected. The magnitude of these distortions depends on the within-subject variation in intake. As a result, relative to a multi-day survey, a one-day survey is likely to generate both overestimates of the prevalence of inadequate intakes and overestimates of the prevalence of excessive intakes, as shown in Figure 4.1.

Figure 4.1 Estimates of the prevalence of inadequate and excessive intakes in the same population

![Figure 4.1 Estimates of the prevalence of inadequate and excessive intakes in the same population](image)


Day-to-day intake is also likely to vary by nutrient. Some nutrients that are found in high concentrations in only a few foods that are less frequently consumed—such as vitamin A and vitamin B12—are subject to greater within-subject variation and require a longer recall period to precisely estimate usual intakes. Nutritionists have found that the length of time required to establish the “usual intake” varies substantially by both nutrient and context, ranging from a few days to several weeks or more (Beaton et al. 1983; Nyambose, Koski, and Tucker 2002; R. Gibson 2005, 132; Alemayehu, Abebe, and Gibson 2010; Presse et al. 2011; Chun and Davis 2012). Presse and colleagues (2011), for instance, found that a minimum of 6 days was needed to assess usual vitamin K intake. Nelson and colleagues (1989) analyzed data from 18 studies reporting a total of 29 nutrients, and found that most nutrients required more than 7 days. Nyambose, Koski, and Tucker reported that in a sample of Malawian women “to estimate true individual intakes within an error range of +/-20 percent required 8–23 days for energy, protein, carbohydrates, and fiber, and 95–213 days for micronutrients” (2002, 1313) (Table 4.1).
Analysts may also want to modify the length of the recall period depending on the specific objective(s) of the nutrition analysis. If, for instance, the focus is nutrients that are not stored in the body or not stored for a long period of time (for example, vitamin C, as opposed to those such as vitamin A, which can be stored in the liver for up to six months), then it is desirable to have a longer recall period to enable greater precision in estimating nutrient intakes of the foods containing the target nutrient. Other things being equal, if HCESs are to be used to analyze micronutrient intakes, in theory a relatively longer recall period is likely to be desirable because it may be able to capture more information and enable better estimates of the usual intakes of more nutrients. However, a longer recall period will not necessarily capture more precise information due to the increased likelihood of forgetting or because with longer recall periods people are more apt to rely more on episodic enumeration (Box 4.2).

Because an adequate human diet requires continuously adequate amounts of at least 51 nutrients (Graham et al. 2007), it should be readily evident that one size will not fit all: one recall period will not meet the objectives of all studies. Even where HCESs are found to be adequately precise and reliable for some nutrients, they may not be for others or in other settings. Where would be a good point to start sorting some of this out in designing a particular study or specifying the best recall period for a particular objective or set of objectives? Reviewing any 24-hour recall (24HR) data available in a particular study country might provide some useful insights as to how much longer a recall period might be necessary to better capture usual intakes. Analyzing existing HCESs could help, too. For instance, earlier HCESs might be analyzed to identify the key food sources of the essential nutrients of interest and to examine variations by seasonality of consumption or by food source or means of acquisition (purchase, own production, gift / in kind). In addition, HCESs with multiple recall periods could be used to compare inter- and intrahousehold variations in consumption by food source or means of acquisition.

### Table 4.1 Number of replications required per individual for 95 percent of observed values to lie within a specified percentage of the true mean in a sample of 184 Malawian women

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy, MJ</td>
<td>10</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Carbohydrates, g</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Fiber, g</td>
<td>21</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Protein, g</td>
<td>23</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Fat, g</td>
<td>188</td>
<td>84</td>
<td>47</td>
</tr>
<tr>
<td>Iron, mg</td>
<td>65</td>
<td>29</td>
<td>16</td>
</tr>
<tr>
<td>Calcium, mg</td>
<td>112</td>
<td>50</td>
<td>28</td>
</tr>
<tr>
<td>Zinc, mg</td>
<td>94</td>
<td>42</td>
<td>24</td>
</tr>
<tr>
<td>Vitamin C, mg</td>
<td>152</td>
<td>68</td>
<td>38</td>
</tr>
<tr>
<td>Vitamin A, µg</td>
<td>213</td>
<td>95</td>
<td>53</td>
</tr>
<tr>
<td>Folate, µg</td>
<td>98</td>
<td>44</td>
<td>25</td>
</tr>
<tr>
<td>Vitamin B-12, µg</td>
<td>274</td>
<td>121</td>
<td>48</td>
</tr>
</tbody>
</table>

Source: Nyamboso et al. (2002, Table 6).
Box 4.2 How interviewees formulate responses: Episodic enumeration versus estimation

When people are asked about a past behavior or event, the common, traditional perception is that the behavior or event is first recalled and then counted. This is referred to as “episodic enumeration.” Psychologists have found that this is the usual cognitive process for many behaviors or events, but when people are asked about more than five behaviors or events, they are more likely to start relying on estimation methods, and they are especially likely to do so for events that are common and recurring (Blair and Burton 1987; Bradburn, Sudman, and Wansink 2004). People are also more likely to use estimation when the behavior is routine and of low salience, such as in the case of food consumption. Thus, despite their low salience, foods—and especially frequently consumed foods, such as staples—are likely to have low rates of measurement error because respondents are likely to use sound estimation methods.

Recognizing that people may use estimation methods rather than episodic enumeration in developing their responses also suggests that longer recall periods do not necessarily result in more data. When longer recall periods are used in order to capture less frequent behaviors (such as eating less frequently consumed foods), they are likely to encourage the use of estimation rather than episodic enumeration and result in more recall error.

Experiment-based work of Menon and associates (Menon 1993; Menon, Raghubir, and Shwartz 1995; Menon and Yorkston 2000; Raghubir and Menon 2005) demonstrated that the way in which respondents make use of estimation techniques requires that the behaviors be similar and regular, or it is likely to result in frequency reports that are underestimated. In the case of infrequent purchases, this work suggests that the use of estimation rather than episodic enumeration to formulate responses will result in underestimates of purchases (or consumption) relative to routine purchases (or consumption).

Source: Authors.

In repurposing HCESs to collect more relevant and reliable food consumption data on individuals, we take note of a prescient observation made by George Beaton and colleagues more than 30 years ago:

It is essential that in the design of nutritional studies the intended use of dietary data be considered, the requisite reliability of the data and the precise nature of the information needed be determined, and then the decision be taken on the method of dietary data collection and the requisite number of replications. It is now clear that there is no ideal method of collecting dietary information. Rather, there may be “preferred” methods for particular uses (1983, 995).

Takeaway Messages

Accurately capturing the usual intake of different nutrients requires different recall periods. Thus, following the guidelines presented in Table 4.1, the ideal recall period depends on what are regarded as the most important nutrients and the most significant nutrition issues, together with the perceived level of accuracy required of the data for the particular analyses or applications.

A Closer Look at Recall Periods and Specifically the Household Consumption and Expenditure Surveys That Collect Purchase Data

Preview

Some HCESs collect exclusively consumption data, while others collect exclusively purchase data and still others a combination of the two. Repurposing purchase data to proxy for consumption generally entails assuming that the food is consumed during the recall period. This section identifies common measurement errors stemming from this assumption and discusses methods for attenuating some of these distortions.
Rates of Consumption and Food Stocks

Most HCESs were originally designed to collect data on food acquired for consumption and not on food consumed (UN 1989). Today roughly 41 percent of HCESs collect data only on purchases, 33 percent collect data on both purchases and consumption, and the residual 26 percent collect only food consumption data (Smith, Dupriez, and Troubat 2014). The collection of purchase data and that of consumption data involve very different cognitive processes and entail different burdens on the respondent—which for most HCESs is generally a single member of the household. Remembering purchases—which in many instances will have been made primarily by the single respondent—is likely to be substantially easier than remembering consumption. In contrast, remembering consumption requires being aware of and remembering the behavior of all household members, as well as being aware of and remembering many more events (since eating food will generally occur more frequently than will purchasing it). It is likely, therefore, other things being equal, that HCESs that report only purchase data will be more accurate and more reliable than those that collect either a mixture of purchase and consumption data, or exclusively consumption data.

However, surveys that collect only purchase data, or a combination of purchase and consumption data, have other intrinsic measurement problems, especially for those who wish to use them to measure food consumption. The use of HCES purchase data as a proxy for consumption entails making important assumptions that are not required in using HCESs that contain only consumption data. The purchase data collected are most accurately described as quantifying food that is “available” as opposed to consumed. Availability, however, is a static concept: it is independent of time and does not enable us to say much about food security or nutrition. In repurposing HCES purchase data to provide a proxy measure for consumption, it is essential to talk about a rate of “apparent consumption.” In using HCES data to proxy consumption, it is generally assumed that the food that is purchased (or consumed from own production or received in kind) during the recall period is consumed during that period. In essence, the recall period is used to define the household’s “usual consumption” period, and the nutrients that are consumed during the recall period (when analyzed in combination with an appropriate food composition table) are used to define “usual intakes.” Although it is obvious that some foods purchased prior to the start of the recall period are consumed during the recall period, and that some foods purchased during the recall period are consumed after the interview (the end of the recall period), HCESs generally do not take food stocks into consideration. Exceedingly few HCESs collect data on food stocks or food inventories, or provide other data with which to calculate food stocks (Box 4.3). As a result, HCESs collecting data on purchases are subject to errors owing to their not accounting for households’ adding to or drawing down food stocks during the recall period.15

In addition, the use of purchase data as a proxy for consumption is likely to result in overestimations of apparent consumption because no account is made for waste, food given to animals, or food gifted to non–household members.16 While the likelihood and extent of underreporting in consumption vis-à-vis purchase data due to differences in recall does need to be kept in mind, given the various additional sources of distortions that may be present in purchase data, we speculate that HCESs collecting purchase data are likely to be less precise than those collecting only consumption data, and are likely to provide less accurate estimates of consumption. In essence, the surveys collecting only consumption data require less extensive repurposing than those collecting only purchase data or a mix of consumption and purchase data. Recognition of the likely potential differences in the accuracy of purchase and consumption data suggests that the degree of confidence in the analyses of purchase-based

---

15 The Nigeria General Household Survey 2012/2013 asked a series of five questions about each food item that enables distinguishing the portion of each consumed item that was consumed from purchases in the past seven days from the portion that came from purchases before the past seven days—the latter constituting the drawdown of existing inventories. To our knowledge, the inventory component of consumption has never been isolated and analyzed.

16 Consumption data may result in overestimates of physical consumption because they may include plate waste. Given that plate waste is only a subset of the potential sources of waste in purchase data, and that purchase data are also subject to plate waste, we conclude that HCESs collecting only consumption data are more accurate at capturing consumption data than those HCESs that collect only purchase data or a combination of purchase and consumption data.
HCES data’s findings and recommendations is less than that in the analyses of consumption-based HCESs. It also suggests that in efforts to distill lessons and recommendations in cross-country analyses of HCESs, the analyses should be stratified by the specific type of data collected.

Box 4.3 Purchase data, food stocks, and recall period

To our knowledge, there have been only two HCES studies analyzing food stocks, Gibson and Kim’s (2012) paper analyzing a 1996 Papua New Guinea survey, and Troubat and Grünberger’s (2016) study of Mongolia. Noting that even when there are nonzero reported purchases, consumption may be understated because food may also have been taken from stocks, Gibson and Kim used an infrequent purchase model to estimate these two potential types of hidden consumption. They found that for each food item, an average of 6 percent of households reported no acquisition of the item during the recall period, at the same time that they reported having consumed some of it from stocks. They also found that, on average, for each food item another 22 percent of households reported consumption levels that were greater than acquisition levels, due to their having consumed from stocks. The authors estimated that consumption from food stocks represented nearly 20 percent of all food consumed. Recognizing the practical obstacles to conducting a similar survey in other countries (with two home visits, to measure starting and ending food stocks), the authors proposed extending the recall period to increase the likelihood of acquisition’s being captured during the survey and thus of reducing the gap between acquisition and actual consumption. They further recommended adding a question about longer-term consumption (for example, any consumption in the past 12 months) to enable distinguishing between households that never consumed the item and those that consumed it only infrequently.

The Mongolia study found that when stocks were not included in the analysis the estimated average daily calories per person from acquisition were 138 kcal lower than the average daily energy from consumption. The authors noted that beginning and ending stocks did not balance out over the enumeration period. They found that this was due to stocks’ being added to at the beginning of the year and at the beginning of the month, and they recommended that the timing of survey fieldwork take this phenomenon into account (Troubat and Grünberger 2016).

Source: The authors.

Recall Period, Large Purchases, and Implausibly High Apparent Consumption

Food expenditure patterns tend not to be normally distributed: households’ food preferences vary and their food stocking practices may vary substantially. As a result, for most food items, some households will have no apparent consumption while others will have very large purchases. At the same time, a limited number of households that have no apparent consumption during a recall period may in fact truly have zero consumption during the recall period, especially if the recall period is somewhat short. Others, however—and this will include the vast majority—will have hidden consumption owing to their consuming stocks of food that were purchased prior to the recall period. One of the challenges of repurposing food purchase data to use as a proxy for consumption is how to distinguish extreme values that are true outliers due to errors in data reporting, recording, or data entry, from those that are outliers only because we are repurposing purchase data and using it as a proxy for consumption. Large quantities of food purchases—especially of staples with a long shelf life—may be entirely plausible, but their use as quantities of food apparently consumed may not be.

What should be done in these instances? Should these observations be dropped altogether? Doing so will compromise the sample weighting scheme, which will affect the point estimates as well as the confidence intervals estimated using HCESs. These extreme values should be flagged and further analyzed. Are there apparent coding errors? Can they somehow be adjusted? For example, is there a way to estimate a maximum plausible average daily per capita amount of consumption of the item? In part, it depends on the application of the analysis.
An alternative to simply dropping the outliers that has been employed involves reviewing the distribution of the same food item from a 24HR survey that was available for the same country (Lividini and Fiedler 2013), and to use the maximum consumption value identified in the 24HR to inform the development of a maximum allowable value in the HCES with which to constrain the outliers. Another method that can be used in HCESs with multiple recall periods is to investigate whether or not the outlier households reported positive purchasing during the second, longer recall period. If so, an alternative average daily rate of consumption can be calculated and substituted for the outlier value (Fiedler et al. 2013a). A third possibility is to develop multivariate demand-type estimates of the food using the HCES to predict the outlier household’s consumption, and then to substitute that estimated value for the outlying value (Smith and Subandoro 2007).

Another type of outlier is also possible: implausibly low levels of consumption. For some consumers, large purchases may result in what appears to be no consumption during a particular recall period. This may occur, for instance, due to large purchases made prior to the recall period resulting in stocks that are drawn down during the recall period. This scenario can result in consumption levels that appear to be implausibly low. (More on this topic below.)

**Recall Period and the Frequency of Purchasing: Distinguishing True Zero Consumption from Reported Nonconsumption**

In a review of a Brazilian survey, Lanjouw found many households with zero values for total food purchases, and noted, “It is unlikely that these households are, in fact, not consuming food. Rather it seems possible that many households in Brazil purchase food on a fortnightly, or even monthly, basis” (2005, 19). The recall period in this instance was too short. This example suggests that another criterion for setting the length of the recall period is that it should reflect food purchasing and consuming habits in the specific country context. As the Brazil example demonstrates, a very short recall period for purchases of a food that is not frequently purchased will capture a great deal of transitory behavior and yield estimates that, due to variability that results in high relative standard errors, are unstable and unreliable. But with different foods purchased and consumed at different rates, how does one go about setting an “optimal” recall period? While estimates based on data collected over a longer period of time are likely to capture more purchases and provide more stable estimates (as already noted), they are also subject to bias—underreporting—due to the increased likelihood of the respondent’s forgetting. Clearly, the selection of the “best” recall period is a balancing act, and how one strikes that balance is likely to be influenced by whether or not the HCES collects purchase or consumption data and by the purpose for which the study is being conducted.

As we have seen, there is a tension between, on the one hand, having a longer recall period—in order to avoid the relatively greater susceptibility of staples to telescoping errors—and, on the other hand, having a shorter recall period—so as to avoid the relatively greater susceptibility of less commonly consumed foods like fruits and vegetables (which are relatively more likely to be important sources of micronutrients) to measurement error. This prompts a question: is there a dichotomy in dietary assessment? In setting a recall period, do we need to choose between better measuring staples and thus calorie/energy intakes, or better measuring micronutrients and thus assessing the quality of the diet? If so, it would appear that the approach regarded as “best” depends on what are perceived as the most important nutrition issues to be addressed or those that need to be most precisely addressed. Box 4.4 discusses an alternative approach to having a single, fixed recall period.

---

17 The relative standard error (RSE) is the standard error of the estimate divided by the estimate itself, multiplied by 100. It is expressed as a percentage of the estimate. A common benchmark for survey estimates to be considered adequately reliable is that the RSE be less than 30 percent. While this may be a useful criterion for setting the length of the recall period for surveys with relatively small sample sizes (for example, as is the case with many 24HR surveys), it is not likely to be discriminating enough in the case of most HCESs because the average HCES sample size is more than 10,000 households and the level of precision will generally be high.
Box 4.4 “Usual week” or “usual month” as an alternative to a longer, fixed recall period

In an effort to avoid the estimation errors and other distortions that a fixed recall period may cause when it does not coincide well with a household’s usual purchasing or consumption patterns, some HCESs attempt to better capture “usual intakes” by asking about a household’s “usual” consumption pattern. Cognitive psychologists, however, have found that questions with specific times generally produce more precise estimates (Bradburn, Sudman, and Wansink 2004). As Hurd and Rohwedder explained, “It appears at first sight that using the wording ‘typical month’ rather than ‘last month’ would be a way of taking advantage of the smaller recall bias of a short reference period while encouraging the respondent to make adjustments to any unusual spending patterns of recent months…. This suggests that for the case of irregular spending patterns, asking about a ‘typical month’ is a masked way of asking about a much longer time period after all, with the only difference being that the respondent is supposed to perform some averaging to arrive at the requested answer. One would expect this requirement to affect data quality in a negative manner” (2009, 442).

In terms of the magnitude and direction of measurement error associated with “usual consumption” patterns, the few studies addressing the topic have had mixed results. Angrisani, Kapteyn, and Schuh (2015) found that significantly lower quantities were reported for “typical” than for specific reference periods. The Survey of Householders Welfare and Labour in Tanzania (SHWALITA) experiment found that the use of a usual month, relative to the benchmark of a single-day diary, resulted in greater measurement error and overestimation than did a fixed recall period of either 7 or 14 days (Beegle et al. 2012). SHWALITA also found that estimating “usual” purchases or consumption also required more interview time. A third study, done in Niger, found that the use of 7-day recall and “usual month” resulted in very similar results (Backiny-Yetna, Steele, and Djima 2014). The authors speculated that although they may be reflective of relatively monotonous diets, the results suggested that 7-day recall would be preferable, as it does not appear to be subject to telescoping and is likely to be subject to less memory lapse than a usual month, especially when its longer interview time and cost requirements are factored in. On balance, these observations suggest that reliance on a “usual” reference period is not an attractive alternative.

Source: The authors.

Purchase Data, Infrequent Purchases, and the Potential Use of Multiple Recall Periods in Estimating Usual Intake

HCESs with multiple recall periods are not uncommon. The 100-country HCES review (Smith, Dupriez, and Troubat 2014) found that 33 countries used more than one recall period. Of those, 16 used more than one recall period for all food items and 17 used different recall periods for different food groups. For example, Zambia asks about consumption and purchases in the past two weeks and the past month, and about 10 percent of countries ask additional questions to attempt to identify common purchasing patterns and “usual intake.” Bhutan has 7DR, 30DR, and last 12 months; Nepal has 7DR and “usual month”; Viet Nam asks about the number of months in the past year and the number of times per month. For every item in its food list, the Guatemalan ENCOVI (Encuesta Nacional de Condiciones de Vida) asks if any member of the household purchased the item for home consumption in the past 12 months, the number of months in which it was purchased, how much is usually spent each month on the item, and the quantity purchased in the past 15 days. Multiple recall periods proffer the opportunity to calculate multiple, distinct consumption rates. With the Guatemalan food list consisting of 116 items, however, asking as many as four additional questions about each item requires a lot of additional time and attention.

For HCESs that collect purchase data, or a mixture of purchase and consumption data, better understanding long-term purchasing behaviors can help to avoid underreporting the consumption of foods that are infrequently purchased either because they are infrequently consumed or because they have a relatively longer shelf life and (as is especially common with staples) are purchased in large quantities. A Zambia study, for instance, which found questionably low levels of sugar consumption using 14DR, turned to 30DR and found that the suspiciously low levels of consumption were overwhelmingly due to
the masking of relatively infrequent large-volume purchases (Fiedler, Lividini, et al. 2012). Whereas use of 14DR suggested that only 54 percent of Zambian households apparently consumed maize, 30DR suggested the percentage was 83 percent, more than 50 percent greater.

This approach might be useful for one-time efforts to establish purchasing patterns and usual intake, or it could be used selectively, for example, to identify food fortification vehicles. We are unaware of any other analysis of whether or not the incremental information obtained has been used or whether it is worth the time and effort, worth risking undermining the quality of the survey owing to the increased time of the interview and the increased likelihood of fatiguing the interviewee or the interviewer and thus risking the interview’s being prematurely ended. In light of the length of time it takes to complete these multipurpose surveys and their increasing rates of rejection (especially in urban areas), it is surprising that (to our knowledge) countries that use such surveys have not empirically established the best recall period and used that information to simplify the interview, shorten the time it requires, and reduce fieldwork costs. HCESs with multiple recall periods provide opportunities to learn more about recall periods and to better understand how consumption patterns of different types of foods vary by duration of recall. More secondary analyses of HCESs with multiple recall periods are needed to better understand the robustness of our food-specific consumption measures. (Appendix 2 contains a list of potential analytical studies that could provide insights for improving the accuracy, reliability, and efficiency of HCESs’ collection of food consumption data.)

Juxtaposing food consumption estimates derived from data collected from the same households reporting for different recall periods provides a useful natural experiment that enables controlling for the host of household-specific variables that might otherwise affect food demand and confound analysis of the impact of recall period length. There remains, however, a potential shortcoming of this design: differences in interviewees’ responses under the two recall periods may be smaller than if the same data were collected for each recall period independently, because respondents will have been sensitized and are likely to strive to make the two responses consistent.

There is a unique, and as yet unexploited, opportunity to learn more about recall periods and how they might vary by food source or type of food, which does not suffer from this potential source of contamination because of its unique design. It is India’s National Survey Sample questionnaire from 2012/2013, the 68th round, in which the survey introduced the use of two distinct interview schedules with different recall periods, and conducted the two surveys using each interview schedule with an independent sample. Schedule Type 1, using what until that time was the standard 30DR period for all 219 food items, was retained, and the new Schedule Type 2 introduced the use of a 7DR period for 87 of the items while retaining the 30DR period for the other 132 items. The items for which the 7DR was introduced were generally thought to be purchased with greater frequency, and the country’s expert group stated, “It is expected to be less prone to (downward) reporting biases due to recall error that are suspected to be present in recording of a household’s consumption of these items by the usual (30DR) method” (NSSO Expert Group on Non-sampling Errors 2003, 15). To our knowledge there has been no analysis of differences in the consumption estimates of individual food items (although aggregate consumption measures of both recall periods have been reported and juxtaposed in poverty assessments). This survey provides an opportunity to assess the impact of different recall periods and to do so without contamination. Still, the generalizability of the results remains an unaddressed issue.

**Takeaway Messages**

Compared with HCESs that collect only consumption data, those that collect purchase data are subject to more distortions due to the repurposing of purchase data to serve as a proxy for consumption. The magnitude of the measurement error associated with those distortions varies by purchasing patterns, which, in turn, are a function of characteristics of the foods consumed, the household’s income and storage facilities, and the number and nature of markets accessible to the household.
The many HCESs with multiple recall periods have largely untapped potential to teach us much more about the complex relationship between purchases, the length of the recall period, and how we might improve the ways in which we proxy consumption from purchases.

Recall Period and Diary-Based Data Collection: Duration, Usual Intake, and Interviewee Fatigue

Preview

This section analyzes two diary-based HCESs to examine the incremental impact of a progressively longer recall period on the trade-offs of better capture of usual intakes versus growing respondent fatigue manifested as increasing underreporting of consumption.

A recent study of Bangladesh data analyzed a series of seven 2-day diaries to investigate the impact of progressively longer reference periods on the capturing of “usual intake” estimates and the prevalence of inadequate intakes, while examining whether or not there are indications of interviewee or interviewer fatigue (Engle-Stone, Fiedler, and Sununtnasuk 2016). The study found that progressively increasing the reference period by 2-day intervals from 2 days to 14 days resulted in very small changes over the seven data collection periods in the apparent consumption of energy, iron, zinc, and calcium (less than 3 percent), and small changes in the prevalence of apparent inadequate intake at the household and individual levels. The study concluded that for purposes of estimating nutrient intake levels or the risk of inadequacy, modifying the length of the reference period between 2 and 14 days would not change the conclusions (Figure 4.2).

Figure 4.2 Prevalence of inadequate individual intakes with seven different recall periods

The authors of this review subsequently undertook additional analyses of the same database. While recognizing that it is not possible to assess individual nutrient requirements from survey data, changes in individuals’ risk of intake inadequacy were analyzed to better understand the stability of these estimates and to try to shed light on the underlying distributions. The analysis revealed that the apparent stability of the total population’s intake adequacy measure over the seven reference periods masked a considerable number of changes in the risk of inadequate intake at the individual level (Table 4.2). Mapping the percentage change against the length of the reference period showed that over time, the changes in individuals’ risk of inadequate intake fell. It also revealed that incrementally extending the reference period would gradually capture more of the changes in individuals’ probabilities of inadequate
intake than would otherwise occur over time. As shown in the last row of Table 4.2, a reference period of 8 days would capture an average of 79 percent of the total changes in individuals’ risk of apparent inadequacy across the five nutrients that would occur over an entire 14-day period; and as shown in Figure 4.3, a reference period of 10 days would capture 89 percent and one of 12 days would capture 95 percent.

Table 4.2 Changes in individuals’ intake inadequacy status by reference period, Bangladesh Household Income and Expenditure Survey, 2010

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Consistency status of prevalence estimates</th>
<th>2 days</th>
<th>4 days</th>
<th>6 days</th>
<th>8 days</th>
<th>10 days</th>
<th>12 days</th>
<th>14 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kcal</td>
<td>Never Inadequate Intake</td>
<td>41</td>
<td>37</td>
<td>35</td>
<td>33</td>
<td>32</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Always Inadequate Intake</td>
<td>59</td>
<td>55</td>
<td>54</td>
<td>52</td>
<td>52</td>
<td>51</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Changes in Status</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Zinc</td>
<td>Never Inadequate Intake</td>
<td>47</td>
<td>42</td>
<td>40</td>
<td>39</td>
<td>37</td>
<td>37</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Always Inadequate Intake</td>
<td>53</td>
<td>48</td>
<td>45</td>
<td>44</td>
<td>43</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Changes in Status</td>
<td>0</td>
<td>10</td>
<td>15</td>
<td>18</td>
<td>20</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Vitamin A</td>
<td>Never Inadequate Intake</td>
<td>25</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Always Inadequate Intake</td>
<td>75</td>
<td>70</td>
<td>67</td>
<td>65</td>
<td>64</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Changes in Status</td>
<td>0</td>
<td>13</td>
<td>19</td>
<td>22</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Calcium</td>
<td>Never Inadequate Intake</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Always Inadequate Intake</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Changes in Status</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4.2 Changes in individuals’ intake inadequacy status by reference period, Bangladesh Household Income and Expenditure Survey, 2010

Percent of all changes in status: 55% 58% 79% 86% 95% 100%

Source: Authors’ calculations.

Figure 4.3 Percentage of total changes in inadequate intake prevalence captured with increasing length of recall period, Bangladesh Household Income and Expenditure Survey, 2010

![Figure 4.3](image)

Source: Authors’ calculations.

Note: DR = day recall.
From the perspective of the total population, the 2-day diary may yield results that look identical to those of the 14-day diary. However, given the changes that we have seen occur in the underlying data that provide the bases for these population-based estimates, it is entirely conceivable that for some analyses of particular subpopulations, the estimates of inadequate intake prevalence would be affected by the length of the diary. While the 2-day option may be appealing because of its obvious savings in time, fieldwork, and resources compared with a 14-day-based survey, depending on the objective of the analysis, a longer, 10- or 12-day, diary might be preferred because it would bring greater stability to the estimates.

The Mongolian study (discussed earlier, Box 4.3), which looked at two different metrics—the average number of items consumed and average daily energy consumption—also found evidence that the length of the diary may affect estimates of usual food consumption (Troubat and Grünberger 2016). The Mongolian HCES survey collected data from three consecutive 10-day diaries and found that households reported consuming steadily decreasing numbers of food items. In addition, while their average daily energy consumption fell by 5 percent from the first to the second diary, it partially recovered in the third, leaving the average in the third period 3 percent below that of the first. This is consistent with the findings of Beegle and colleagues (2012) and confirms what Crossley and Winter noted: “Because diary fatigue seems to lead to decreasing compliance throughout the recording period, designers of diary surveys face a trade-off not unlike that faced by designers of recall surveys. Shorter recording periods will lead to less bias in the estimation of mean expenditures, but, because of infrequency, higher variance” (2015, 38).

**Takeaway Messages**

Although there is no optimal recall period, the choice of a shorter or longer recall period involves a series of trade-offs, including (1) precision versus capturing usual intake, (2) instability versus underestimation, (3) telescoping versus omission, (4) whether respondents rely on episodic enumeration (that is, counting) versus estimation algorithms in formulating their food quantity responses, and (5) having a larger number of items in the food list versus limiting interviewer and interviewee fatigue and cost. Setting the ideal recall period consists of striking the best balance between these countervailing biases, which is not an easy task. The size of the under- and overestimations that we are trying to balance are hidden because we do not know the level of “true consumption.” We know the direction of the biases, but we do not know their size.
5. STUDY TOPIC 3: MODIFYING THE FOOD LIST TO BETTER CAPTURE IMPORTANT FOOD SOURCES OF ESSENTIAL NUTRIENTS

Preview

This section describes key characteristics and uses of HCES food lists in estimating food consumption, together with their common pitfalls and limitations.

The length and composition of the food item list are important determinants of the accuracy of food consumption reporting (Grosh and Glewwe 2000; Deaton and Grosh 2000; Jolliffe 2001; J. Gibson 2005). Longer lists are presumed to result in higher reported consumption because more, and more disaggregated, food items help respondents remember their consumption more completely and accurately (Beegle et al. 2012; Carroll, Crossley, and Sabelhaus 2015; Comerford, Delaney, and Harmon 2009). While the evidence from low- and middle-income countries is limited (and much of it now quite dated), studies in El Salvador, Indonesia, Jamaica, and Ecuador show that a longer food list results in higher estimated levels of consumption, and conversely that short, highly aggregated food lists contribute to underreporting of food consumption (Jolliffe 2001; Pradhan 2001, 2009; Jamaica, Statistical Institute and Planning Institute 1994; Beegle et al. 2012). The Jamaican study (1994) found that the shortened consumption module (which included more than just food) produced estimated mean per capita consumption levels 20 percent below those of the standard module. Similarly, Jolliffe’s (2001) study of a between-groups-designed experiment in El Salvador found that going from a list of 18 food items to one of 72 items increased average per capita food consumption expenditure estimates by 20 percent.

The Indonesian National Socioeconomic Survey, or SUSENAS, makes use of an alternating design in its questionnaires. Every year a short questionnaire with only 15 food items is administered. Once every three years, one-third of the sample receives a detailed consumption module that includes 218 food items. The Pradhan studies (2001, 2009) made use of this experiment and analyzed three years of data, comparing the short and long versions of the food list in three years. The author found that the degree by which average per capita food consumption was underestimated varied by year from 3.5 percent in 1993 to 11.4 percent in 1996 to 6.5 percent in 1999. Although the author did not measure the time required to collect information for just the food lists, he did measure how much time was required for the entire short and long lists of all expenditure items (which numbered 23 and 320, respectively, roughly the same relative sizes as the food-only list, with the short list being equivalent to 7 percent of the long list). The 23 items required 52 minutes, and the 320 items required 82 minutes. Thus the longer list resulted in a much smaller than proportionate increase in administrative time relative to the increase in the number of items. The study also found that the fraction by which consumption was underestimated increased as consumption rose, and concluded that whether or not a short questionnaire might be regarded as adequate “depends on the level of economic development, which is directly related to the variety of consumption patterns” (Pradhan 2009, 415).

The Survey of Household Welfare and Labour in Tanzania (SHWALITA) experiment found that administering a 58-item food list using seven-day recall (7DR) required 49 minutes, and reducing the list by 70 percent, to 17 items, resulted in reducing total interview time by just 8 minutes (Beegle et al. 2012). Noting the substantially lower reported mean consumption levels from the 17-item list, and that the longer list tracked the gold-standard supervised one-day diary much more closely, the researchers concluded that the longer list was preferable.

Given evidence that HCESs generally underreport food consumption relative to the more precise, benchmark-providing 24-hour recall (24HR) (Dary and Rambeloson 2012; Serra-Majen et al. 2003; Moursi et al. 2012; Carroll, Crossley, and Sabelhaus 2015) and that the incremental interview time requirement of adding additional items is relatively small, other things being equal, the evidence would suggest that having a longer list is generally preferable.18

---

18 Fricker and colleagues (2015) noted that in United States–based studies, the general impression that HCES estimates are
But obviously, a food item list can also be too long, and the bulk of the analyses that found longer food lists to be associated with higher consumption had food lists that were relatively short—fewer than 100 items, less than the average of 125 food items reported in the 100-country HCES review (Smith, Dupriez, and Troubat 2014). The common finding—that longer lists are associated with higher consumption—may reflect this lower average number. Longer lists require more interview time and higher costs, and they risk increasing respondent and enumerator fatigue—which can result in increasing the likelihood of enumerators’ cutting corners (Statistics Indonesia and World Bank 2014; Finn and Ranchhod 2015) and undermining the quality of the data collected—and respondents’ prematurely terminating the interview or refusing to participate (Deaton and Grosh 2000; R. Gibson 2005).

Longer lists have also been found to result in underreporting due to respondents’ “learning to say no” in order to complete the interview more quickly (Kreuter et al. 2011; Eckman et al. 2014). Such “motivated underreporting” is more likely to occur when a questionnaire has a cascading structure, which is a fairly common characteristic of HCES food consumption modules, and it has been found to be differential across items, depending on their sequence and format (Crossley and Winter 2015). Many HCESs, for instance, have a screen for each item in the food list that asks whether any of the item has been consumed or purchased during the recall period. If the response is positive, there is a series of follow-up questions asking how much and from where the food was obtained. In an interleaved format, if the response is positive, these follow-up questions are immediately asked. In a grouped format, the follow-up questions are asked only after multiple filters have been administered. The interleaved format has been found to result in more motivated underreporting by respondents hoping to shorten the questionnaire and interviewing process (Eckman et al. 2014).

What is needed, therefore, is a balance between the shorter food list’s reduced memory lapses and reduced interview time and costs, and the longer list’s better recall and more comprehensive reporting. How to strike that balance and what the ideal trade-off entails is uncertain. The issue is not solely one of how many items to add, but also the types of items to consider adding and the possibility of dropping others. For example, to what extent is it necessary that an HCES food list include processed foods, which may require developing recipes, increasing data processing and analysis time, and introducing new sources of complexities and uncertainties? The best way to capture processed foods and composite foods is likely to depend on a host of variables, including a country’s general level of development, in particular the structure of its food system and the nature of its value chains; the proportion of food consumed that is from home production; the frequency of purchasing patterns; the average shelf life of common foods; the importance of food consumed away from home; the variability in dietary patterns and food preparation methods; and the intrahousehold variability in the practice and importance of eating outside the home.

The objective of revising food lists in repurposing HCESs is to provide more and better nutrition information, leading to more accurate estimates of the quantities of food(s) or nutrient(s) consumed while not adding to survey length or complexity, or excessively complicating survey processing. Clearly, this is an area in which nutritionists have a particularly important role to play.

How to go about revising the food list also depends, in part, on how the food consumption data are going to be used. From a nutrition perspective, the most demanding application of HCESs is in the quantification of nutrient content. If the intention is to go beyond identifying food types and food groups (dietary diversity) to estimate nutrient availability or apparent nutrient intakes as precisely as possible, researchers must be sure to include foods that are important sources of key nutrients while avoiding those that are nutritionally insignificant and add little understanding of nutrient availability or apparent intakes. Yet the latter still must be captured in a broader food item category, and the survey must still cover items dominated by underreporting is based primarily on comparisons of HCESs with national income and product account estimates. While most of the observations about underreporting that have been made in low- and middle-income countries have come from comparisons of HCESs and 24HR estimates, Fricker and colleagues’ suggestion that the magnitude and direction of errors in HCESs depend on the expenditure category analyzed, the characteristics of the respondent, the survey protocol, and the method used to assess measurement error is pertinent for analysts wanting to better understand variations in food expenditures worldwide.
that are important for continuing to fulfill the original purpose of the HCES or to meet other HCES stakeholders’ needs. Fulfilling these requirements entails concurrent consideration of the food list and the specific food composition table (FCT) that will be used to transform food quantities into nutrient quantities (FAO and INFOODS 2012b). Is there a recent, country-specific FCT available? If not, does a neighboring country with similar dietary patterns have one that could be used? It is also important to note that the FCT must be reasonably recent because the nutrient content of food varies over time, by where it is grown, and by the type and condition of the soil in which it is grown.19 If a recent FCT does not exist, one will need to be developed. Finally, is the HCES food list specific enough to enable capturing the consumption of the nutritionally most important foods and matching them unambiguously to an entry in the FCT? Box 5.1 discusses some additional food list consideration that are specific to fortification and biofortification policymakers.

**Box 5.1 Some food list–specific fortification and biofortification considerations**

<table>
<thead>
<tr>
<th>Source: Authors.</th>
</tr>
</thead>
</table>
| If one of the objectives of nutrition policy makers is to collect data with which to assess potential fortification vehicles, it will be important that the food list include processed foods that contain significant quantities of fortified food so that the coverage and additional intake of foods may be adequately measured. For example, if the fortification of wheat flour is being considered and the intention is to require fortification of all wheat flour—not only that produced and sold as a final consumer product, but also that used as an intermediate input to produce another food, which is generally the case—then it will be important to include not only wheat flour in the food list, but also those items containing significant quantities of potentially fortified wheat flour, such as white bread, whole wheat bread, pasta, cookies, crackers, pies, bakery items, and the like.

Similarly, if an objective of the food analysis is to estimate the coverage and impact of biofortified orange sweet potato, then it may be important to include not just “tubers” or not just “sweet potatoes” in the food list, but “orange sweet potatoes,” and perhaps also yellow and white sweet potatoes if they are also consumed in significant quantities. Many biofortified crops, however—such as iron-biofortified beans and zinc-biofortified wheat—have invisible traits that preclude modifying the food list to enable tracking their consumption with adequate precision. Such crops will require other monitoring and evaluation methods. |

Inadequate articulation of the food list makes it difficult to match the food item with an FCT item and introduces uncertainty (and if the matching process is not explicitly explained and standardized, becomes a likely source of unreliability in how the match is made and in estimates of nutrient content). For example, the Uganda HCES food list includes “beans.” As Table 5.1 shows, however, the Uganda FCT has 16 different combinations of bean type and preparation method with calcium estimates per 100 g that vary by a factor of 12, iron estimates that vary by a factor of 9, zinc estimates that vary by a factor of 6, and vitamin A contents varying from a low of 0.0 IUs to 209.0 IUs.

---

19 Studies to date have focused primarily on the iodine and selenium contents of soils (Chalimba et al. 2011; Johnson 2003; Watts et al. 2015), but there have also been notable discussions about iron and zinc. For instance, studies have considered the nutritional implications of the high iron content of water (in Bangladesh in particular), the high iron content of soils in Ethiopia, and the growing zinc depletion of soils in Bangladesh and India for reducing the attractiveness of zinc-biofortified crops (such as rice).
Table 5.1 Micronutrient content of different types and preparations of beans, Uganda

<table>
<thead>
<tr>
<th>Bean type</th>
<th>Calcium (mg)</th>
<th>Iron (mg)</th>
<th>Zinc (mg)</th>
<th>Vitamin A (IU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEANS,K20 (PINTO),FRESH,RAW</td>
<td>35</td>
<td>3.0</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,K20 (PINTO),FRESH,BOILED</td>
<td>32</td>
<td>2.7</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,K20 (PINTO),DRIED,BOILED</td>
<td>46</td>
<td>2.1</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,BLACK,DRIED,RAW</td>
<td>123</td>
<td>5.0</td>
<td>3.7</td>
<td>0.0</td>
</tr>
<tr>
<td>BEAN,WHOLE SEED,WHITE (NAVY),FRESH,RAW</td>
<td>73</td>
<td>3.2</td>
<td>1.1</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,WHITE,DRIED,RAW</td>
<td>240</td>
<td>10.4</td>
<td>3.7</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,WHITE,DRIED,BOILED</td>
<td>90</td>
<td>3.7</td>
<td>1.4</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,MUNG (GREEN GRAM),DRIED,RAW</td>
<td>132</td>
<td>6.7</td>
<td>2.7</td>
<td>114.0</td>
</tr>
<tr>
<td>BEAN,WHOLE SEED,CRANBERRY (KANYEBA),FRESH,RAW</td>
<td>29</td>
<td>1.1</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>BEANS,CRANBERRY (KANYEBA),DRIED,BOILED</td>
<td>50</td>
<td>2.1</td>
<td>1.1</td>
<td>0.0</td>
</tr>
<tr>
<td>BEAN,WHOLE SEED,KIDNEY,FRESH,RAW</td>
<td>39</td>
<td>2.2</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>BEAN,WHOLE SEED,KIDNEY,FRESH,BOILED</td>
<td>31</td>
<td>1.7</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,KIDNEY (ALL TYPES),DRIED,RAW</td>
<td>143</td>
<td>8.2</td>
<td>2.8</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,KIDNEY (ALL TYPES),DRIED,BOILED</td>
<td>35</td>
<td>2.2</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>BEANS,LIMA (BIGAGA),FRESH,RAW</td>
<td>34</td>
<td>3.1</td>
<td>0.8</td>
<td>(200.0)</td>
</tr>
<tr>
<td>BEANS,LIMA (BIGAGA),DRIED,RAW</td>
<td>81</td>
<td>7.5</td>
<td>2.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Variation: Highest as a Percent of Lowest: 829% 916% 622% ∞

Source: Hotz et al. (2012).
Note: Circles indicate highest nutrient content; Boxes indicate lowest nutrient content.

When the HCES food list item is too “crop-like” or “commodity-like” and its specific form or preparation method is not identified, food specificity and FCT-matching considerations suggest that, other things equal, a longer food list is preferable. A food list may contain, for instance, “corn” as an entry, without differentiating corn on the cob, green corn, dried corn, roasted corn, steamed corn, boiled corn, or other characteristics. Another example, one that is common with short food lists, is food list entries that are broad food group categories, such as “fruits” or “vegetables.” Another common problem is the use of “residual” or “bucket” categories to capture a number of less frequently consumed items that may be compositionally diverse, such as “other fruits,” “other meats,” or “other vegetables.” Quantifying the frequency with which these residual “other” categories are reported in existing HCESs would be a useful starting point for informing where further specificity might be needed.

There are several different ways in which these “other” categories might be used in combination with an FCT to estimate nutrient contents. The typical situation arises when the HCES food item potentially could be matched with a number of different FCT entries. Possible approaches include the following:

- A single food item in the FCT is chosen either because it is the most commonly consumed specific item or because all other such items are compositionally similar.
- All of the plausible matching FCT entries are used, and a simple unweighted mean of the entries’ composition is calculated because there is no additional information by which to more accurately gauge the relative frequency of consumption of the individual items.
- All of the plausible matching FCT entries are used, and a weighted mean of the entries’ composition is calculated. The weightings capture the relative importance of each FCT entry, taking into account that their eating habits and preparation methods are substantially different. Data to inform the construction of the weights could be expert opinion or individual dietary surveys. Weightings might vary by region of the country to capture regional variations in eating habits or preparation method.
These are situations in which the survey data processor is required to exercise considerable judgment. Food list matching is a gray area, where survey design and survey processing overlap and affect one another. All of these processing steps should be explicitly explained to ensure that they are as transparent, well understood, and as contestable as possible so that they may be more readily improved.

An unpublished, informal study of the same Uganda HCES data used in Table 5.1, conducted by HarvestPlus and the FAO Statistics Division, illustrates the impact of ambiguities in matching an inadequately articulated HCES food list with an FCT, as well as how other ambiguities in data cleaning and processing can result in widely differing estimates of nutrient intakes. Initial comparisons found that the variation in two sets of estimates averaged about 25 percent across the four nutrients analyzed (energy, iron, zinc, and vitamin A). The variances stemmed from several sources: different quantity estimates, different data cleaning rules, different ways for dealing with outliers, and different matches of the food list with the same FCT. When the variances attributable only to differences in matches of the food list with the FCT were isolated and calculated, the differences in the two sets of estimates ranged from 5 percent for zinc, 7 percent for energy, and 17 percent for iron to 201 percent for vitamin A. Clearly getting the food list “right” is essential for better estimation of food consumption. Therefore if HCESs are to be used to develop quantitative estimates of nutrient intakes, an important component of that challenge is how well it enables reliably matching the food list to the FCT.

The Food List Structure and Recall: Grouping and Sequencing Questions

Another aspect of getting the food list “right” concerns the way in which the list is structured in the questionnaire. Bradburn (2010) noted that grouping questions (and food types) can help to minimize the cognitive effort required for respondents to recall the requested information, and is thereby conducive to lower recall error. This methodology is intuitively appealing and is also important to help ensure that food is not double counted. Bradburn also pointed out that people tend to organize memory as event sequences that are hierarchical, and suggests “that grouping questions that tap into the same type of event sequences would aid recall. For example, grouping items that would be purchased on a shopping trip to a bakery, to a butcher shop, to a vegetable market, to a grocery store, and so on would be likely to improve the recall of similar items purchased. If certain types of expenditures are typical in distinct settings, grouping items by those settings, or at least cueing the setting in the body of the questions, may improve recall” (2010, 8). This suggests that questions about food away from home (FAFH) should be grouped and that, to the extent to which the locations where such food is consumed are identified, if there are multiple questions pertaining to a particular location, they should be grouped by that location (Bradburn, Sudman, and Blair 1979). This may be especially important for organizing and providing memory cues for FAFH of different types (particular types of common snacks versus breakfast or dinner) or in different types of settings.

Another implication of the recognition that people organize memory as hierarchical event sequences is that people must first remember purchasing or consuming a particular type of food before they can remember the quantity of the particular food they purchased or consumed. Thus, as Comerford, Delaney, and Harmon (2009) noted, the recall of expenditures (or quantities of foods purchased) is likely to be even more biased than the recall of whether the item was purchased (or consumed). Moreover, it is likely that the memory processes involved in recalling these two distinct facts are not the same (Bradburn 2010).

We would therefore expect that the data required to select a fortification vehicle, or to estimate the coverage of a fortified food, or to construct a dietary diversity index—which all require only a simple dichotomous yes/no response about purchasing or consuming one or more particular foods—would need to be retrieved from memory before any memory about the quantity of the food(s) purchased or consumed, and thus would be the more accurate of these two distinct measures.20

20 The 100-country HCES review (Smith, Dupriez, and Troubat 2014) did not acknowledge these differences and did not distinguish the relevance or reliability of these two measures, differences in their causes, or different implications for measures that used one versus the other. This oversight resulted in the review’s concluding (erroneously) that dietary diversity measures were no more reliably estimable with HCESs than measures of nutrient content. That was a major shortcoming of the review.
Just how long the food list needs to be to capture “all” important food sources and to be able to unambiguously match the food list with the FCT will depend on dietary patterns and food preparation habits in the country in question, and on the degree to which the form in which the foods are asked about in the HCES matches with that of the FCT (FAO and INFOODS 2012a). Analyzing an existing HCES (or other dietary survey data) can provide useful information for developing a first approximation. Such an analysis would quantify the frequency and relative importance of each food item as measured by its incidence of consumption (percentage of households consuming it), the amount of it consumed, and the amount and share of total nutrients it provides. How well it would do so would of course be constrained by how good the current food list is, but reviewing these data would be a simple way to start an empirically based process. Analyzing any existing 24HR or food frequency questionnaire would be helpful, as well, in order to ensure that more popular and nutritionally important foods are captured as precisely as possible. One approach that might be useful for delineating food lists in such a way as to minimize aggregation bias would be to conduct formal separability tests, which could be designed to determine how best to split important food categories in the food list, such as Salvanes and DeVoretz (1997) conducted in their demand analysis of different fish and meat.21

Indonesia’s statistical institute—which still alternates use of the two different questionnaire designs as noted earlier—is currently implementing a study that will address some of the same issues about the length of the food list and under- or overreporting bias raised by Pradhan (2001, 2009). It will conduct what is, to date, one of the most sophisticated national-level experiments that has been done concerning a food list (Statistics Indonesia and World Bank 2014; Jolliffe 2014). The study was prompted by Statistics Indonesia’s desire to empirically inform its selection of a fixed number of items in its food list. (In the past it had used three different food lists, alternated annually.) The study is an experiment with a benchmark (control) group and five design arms. Six different food consumption questionnaires will be fielded. Each will be randomly assigned to 500 households across Indonesia and will be representative at the provincial level, not just the national level. The key parameters differentiating the arms are identified in Table 5.2. The control group will have a 1DR period, intended to be the closest possible approximation to “true consumption,” and will serve as a benchmark against which the other modules will be compared. The results are due in mid-2016.

Table 5.2 Key parameters of the Indonesian National Socioeconomic Survey experiment

<table>
<thead>
<tr>
<th>Design</th>
<th>Number of food items</th>
<th>Recall period</th>
<th>Question sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark (Control)</td>
<td>229</td>
<td>1DR</td>
<td></td>
</tr>
<tr>
<td>Treatment Group 1</td>
<td>229</td>
<td>70R</td>
<td></td>
</tr>
<tr>
<td>Treatment Group 2</td>
<td>126</td>
<td>70R</td>
<td></td>
</tr>
<tr>
<td>Treatment Group 3</td>
<td>126</td>
<td>70R</td>
<td>Ask yes/no consumption of all food items before moving on to details of each food</td>
</tr>
<tr>
<td>Treatment Group 4</td>
<td>Shorter Alternative</td>
<td>70R</td>
<td></td>
</tr>
<tr>
<td>Treatment Group 5</td>
<td>Shorter Alternative</td>
<td>1DR&amp;7DR</td>
<td></td>
</tr>
</tbody>
</table>

Source: Jolliffe 2014.
Note: DR = day recall.

21 Salvanes and DeVoretz (1997) estimated demand equations for alternative aggregation levels of food items and, guided by the economic theory of substitutes and complements (that is, the relative magnitudes of cross-price elasticities), used statistical tests to determine the appropriate level of disaggregating food items. Haines, Guilkey, and Popkin (1988) showed that the determinants of the decision to consume a particular food group are often different from those of which particular food item to consume from within that group, particularly for more highly specified commodities or food groups. They warned that ignoring the two-step decision process (discussed in Section 7) risks missing the true behavioral patterns and leads to erroneous conclusions, underscoring the need to better understand the specific characteristics of foods that influence choice and the criteria by which households define substitutes.
Added Value or Added Burden: A Reliability Assessment Tool or Unnecessary Duplication?

A common new feature of many of the Living Standards Measurement Study–Integrated Surveys on Agriculture country questionnaires has been the addition of a subsection following the food list in the consumption module that asks about the consumption of any food items in the past week from each of the 11 food groups (Table 5.3). Given that the food list responses were for the same recall period (seven days) and could have been used to construct the same information, it is not clear whether the introduction of this new section, which requires respondents to categorize the foods they have reported consuming during the recall period, is intended to imply a validation test, a reliability test, or a pilot test of whether this information is adequate for a dietary diversity index, or whether it is simply redundant. The value-added of the additional questions should be empirically investigated.

Table 5.3 Example of a survey subsection on specific food items consumed in the past week

<table>
<thead>
<tr>
<th>Section 10C: AGGREGATE FOOD CONSUMPTION OVER PAST ONE WEEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Over the past 7 days, how many days did you or others in your household consume any [...]? IF NOT CONSUMED, RECORD ZERO.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>Grains and Flours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Maize grain/flour; Yam flour; Cassava flour; Rice; Millet; Guinea corn/Sorghum; Wheat flour; Bread; Other grains and flour)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Starchy Roots, Tubers, and Plantains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Cassava tuber; Gari; Sweet potato; Yam; Irish potato; Plantain; Other roots and tuber)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>Pulses, Nuts, and Seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Soya bean; Brown beans; White beans; Groundnuts; Other nuts/seeds/pulses)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>Vegetables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Onion; Garden egg/eggplant; Okra fresh and dried; Pepper; Tomato fresh and canned; Leaves/cocoyam and spinach; Other vegetables/leaves)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>Meat, Fish, and Animal Products</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Eggs; Dried/Fresh/Smoked fish (excluding fish sauce/powder); Beef; Goat meat; Pork; Mutton; Wild game; Chicken; Duck; Other meat)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F</th>
<th>Meat, Fish, and Animal Products used as condiments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fish sauce/powder; Meat sauce/powder, etc. used in small amounts on top of meals as flavour</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>G</th>
<th>Fruits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Mango; Banana; Orange/tangerine; Pineapple; Papaya; Avocado; Canned fruit; Other fruit)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H</th>
<th>Milk/Milk products</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Fresh/Powdered/Tinned milk; Yogurt; Other milk product - excluding margarine/butter or small amounts of milk for tea/coffee)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>Oil and Fats</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Palm oil; Butter; Margarine; Groundnut oil; Other oil and fat)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J</th>
<th>Sugar/Sugar Products/Honey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Sugar; Sugar cane; Honey; Jam; Other sweets and confectionary)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L</th>
<th>Spices/Condiments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Tea; Coffee/Chocolate drink/Milo; Salt; Spices; Pepper; Tomato sauce; Fish powder/sauce; Other condiment - including small amounts of milk for tea/coffee)</td>
</tr>
</tbody>
</table>

Takeaway Messages

The accuracy of HCES food consumption estimates depends on the length, specificity, and structure of the food list, the sequencing of its individual food items, and the ease with which those items can be unambiguously matched to a relevant FCT.

The food list should include roughly 100–125 items. The ideal food item list depends on dietary patterns and the distribution of nutrient content within food sources, and is likely to vary by nutrient. Existing HCES food lists (and other sources of dietary intake data) should be reviewed to identify the key food sources of the nutrients of interest and to assess the adequacy of the degree of specificity of the food item and the degree of ambiguity in its paired FCT entry.

More important food sources of nutrients merit greater specificity. In order to improve the precision of nutrient estimation, residual food categories (for example, “other fruits”) should be reviewed to determine whether they warrant unpacking to enable more granular understanding of specific food items and to improve the precision with which they can be matched to a specific FCT entry. To neutralize the adverse impacts this procedure may have on the length of the food list, less important food sources of nutrients may be collapsed into other food item categories to the extent that doing so does not undermine the work of other HCES stakeholders.
6. STUDY TOPIC 4: IMPROVING NUTRIENT IMPUTATION METHODS FOR FOODS WITH ONLY EXPENDITURE (NOT QUANTITY) DATA

Preview
Some HCES food items are reported only in values or expenditures, not quantities. These are usually composite or processed foods, which are particularly common among foods consumed away from home. This section discusses measurement errors associated with the way in which the nutrient content of these foods has traditionally been imputed, and examines alternative imputation methods.

We echo the disclaimer of Deaton and Grosh:
It should first be noted that imputation is an inherently difficult and error-ridden process. Imputation is likely to work best where there is relatively little need for it—when the economy is highly monetized but there is a relatively small amount of own-production (such as vegetable gardens) involving goods that have clear market equivalents.
Imputation works badly in economies in which a large share of transactions do not pass through the market (2000, 117).

Calorie imputation is not a new practice. It has long been used to analyze HCESs that include only expenditures (values), as well as others that collect quantity data but may include some food items for which only values (not quantities) are reported or that include composite processed food items for which there is no nutrient content information (although they may have precise quantity measures). In these instances, the approach to quantifying the nutrient content of these foods has been to follow the two-step procedure pioneered by Subramanian and Deaton (1996). The first step consists of calculating the average cost per nutrient unit for all those foods for which there is existing nutrient content data (primarily unprocessed foods, but some composite, processed foods as well). The second step consists of multiplying the expenditures on the food in question by the reciprocal of the average cost per nutrient unit to provide an estimate of the quantity of the micronutrient contained in the food. This is the current methodology that is used to estimate caloric availability (or “apparent caloric intakes”). We propose extending this approach to other nutrients and conducting sensitivity analyses of alternative imputation methodologies. Among the methods that should be examined is a set of household income-differentiated average costs per nutrient unit (for example, based on income deciles)—rather than a single national-level average cost per nutrient unit—in an effort to address the systematic bias that would otherwise distort the imputed values due to richer households’ purchasing more expensive food, making it appear as though they are consuming more than they do, and conversely for poorer households.

When food lists are short, they usually include highly aggregative commodity-like categories. Several researchers have noted that when the food list is highly aggregated and prices change, the change in nutrient availability based on expenditure imputation is likely to be overestimated (Deolailikar and Behrman 1988; Gibson and Kim 2011; Gibson 2012). This result reflects the fact that households have a tendency to replace more expensive calories with cheaper ones when prices rise and real income falls: households substitute lower-quality varieties of the same foods, resulting in lower expenditures but no change in calories (and do the opposite when prices fall). The use of expenditure data to impute nutrient content in such instances results in overestimating the impact of rising prices on reductions in the quantities of food and nutrients. This suggests that how well nutrient imputation methods may work, and what the preferred methodological approach to nutrient imputation may be, will be shaped by both how well the food list is constructed and a sound understanding of food markets and consumers’ views of food—in particular which foods consumers view as substitutes.

One aspect of becoming familiar with food markets that will be useful in determining how best to impute nutrients (and also in designing the “best” food list) will involve analyzing price variability. Greater variability in the price of a given food item is likely to compromise the accuracy with which caloric intakes are estimated from expenditure data using an estimated cost per calorie that is calculated from data captured on less processed foods, for which HCESs more commonly collect quantity data. One
way in which the price variability of a food item might be reduced is by further refining the category, splitting it into two or three categories, a method that will be particularly consequential for important foods in the diet that vary a great deal in quality. In Bangladesh, for instance, where rice accounts for 71 percent of all energy intakes, the Household Income and Expenditure Survey identifies “coarse rice,” “medium rice,” and “fine rice,” as well as “puffed rice.”

The price data collected using food lists with broader food categories that do not reflect qualitative differences within items are likely to contain substantial variability in the prices collected for a given food item owing to the heterogeneity in the quality of foods included in the category. This suggests that surveys with shorter food lists will have more aggregative items that are likely to have more intragroup variability in any given food item. One implication of this observation is that, other things being equal, a longer, more qualitatively differentiated food list is preferable. A second implication is that in applying the expenditure- (or value-) based approach in imputing the nutrient content of foods for which no quantitative data are collected, it would be wise to examine alternative methodologies that attempt to control for interhousehold variability. For instance, rather than basing the imputation on nationwide averages of processed foods’ nutrient content, additional variables capturing differences in price-affecting markets characteristics (such as urban-rural residence, state, or region) should be taken into account, and additional variables allowing for better capturing of household determinants of demand (such as household size, composition, and income) should be incorporated.

Clearly it is not feasible to capture all price-related variation in food quality by adding more individual food items to the list. Doing so would be prohibitive in cost and time. A strategic approach to dealing with this quality-price variability would consist of also being guided by taking into account relevant nutrition criteria in determining the “appropriate” extent of disaggregation of the food list. For instance, an additional criterion might be the variability in nutrient content across more disaggregated food items: where there is little variation in nutrients, there is little need for more refined food categories. It should be recognized, depending on the specific nutrient content of interest, that this consideration may result in different approaches to how best to impute nutrient content and how best to structure the food list.22 There is also substantial evidence that the nutrient content of one of the two major categories of foods for which it is necessary to impute nutrient content—food consumed away from home—is different from the nutrient content of food usually consumed at home, as will be discussed in Section 7.

Takeaway Messages

The nature and significance of imputation methods have been relatively neglected topics, at the same time that they are very likely to be of growing consequence in terms of their potential impact on the accuracy of HCES-based food consumption and nutrient intake estimates.

This topic warrants greater attention to better understand the trade-offs between the additional processing time required and the introduction of methods that allow for greater variability in the cost per unit of a nutrient by taking into account additional variables that affect food prices (such as income, rural/urban residence, and geographic area).

---

22 The significance of qualitative differences in food within a food item category was not mentioned in the 100-country HCES review (Smith, Dupriez, and Troubat 2014). The review asserts (without evidence) that nutrient content imputation works well for estimating caloric content but cannot be used to develop micronutrient estimates (due to the relatively greater variability of micronutrient content).
7. STUDY TOPIC 5: BETTER CAPTURING THE CONSUMPTION OF FOOD CONSUMED AWAY FROM HOME

Preview

The inadequate capture of food away from home (FAFH) is regarded as one of the most important threats to the validity and reliability of HCES-based estimates of food consumption. This section discusses the different ways in which FAFH is defined and the different approaches that have proliferated in recent years in the attempt to incorporate it, noting their relative strengths and weaknesses. We endorse a particular definition—food that is prepared and consumed away from home—and encourage adoption and further testing of a unifying conceptual framework consisting of individual household member–specific questions about meal eating behaviors during the recall period as a means to reduce various sources of measurement error in capturing FAFH.

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). China may be an extreme case, but it is not an exception. The importance of FAFH has grown rapidly in the past decade and is expected to continue doing so globally throughout the foreseeable future. As it does, so too does the urgency of the need to measure FAFH more precisely. Not doing so risks HCESs’ losing their relevance as an important source of data for better understanding food and nutrition issues and informing nutrition policy making and programming. It is only in the last decade that most HCESs have even begun to ask about FAFH. As the practice of eating away from home has become increasingly common worldwide, countries have come to recognize that adequately capturing food expenditures requires collecting accurate data on FAFH. The 100-country HCES review (Smith, Dupriez, and Troubat 2014) found that 90 percent of countries now collect some information about FAFH. The ways in which they do so, however, vary substantially, reflecting the fact that most of the approaches are ad hoc and unsatisfactory (Smith, Dupriez, and Troubat 2014).

Although there is little argument about the importance of FAFH, or that its prevalence and significance are increasing, there remains a paucity of knowledge about even its major characteristics. More fundamentally, there is little discussion about how it should be defined. As its title suggests, FAFH is commonly defined by where food is consumed, regardless of where it is prepared—in which case, food that is prepared at home and taken outside of the home to be eaten—at work or school, for instance—is considered FAFH. In other instances, it is defined by where food is prepared, regardless of where it is consumed. In this paper, we define FAFH to include only one of the four possible combinations of where food is prepared and consumed: we include only food that is both prepared and consumed away from home. Food prepared at home and consumed at home, food prepared at home and consumed away from home, and food prepared away and consumed at home are all captured in home consumption and expenditures data, and as such are not FAFH.

Many of the approaches that countries have adopted lack face validity, particularly in terms of their being able to comprehensively capture FAFH. Some HCESs collect information only about the level of expenditures on FAFH, while others collect expenditure data as well as quantities of specific foods or the numbers or types of meals consumed away from home. As shown in Figure 7.1, (Ethiopia 2010/2011) some have a single question asking the amount of money spent on meals in hotels and restaurants.
Figure 7.1 The diversity of household consumption and expenditure survey approaches to capturing food away from home

<table>
<thead>
<tr>
<th>Bangladesh, 2010</th>
<th>Malawi, 2010/11</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Dining out (Food outside) 160</td>
<td>82 Cooked Foods from Vendors</td>
</tr>
<tr>
<td>Meals (Rice/Briani) 181</td>
<td>83 Maize - boiled or roasted (vendor)</td>
</tr>
<tr>
<td>Fish 192</td>
<td>84 Chips (vendor)</td>
</tr>
<tr>
<td>Meat 183</td>
<td>85 Cassava - boiled (vendor)</td>
</tr>
<tr>
<td>Patties/Cake 184</td>
<td>86 Eggs - boiled (vendor)</td>
</tr>
<tr>
<td>Sandwich 165</td>
<td>87 Chicken (vendor)</td>
</tr>
<tr>
<td>Burger 166</td>
<td>88 Meat (vendor)</td>
</tr>
<tr>
<td>Hotdog 167</td>
<td>89 Fish (vendor)</td>
</tr>
<tr>
<td>Pizza 168</td>
<td>90 Mandazi, doughnut (vendor)</td>
</tr>
<tr>
<td>Samucha/Singara/Puri/Cake 189</td>
<td>91 Samosa (vendor)</td>
</tr>
<tr>
<td>Tea 191</td>
<td>92 Meal eaten at restaurant</td>
</tr>
<tr>
<td>Coffee 192</td>
<td>93 Other (specify)</td>
</tr>
<tr>
<td>Soft drinks/bottle water 193</td>
<td></td>
</tr>
<tr>
<td>Other 194</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>India, 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooked Meats</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethiopia, 2010/11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure on Hotels &amp; Restaurants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nigeria, 2012/13</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEALS PREPARED AND CONSUMED OUTSIDE THE HOME</td>
</tr>
<tr>
<td>1. In the past 7 days, did members of this household consume any of the following meals or drinks away from home?</td>
</tr>
<tr>
<td>2. How much did you or other household members pay, in total in the last 7 days for [MEAL]? If free, please estimate what it would have cost if you had to pay.</td>
</tr>
<tr>
<td>Breakfast</td>
</tr>
<tr>
<td>Full meals (e.g. rice and stew, pounded yam and egusi, etc.</td>
</tr>
<tr>
<td>Side dishes like pper soup, nkwobi, suya, etc.</td>
</tr>
<tr>
<td>Snacks such as sandwiches, biscuits, meatpies, donuts, polpof, etc.</td>
</tr>
<tr>
<td>Dairy-based beverages such as milk, yoghurt, etc.</td>
</tr>
<tr>
<td>Vegetables and roasted such as (carrot, pears, roasted corn and plantain, sugar cane)</td>
</tr>
<tr>
<td>Nonalcoholic drinks</td>
</tr>
<tr>
<td>Alcoholic drinks</td>
</tr>
<tr>
<td>Other juice</td>
</tr>
</tbody>
</table>

Some surveys (Figure 7.1) ask about expenditures in restaurants on very broad categories of food items. For example, the survey from Uganda (2010) asks about “food,” “soda,” and “beer.” Others ask about expenditures on food consumed away from home that is purchased from different sources, such as vendors and restaurants, as in the case of Malawi (2010/2011). Some ask only about the total number of meals eaten away from home, while others ask about the number and types of meals eaten (for example, breakfast, lunch, dinner, snacks). One of the more articulated HCES FAFH sections is that of the Bangladesh Household Income and Expenditure Survey (2010). It includes questions about both the number of meals and the frequency with which specific types of foods are commonly eaten away from home. The diversity of the approaches and the readily apparent inadequacy of many countries’ approaches are self-evident.

There are other sources of HCES measurement error in FAFH as well, beyond the ambiguities of definition and the limited number of questions asked about it. HCESs commonly rely on a single, key respondent to report all household members’ food consumption and expenditures. FAFH’s becoming more common has complicated the key respondent’s task of accurately reporting household consumption: the key respondent is increasingly unlikely to be aware of events that occur outside of his or her purview. Larger households, and especially those with more adults; persons living in urban areas; those with more complex lifestyles and greater physical mobility; and persons who commonly spend more than 24 consecutive hours away from home, are particularly likely to have their consumption underreported if FAFH is ignored. Additional sources of measurement problems associated with FAFH include the need to capture multiple foods from multiple places with different menus and prices, the challenge of estimating the quantities of the different types of foods contained in processed foods, and the difficulty of measuring how much food was consumed.

There is also considerable evidence that the nutrient content of FAFH is different from that of food usually consumed at home. Most comparative analyses of the energy intakes find them to be higher on average among persons who eat away from home relative to those who do not (French, Harnack, and Jeffery 2000; Bowman and Vinyard 2004; Bezerra et al. 2012; Lachat et al. 2012), and they find that the energy density of foods eaten away from home is higher than that of food eaten at home (Lachat et al. 2012). A study in China found that FAFH resulted in a more than doubling of meat consumption and reduction in fruit consumption by one-third compared with food eaten at home (Wang et al. 2011, cited in You 2014).

There is also growing evidence of substantial heterogeneity in the consumption of FAFH both within and across households (Fiedler and Yadav 2016; Fiedler 2015; Naska et al. 2015), which suggests that it may be important to find methods to impute the nutrient content of food that adjust for a household’s lifestyle (Naska et al. 2015). Among the evidence of heterogeneity across households is data from the 68th round of India’s HCES (2011/2012), showing that households reporting they had hosted a ceremony in the previous 30 days accounted for a disproportionate share of all food served to non–household members. Those that had hosted a ceremony comprised only 1.4 percent of all households, but they served 20.2 percent of all of the meals served to non–household members. Households that hosted a ceremony provided a mean of 82.4 and a median of 28.0 meals to others, compared with 5.0 and 0.0, respectively, for households that did not host a ceremony.

The same Indian survey also provided evidence of considerable variability within households. This survey, one of the few that collected individual household member–level information on meal consumption patterns, found that the 6 percent of persons who spent at least 24 hours away from home accounted for 36 percent of all meals consumed away from home (Fiedler 2015). In a recent unpublished analysis of one of the few other datasets that reports individual-level patterns of meals and FAFH (the 2011/2012 Bangladesh Integrated Household Survey), we found even more striking evidence

---

23 In light of the findings that the nutrient content of FAFH and food eaten at home are significantly different, combined with evidence of considerable heterogeneity (both within and across households) in who eats FAFH, it is likely that in countries where FAFH is relatively important, the use of the adult male equivalent will be a less accurate proxy for the intrahousehold distribution of food. This is a hypothesis that merits testing.
of both heterogeneity in the patterns of meals away from home among individuals within households and a skewed distribution of meals away from home, caused by the high concentration of those meals in a few individuals. The 718 individuals who reported missing all three meals at home in the past 24 hours because they were “currently staying away” constituted just 2.7 percent of all individuals in the survey but accounted for 79 percent of all meals missed due to “currently staying away” and 39.3 percent of all missed meals.

These findings regarding the heterogeneity of FAFH are unsettling. They raise a number of concerns. First, they underscore growing recognition of the importance of understanding which household members stayed away from home and the duration of their stay for better capturing FAFH—which is one of the key areas for which we impute the nutrient content of food. These findings also raise questions about how well the adult male equivalent is likely to proxy the intrahousehold distribution of food in households where there is considerable consumption of FAFH—especially where much of the FAFH is consumed by only individual household members. They also suggest that a potentially useful alternative approach would be to calculate the average expenditure per nutrient unit not on the basis of households with similar characteristics, or on the basis of a set of market characteristics or a market area, but instead on the basis of the individual household. More specifically, this approach would develop an estimate of the quantity of the micronutrient contained in a food for which the nutrient content is unknown by calculating each individual household’s average expenditure per unit for each nutrient from all of the foods (both food at home and FAFH) that the household reported consuming and for which there are nutrient content data, and then multiplying the individual household’s expenditures on the food in question by the reciprocal of the average cost per nutrient unit. This approach could be regarded as controlling for household “tastes,” as well as the household’s other demand-conditioning characteristics, which otherwise would confound the analysis.

A large variety of factors have been shown to be associated with eating away from home. Evidence suggests that household income, pocket money, household size, and level of urbanization are positively associated with FAFH participation and the amount of food (or level of expenditures) involved (Lachat et al. 2009; Meenakshi and Ray 1999; Mihalopoulos and Demoussis 2001). These characteristics have also been shown to increase the likelihood of and level of expenditure on consuming breakfast away from home, but to reduce the likelihood of and level of expenditure on lunch away from home, and also to affect the consumption of specific types of foods (Yen and Jones 1997; Mutlu and Gracia 2006; Meng et al. 2012). Older households and societies have been found to have lower probabilities of and expenditures on FAFH (Redman 1980; Meng et al. 2012; Liu et al. 2015). How do we begin to sort through these issues to better understand what is happening to FAFH, household consumption levels, and the quality of diets?

**Multivariate Tools and Methodological Approaches for Investigating Food Away from Home**

Multivariate techniques have long been used to study households’ decisions about food purchases and food consumption, and have figured prominently in efforts to sort out the size and importance of specific correlates of FAFH. Most HCES consumption or purchase data consist of a large number of zeros because many households report no purchases (consumption) of a particular item during the recall period—either because they do not (ever) purchase (consume) a particular item, or because they had no purchases (consumption) during the recall period. The use of ordinary least squares in such instances results in biased and inconsistent estimates.

Alternative approaches that have been adopted to estimate expenditures (consumption) have commonly included the use of a single-equation Tobit model or a two-equation approach. There have been two distinct types of two-equation approaches. The first is the infrequent purchase model (IPM), which uses a logistic or probit model to identify the determinants of whether or not spending occurs in the recall period, followed by a Tobit model for the spending level, which takes into account all households, including those with zero consumption (Deaton and Irish 1984; Blundell and Meghir 1987). The second is
the double-hurdle model (Cragg 1971), a sequential, two-step process (also known as the Heckman selection model). The first hurdle in the double-hurdle model is referred to as the participation decision. It consists of the household’s making a binary decision of whether or not to eat away from home, and it is estimated using a probit model. The second hurdle, which is conditional upon the decision to eat away from home (and is estimated using only those subjects who chose to do so), involves deciding how much to spend. This second step is referred to as the consumption level decision (Liu et al. 2015).

A variety of estimation techniques have been used to estimate the second step of the double-hurdle approach, with the choice of method and the empirical specification based on other considerations. The most important of these other considerations is whether it is thought that the participation and consumption level equations are dependent or independent, and how to take into account nonnormal errors. Generally, the Box-Cox transformation has been adopted to take into account nonnormal error distributions.

The IPM approach is designed to capture not only consumption in the presence of zero reported purchases (due to consumption of a purchase that may have been made before the recall period) but also underreporting of household consumption during the recall period when the household may have made and reported some purchases but also consumed a quantity over and above the amount purchased during the recall period, by drawing down food stocks that existed at the start of the recall period.

As noted earlier, the cognitive act of remembering involves a sequential, two-step process—one very similar to the two-step process involved in modeling the decision-making process regarding consuming FAFH—suggesting that the participation decision will be subject to less measurement error than the consumption level decision. It was also noted in the recall period discussion that the use of multiple recall periods may help reduce the number of false reports of apparent nonconsumption by capturing more infrequent purchases for consumption. That same approach could also help to improve estimates of the incidence of FAFH when some FAFH might otherwise be “hidden.”

The determinants of occasionally eating out have been found to be distinct from those of more regularly eating out (Naska et al. 2015; Liu et al. 2015; Orfanos et al. 2009, 2013). Among people for whom eating out is an occasional behavior, it is not easy to determine whether the days that individuals or households do not report eating out reflect their usual habits or not. Depending upon the frequency of “usual” eating out patterns, the length of the recall period may exert undue influence on the estimates, and if too short will result in inaccurate and unstable estimates. This suggests that to be better able to design a survey to capture FAFH there is a need to better understand the frequency and general patterns of eating FAFH. A number of studies have found the location of out-of-home consumption, for instance, to be systematically related to the probability of consuming FAFH, the frequency of out-of-home consumption, the composition of foods consumed, and nutrient intakes (Naska et al. 2015; Liu et al. 2015; Orfanos et al. 2009, 2013; Vandevijvere et al. 2009).

24 The double-hurdle approach has been shown to be a more generalized approach, which includes the Tobit as a special case (Jones and Yen 2000).

25 In the IPM approach, total consumption is estimated as the product of observed spending multiplied by the odds of a purchase’s being observed, and the odds of a purchase’s being observed is the ratio of the length of the survey period to the length of the purchase period (Gibson and Kim 2011). (For instance, if a household usually buys an item once every four weeks and the recall period is one week, then the odds of a purchase’s being observed are 1/4 = 0.25.) As the recall period increases, the level of spending reported increases and eventually comes to equal the amount of consumption.

26 It should also be noted that IPMs can be helpful for better capturing usual intakes, as has been demonstrated (although the methods have been developed in disciplines other than economics, and go by a different name, referred to as the “multiple source method”) (Hubrock et al. 2011. Souverein et al. 2011).

27 In light of the findings that the nutrient content of FAFH and food eaten at home are significantly different, combined with evidence that there is considerable heterogeneity (both within and across households) in who eats FAFH, it is likely that in countries where FAFH is relatively important, the use of the adult male equivalent will be a less accurate proxy for the intrahousehold distribution of food. This is a hypothesis that merits testing.
Another way to develop a better understanding of eating FAFH is to ask about the meal consumption patterns of individual household members. For example, India’s HCES questionnaire includes a seven-question section that asks about the number and type of meals consumed by each individual household member (Figure 7.2). The responses to these questions provide the wherewithal for developing a detailed understanding of the composition of meals away from home (MAFH) and total meal consumption patterns of each household member. In the Indian approach, the place of FAFH is not identified, but distinct types of meals are, and these data may provide similar insights into systematic differences in individuals’ and households’ FAFH, which in turn can be helpful in devising ways to better capture FAFH.28

Figure 7.2 The Indian household consumption and expenditure survey questionnaire’s sections on meals and meal participation

<table>
<thead>
<tr>
<th>[4] demographic and other particulars of household members</th>
</tr>
</thead>
<tbody>
<tr>
<td>srl. no.</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>(1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>item</th>
<th>code</th>
<th>consumption out of home produce</th>
<th>total consumption</th>
<th>source</th>
<th>value (Rs.)</th>
<th>value (Rs.)</th>
<th>value (Rs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cooked meals purchased (no.)</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cooked meals received free in workplace (no.)</td>
<td>281</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cooked meals received as assistance (no.)</td>
<td>282</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cooked snacks purchased [samosa, puri, paratha, burger, chowmein, idli, dosa, vada, chops, pakoras, pa bhaji, etc.]</td>
<td>283</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other served processed food**</td>
<td>284</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>served processed food: sub-total (280-284)</td>
<td>289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: India, NSSO 2014.
Note: The instructions to field staff defined a “ceremony” as “an occasion on which a large number of meals (not just snacks) were served to non–household members, which significantly affects the household’s expenditure during the last 30 days. The occasion need not be religious” (India, NSSO 2014, sect. 3.3.14). HCES = household consumption and expenditure survey.

---

28 This is not to suggest that the data collected on MAFH have never been analyzed before. They have been analyzed in several papers (Minhas 1991; Subramanian and Deaton 1996; Tandon and Landes 2011).
This type of approach could also help to provide a mechanism for better ensuring the consistency and comprehensiveness of food consumption and meal reporting. It could, for example, help to avoid oversights such as the one that was recently uncovered in the analysis of food security in Brazil, when researchers realized that the 2008/2009 HCES (Pesquisa de Orçamentos Familiares) had failed to capture food received by children through an enormous school meals program (Borlizzi, del Grossi, and Cafiero 2016). The oversight resulted in a 10 percent underestimate of daily per capita energy availability.

Finally, particular attention should be given to the specific categories of FAFH that might be included in an HCES. While a great deal of attention has been paid to food that household members purchase and, to a lesser extent, food (or meals) that household members receive 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. 2008, 2010; Fiedler and Yadav 2016). Given its unique characteristics—food provided free and by a private individual or household—there is a risk that this category may be overlooked, resulting in systematic underestimation of consumption. 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 MAFH, and 36 percent of all persons with at least one MAFH reported having at least one hosted meal provided by another household. Clearly, understanding the sociology of MAFH in a country is important input into determining how best to capture these potentially important sources of consumption.

**Takeaway Messages**

There is a need to adopt a standardized definition of FAFH and to begin assessing the diversity of approaches that have been introduced over the past few decades as countries, unaided by international guidelines, have sought to capture this rapidly growing source of consumption.

Growing evidence of heterogeneity in meal consumption patterns within households suggests that it is increasingly important to capture meal consumption patterns by individual household member. Furthermore, there is considerable evidence of heterogeneity in the nutrient content of FAFH and food at home, suggesting that it is increasingly important to distinguish these two types of consumption.

The introduction of roughly seven individual household member–based questions about meal eating behaviors during the recall period would help deal with the heterogeneity from both of these sources and thereby help to reduce the high level of measurement error currently associated with both of them. This section can also be used to cross-check several other portions of the data, including specific household-level food list entries capturing FAFH, and to provide measures for adjusting for meal participation (further discussed in the next section). Given its potential importance to capture new data and provide a cross-check for other data, we regard the addition of these questions as a “best practice” and encourage all countries to adopt the approach on a pilot basis and to assess its contribution.
8. STUDY TOPIC 6: ADJUSTING FOOD CONSUMPTION ESTIMATES FOR MEAL PARTICIPATION

Preview

HCES data generally do not make any adjustments for household members who do not eat one or more meals or for non–household members who receive one or more meals from a household. This section discusses different types of adjustments in the quantity of food consumed that can be made in order to be better able to measure the adequacy of a household’s consumption of food. It also discusses the possibility of collecting data on the number and type of meals usually eaten and of weighting the relative importance of meal types to fine-tune the adjustment. Finally, it reviews the scant empirical evidence of the need for and significance of these adjustments.

Meal participation is at the very crux of the question of how well data about household food availability can be transformed into a proxy measure of the adequacy of food consumption and nutrient intakes. To assess the adequacy of household food availability, it is essential to know the quantity of food consumed by the household during a specific time period and the household’s food requirements. A time dimension is required in order to calculate a rate of consumption, which then enables relating the quantity of food and its nutrient content consumed (or apparently consumed) to the adequacy of the nutrients required for the persons consuming them.

In order to create a rate of consumption, the vast majority of HCES analyses of consumption assume that all of the food acquired during the recall period is consumed during that time period. Given a rate of consumption, the adequacy of the consumption of the household’s food can be divided into two issues—how much food is being consumed and who is consuming it—neither of which is directly measured by most HCESs. There are a number of ways in which these two essential pieces of information may be measured and adjusted to provide more refined composite measures of adequacy. Earlier discussions in this review noted how HCESs measure food quantities and household food availability, and identified various limitations on how those measures may be used to proxy household food consumption, including distortions (both leakages and additions that may be) due to food inventories, waste, spoilage, food given to animals, or food served to non–household members. Those shortcomings are equally applicable here, and need not be reiterated. The discussion here focuses instead on how adjustments might be made in the other component of the adequacy metric, the denominator measuring the number of household members consuming the food.

The number of household members consuming the food may be measured in a number of ways: most simplistically as the number of household members. This may be an undifferentiated per capita measure, or alternatively, it can be based on the adult male consumption equivalent (AME) and can take into account differences in the household’s composition to capture variations in the food and nutrient requirements of individual household members due to differences in, at minimum, age and gender, and if available, body size and physical activity level. All HCESs collect data on age and gender—which are the key components of the AME—but few collect data on body size or physical activity level. The individual household members’ requirements can then be aggregated to the household level and compared with the household’s average daily consumption or, if combined with an assumption about the distribution of food among its members (using, for example, the AME), the analysis can be done at the individual level.

Another possible refinement is to attempt to unpack who participates in each meal and to measure the number of household members who were at a particular meal, or the total number of persons (household members and others) who were at a particular meal (exclusive of household members not in attendance). Generally it is assumed that all household members eat “their fair share,” which implicitly assumes that all household members are present in the household throughout the recall period and that they participate in each meal served during the recall period. The possibility of the absence of some household members during the recall period or their nonparticipation in some of the household’s meals generally is not addressed, nor (to reiterate) is the presence of non–household members during one or more meals. To the extent to which individuals’ presence or absence at meals is noted, rarely is
consideration given to the relative importance of the persons’ missing the meal in terms of how much they consume on average, with commensurate adjustments in the consumption—downward for the individuals missing the meal and upward for those attending it. Nor is there generally any consideration for the type of meal missed (namely, the meal’s relative importance in terms of the proportion of the household’s total daily nutrient intakes that have been missed), such as applying the equivalent nutrition unit (ENU) adjustment as described by Weisell and Dop (2012). The ENU weights the AME value according to the percentage of meals attended. It is equal to the AME if the consumer attends all meals of the day or recall period and is less than the AME if some meals are missed. Noting that the exact relative importance of meals depends on cultural factors and on the number of meals usually eaten in a day, they report that the Food and Agriculture Organization of the United Nations (FAO) has adopted weights for the four daily meals of breakfast, lunch, afternoon snack, and dinner as 0.16, 0.43, 0.11, and 0.30, respectively. They provide plausibly common examples demonstrating how to construct the ENU and, after comparing them with AME only–based adjusted and nonadjusted estimates, conclude the following:

If family and household members and guests are not taken into consideration in the estimation of household food supply, the conclusions drawn as to its adequacy will most likely be inaccurate. One can see that the accurate calculation of energy requirements of those consuming the household meals and the correct recording of those present and absent at the various meals, including guests, strongly influence the analysis of the food intake. **Even if the data are not available for a precise calculation of these factors, an approximate estimate of energy requirements of those attending the meals will provide an added value in the analysis.** (Weisell and Dop 2012, S162; emphasis added)

While application of the AME has become quite common (in particular a more limited version, which includes only age and gender considerations), the use of ENUs has not. For most HCESs, however, making these adjustments is precluded because the requisite information—the number and types of meals household members missed and the number of meals provided to non–household members—is simply not collected. Moreover, even disregarding the political battle that introducing the numerous and onerous battery of additional questions to capture this information would entail, collecting the requisite data may not be a viable option. Most HCESs rely on a single respondent to provide information on the household’s food consumption (or purchases and consumption), and the cognitive burden this level of detail would entail is onerous and likely prohibitive because of the attendant level of recall and measurement error that would plague any such attempt. For example, to provide complete meal participation data in the average-sized household for the modal recall period requires a minimum of 210 data points (= 5 persons * 14 days * 3 meals/day). It is not reasonable to expect that a single household respondent is likely to be aware of, and able to recall, all of those data points accurately. Moreover, given the time and resource constraints of the respondents and the enumerators, attempting to collect that level of meal consumption specificity for household members alone is likely to prompt a backlash and result in sizable increases in prospective interviewee rejections or premature terminations of interviews (and then there are the additional non–household member participation considerations). Detailed person-level meal participation data are routinely available in 24-hour recall (24HR) surveys, but the nature of those surveys—person-specific, detailed, food-specific data on a single day—make for an altogether different, much more circumscribed and manageable situation.

Even if the requisite data could be collected, the resulting nutrition intake estimates would likely remain incomplete. The data would probably be less than comprehensive due to the absence of any understanding of any systematic variation in the specific content of the meals eaten and those missed. While it would be possible to apply, for instance, a measure like the ENU discussed earlier—which takes into account systematic differences in the energy content of the specific meals missed—no comparable constructs exist for the various other nutrients of interest, although they could be constructed. These adjustments could be made and would likely improve the accuracy of estimates of food consumption and nutrient intakes for households, household members, or both, but would require considerable work—in questionnaire design, enumerator training, and additional analytical steps.
Furthermore, it is not clear how the requisite individual household member–level data would be collected or whether they would be used to adjust household-level estimates or household member–level estimates. The results would likely be, at best, a relatively crude approximation, which raises additional questions: How much additional information would be good enough? Do we need the cognitively demanding detail of the specific foods, the specific meals, and the specific days of the specific recall period for each individual household member? Or might it be possible to collect, instead, additional information on individual household members’ “usual meal patterns” (not unlike the quest for estimating “usual intake”), or would it be best to tie the measure to a particular experience and to inquire about a summary of meal patterns during the recall period (using an approach, for example, such as the one India uses)?

If it is not feasible to collect all of the additional data required to develop an individual-specific meal participation accounting of household food consumption throughout the recall period, then what is the best way to proceed? Which of these possible alternative ways of addressing the meal participation information gap is worthwhile to pursue? Which would generate the most useful additional information at the least additional burden while minimizing the risk of introducing additional measurement error? What are the trade-offs involved in adding additional sets of questions and obtaining data that significantly affect estimates of the adequacy of consumption and are most likely to result in improved public health nutrition policies? We don’t know. As noted, there are no HCESs that collect all of the data necessary to make all of these adjustments. As a result, there are no HCESs with which this question can be definitively empirically answered. The ideal approach to answering this question would be a series of experiments collecting different combinations of these additional data points in different countries. In the absence of the funding that would be required to implement such an approach, the only alternative available to us is to survey existing HCESs to identify and analyze those that include questions that collect some of the more comprehensive data necessary to try to ferret out the incremental contribution of additional questions and to attempt to ascertain the trade-offs their implementation entails—that is, to quantify how different the results would be if these adjustments were not made. The nonexperimental nature of the approach renders it vulnerable to threats of rival hypotheses, but it is the best that can be done.

Given how demanding it would be to collect all of the data required to make a comprehensive set of adjustments, where are we now? What information is actually currently collected, and what adjustments could presently be made? Do they appear to potentially have much impact? In short, are they likely sources of significant measurement error, and do they appear to be worth being concerned about? A recent 77-country partial update of the database from the 100-country HCES review (Smith, Dupriez, and Troubat 2014) found that only 12 (16 percent) of the countries’ HCESs contained any information about household meal participation (Nathalie Troubat, FAO Statistics Division, personal communication, November 4, 2015). Moreover, most of those that did asked only about the number of household members who were present in the home during the recall period: they did not ask specifically about meal participation (as shown in Table 8.1). Nor is information generally collected about the household’s usual daily meal pattern. Given these limitations of current HCES questionnaires, the most that can be done to adjust for meals on an ex post basis in most countries would be to multiply the number of days at home by the individual’s or the household’s reported average number of meals per day or, if not available, an assumed average number of meals per day.

---

29 This is the dataset used by Conforti, Troubat, and Grünberger (2015), which was shared with the authors. Several additional HCESs that contain the information, as well, were subsequently identified and have been added to Tables 8.2 and 8.3.
Table 8.1 Household consumption and expenditure surveys (HCESs) collecting information on meal participation, as identified in a review of 77 countries’ HCESs

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Year</th>
<th>Information collected on meal participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>National Risk &amp; Vulnerability Assessment</td>
<td>2007/08</td>
<td>Number of HH member-residents who ate at least dinner regularly in the household during the last 7 days?</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>Household Budget Survey</td>
<td>2011</td>
<td>Number of household members present in the household the period the diary was administered</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>Household Budget Survey</td>
<td>2006</td>
<td>Number of household members present in the household the period the diary was administered</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Household Income and Expenditure Survey</td>
<td>2001</td>
<td>No specific question but in the module on daily consumption they asied number of male/female &lt; 10 years</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Household Income and Expenditure Survey</td>
<td>2006</td>
<td>No specific question but in the module on daily consumption they asied number of male/female &lt; 10 years</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Household Income and Expenditure Survey</td>
<td>2010</td>
<td>No specific question but in the module on daily consumption they asied number of male/female &lt; 10 years</td>
</tr>
<tr>
<td>Cape Verde</td>
<td>Inquerito as Despensa e RECEITAS</td>
<td>2001/02</td>
<td>Number of household members present in the household at the time the diary was administered</td>
</tr>
<tr>
<td>Chad</td>
<td>Enquête sur la Consommation et le Secteur informel au Tchad</td>
<td>2009</td>
<td>Number of meals household usually take per day per children and adults (male/female)</td>
</tr>
<tr>
<td>Congo, DR</td>
<td>Enquête Nationale du Type 1-2-3 auprès des Ménages</td>
<td>2004/05</td>
<td>Number of people present in the household during the reference period (15 days)</td>
</tr>
<tr>
<td>Lithuania</td>
<td>Household Budget Survey</td>
<td>1999/00</td>
<td>Number of people who consumed meal made at home and type of meal: breakfast, lunch and dinner.</td>
</tr>
<tr>
<td>Malawi</td>
<td>Third Integrated Household Survey</td>
<td>2011</td>
<td>Over the past one week (7 days), did any person that you did not list as household members eat any meals in your household?</td>
</tr>
<tr>
<td>Mali</td>
<td>Enquête Légère Intégrée auprès des Ménages (EUM)</td>
<td>2006</td>
<td>Count of number of meals served</td>
</tr>
<tr>
<td>Mongolia</td>
<td>Household Socio-Economic Survey</td>
<td>2007/08</td>
<td>Did any visitor stay here with your household for the last month? If yes, the number of days (person/days)</td>
</tr>
<tr>
<td>Niger</td>
<td>Enquête Nationale sur le budget et la consommation des ménages, ENBC</td>
<td>2007/08</td>
<td>Did any non-household members older than 1 year participate in the meal?</td>
</tr>
<tr>
<td>Uganda</td>
<td>Uganda National Household Survey</td>
<td>2010/11</td>
<td>Very exhaustive survey on meals prepared and consumed at home. Food was weighed.</td>
</tr>
</tbody>
</table>


Table 8.2 presents greater detail about meal participation data that the review by Conforti, Troubat, and Grünberger (2015) reported was collected for non–household members. The review revealed that greater detail is captured about the meal participation of non–household members than that of household members. It also showed that what is collected varies substantially by county and that few countries collect any of the measures identified. We found only four studies in which estimates of consumption (intakes of energy, micronutrients, or both) were adjusted for meal participation. One of those was done using a unique HCES from Cape Verde, which included a food weighing module for measuring household food consumption (Dop et al. 2012). Household food consumption was measured in the home by enumerators who visited before and after each main meal over seven days, and daily attendance of household members, as well as the presence of guests, was recorded for each of four meals (breakfast, lunch, afternoon snack, and dinner). Unfortunately, the highly atypical Cape Verde HCES study did not provide any insights into the significance of controlling for meal participation because, while it explained the methodology and made the adjustments, it did not report on the sizes of the adjustments or provide any other indication of their importance.

Table 8.2 Types of information collected on food given to non–household members

<table>
<thead>
<tr>
<th>Type of information collected</th>
<th>Number of surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Presence and/or household meal consumption of non-household members during the recall period</td>
<td>15</td>
</tr>
<tr>
<td>2. Number of visitors in the household</td>
<td>11</td>
</tr>
<tr>
<td>3. Visitors’ length of stay</td>
<td>5</td>
</tr>
<tr>
<td>4. Number of meals consumed by visitors/guests</td>
<td>10</td>
</tr>
<tr>
<td>5. Type of meal consumed by visitors/guests (breakfast, lunch, dinner)</td>
<td>7</td>
</tr>
<tr>
<td>6. Age of visitors/guests</td>
<td>7</td>
</tr>
<tr>
<td>7. Sex of visitors/guests</td>
<td>6</td>
</tr>
</tbody>
</table>

Source: Data from Smith, Dupriez, and Troubat (2014).
A study of India’s 1983 HCES also made adjustments for meal participation (Minhas 1991). Analysts used the meal patterns identified (using an earlier version of the questions presented in Figure 7.2) to develop a household-level adjustment factor, which was applied to the aggregated individual household members’ meal participation patterns over the previous 30 days, and to factor in free meals and meals served to others. The adjustment factor was as follows:

\[ \frac{(\text{Meals at home} + \text{Free meals})}{(\text{Meals at home} + \text{Meals served to guests} + \text{Meals served to employees})} \]

Table 8.3 shows the composition of total meals and reveals their relative numerical importance, by urban-rural residence. Assuming the caloric content of meals is roughly the same, the adjustment is likely to result in a relatively larger increase in the percentage of urban households with adequate estimated caloric intakes relative to rural households (because, on average, urban households received relatively more free meals than meals they served to others and to employees).

<table>
<thead>
<tr>
<th>Number of meal per HH served to:</th>
<th>Number of meals per HH taken AFH</th>
<th>Total meals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guests during ceremonies</td>
<td>Other guests</td>
<td>Employees</td>
</tr>
<tr>
<td>Urban</td>
<td>3.33</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>1.0%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Rural</td>
<td>5.97</td>
<td>6.99</td>
</tr>
<tr>
<td></td>
<td>1.7%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

Source: Constructed by authors from Minhas (1991).
Note: HH = household; AFH = away from home.

The only other two studies we found that reported meal participation adjustments both analyzed the meal participation only incidentally as it related to other issues or characteristics of HCES data that were their main focus. The objective of a Bangladesh study by Coates and colleagues (2016) was to assess the consistency of the estimates of inadequate intakes using 24HR-based and AME-based estimates of inadequate intakes. It found that accounting for meal partakers did not improve the consistency of the estimates. It did not, however, report the significance of the adjustment—namely, the frequency of households for which the adjustment was made, or the magnitude of the adjustment—in an absolute sense.

Conforti, Troubat, and Grünberger (2015) conducted a multivariate analysis of the impact of key survey design characteristics in 77 countries. They found that 17 (22 percent) of the countries collected data on the number of meal partakers, but that (controlling for the national dietary energy supply) the number of meal partakers had no impact on either the estimated national average level or the coefficient of variation of apparent energy intakes. (The authors used the HCES-processing software ADePT [Moltedo et al. 2014], which is designed to allow for adjusting for the number of partakers. This is noteworthy because ADePT’s inclusion of this feature will both encourage and facilitate others to make the adjustment.)

It is important to explicitly state that although the statistical tests of the significance of the adjustments for meal participation in these studies find the adjustments to be insignificant, one must be careful not to conclude that meal participation is not an important issue that still needs to be addressed.
The HCESs that currently attempt to make these adjustments are few and highly diverse, and most are plagued by the same shortcomings as food away from home (FAFH), which is one of the factors that figures into meal participation. In several of the countries, the questionnaires appear to capture only a portion of the requisite information (that is, they lack face validity) and the results are likely subject to considerable measurement error.

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. We believe the sources of measurement error related to capturing meal participation—in particular, the increasing heterogeneity of individual household member behaviors with their derivative increased recall burden for the HCES key respondent—will continue to increase over time, underscoring the urgency of the need to address these shortcomings. How, for instance, can HCESs better capture FAFH, which is consumed disproportionately by household members who stay away from home for more than 24 hours at a stretch during the recall period, such as, for example, in India the 6 percent of the population that reported staying away from home in the previous 30 days and ate 36 percent of all meals consumed away from home?

HCESs have one primary respondent, who is increasingly likely to be challenged to be aware of, remember, and report the meal participation rates of all members of the household for all days and all meals of the recall period. It would seem that the best approach would be a household member–specific set of questions to capture this information. How to structure that question (or set of questions) to enable best capturing the requisite information, while minimizing the additional interviewing time required, is still an outstanding issue. These are difficult issues to address with any degree of confidence in the absence of conducting an experiment, or better yet—in lieu of the likely importance of contextual influences—a series of experiments in different countries.

It is noteworthy that, by implication, these studies may not provide an accurate portrayal of the actual situation in several of the studied countries: that is, 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 in which there is greater travel away from home and where, more generally, there is more widespread practice of eating away from home. We conclude that it is premature to make any definitive assessments about the value of making adjustments for meals or, more fundamentally, (even) about the feasibility of being able to collect the requisite information to make such assessments, given the inherent difficulties of doing so in a household-based, single-respondent, multipurpose survey, and the trade-offs entailed in terms of the additional time that would be required to collect those additional data. The major contribution of this proposed study in addressing this issue, then, is to document the different approaches and specific questions that have been introduced, and to attempt to assess the potential value-added of these initiatives, recognizing the questionable external validity of their findings.

A review of selected questionnaires from more recent HCESs reveals that over the course of the past five years a number of countries—India, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda, among others—have introduced new questions to identify who among the household’s members were meal takers, which meals were eaten away from home, and by whom. Several of the surveys have also asked about meals that households provided to non–household members. These modifications appear to have been motivated by concerns that FAFH is increasing in frequency, is significantly underreported, and is distorting the precision of food security analyses. The number of questions that have been added to capture FAFH and their specific wording vary substantially. Other than the Minhas (1991) study, we could find no evidence of these questions’ actually being used to examine the significance of or to adjust for meal participation. It is uncertain how or whether the national statistical agencies of these countries are using the additional information.

The number of meals served to non–household members varied dramatically depending on whether or not the household sponsored a ceremony (as noted earlier).30 Households that sponsored a

---

30 The India National Survey Sample Organization’s Round 68 instructions to field staff defined a “ceremony” as “an
ceremony are likely to be extreme outliers in terms of the overestimation of their consumption if meals they serve to others are not identified and subtracted from estimates of the food the household’s members consume. Only a few countries—Bangladesh, India, and Viet Nam—collect data about special occasions such as ceremonies, festivals, or holidays, which are likely to be associated with these large distorting influences. This is likely to be a country-specific topic that warrants clarifying whether or not it should be asked about.

**Takeaway Messages**

Relatively few HCESs collect data with which to make any adjustments for meal participation. Those that do generally collect only information on the number of household members who were present in the household during the recall period or about the number of meals usually eaten and the number eaten away from home. While some HCESs collect data on the number of non–household members present, the extent to which they have used those data to adjust for meal participation is not known.

Of the five HCES-based studies reviewed, there is evidence from only one that adjusting for meal participation has a significant impact on a household's estimated level of consumption, although the studies were designed to address other issues, and the participation issue was addressed only tangentially.

The number of additional questions that would be required to make comprehensive adjustments is viewed as too burdensome, and an alternative, less demanding strategy is to capture information on more systematic behavioral patterns with which to make such adjustments. Two such patterns that look to be important are the hosting of special events (for example, weddings) and household members’ staying away from home for at least 24 hours.

occasion on which a large number of meals (not just snacks) were served to non–household members which significantly affects the household’s total expenditure during the last 30 days. The occasion need not be religious” (India, NSSO 2014, sect. 3.3.14).
9. PRIORITIZING THE SIX KEY STUDIES

What are the next steps in addressing the HCES agenda? What criteria should be used to decide how the work of undertaking the six studies this paper proposes should proceed? An obvious initial knee-jerk response might be that it should start with what is regarded as the most important study. This of course raises the question, how might we gauge the relative importance of these studies? Is a particular study “the most important one” because it addresses what is regarded as the source of the largest measurement error among the studies? Or is it “the most important one” because it is viewed as being able to result in the largest reduction in measurement error? One may question whether we even know enough about the measurement of food consumption and nutrition using HCESs to be able to gauge the magnitude of the measurement error involved in many of these studies—in either absolute or relative terms—well enough to develop a measurement error–prioritized rank ordering of them using either of these approaches. There are other considerations in plotting a way forward, as well.

From Whose Perspective? Judging the Relative Importance of Household Consumption and Expenditure Survey Shortcomings, and Rank Ordering the Proposed Studies for Global versus Country-Level Agendas

In part, the relative significance of the proposed studies depends on whether HCESs are being assessed from a global perspective—considering all HCESs together—or from the perspective of an individual country. The 100-country HCES review (Smith, Dupriez, and Troubat 2014) prioritized these particular studies due to a combination of how commonly (that is, the number of countries in which) they pose problems and the magnitude of the measurement error associated with them. The priority rankings of these studies is likely to vary across countries because the identified shortcomings are systematically related to various characteristics of the country, including dietary patterns, standard of living, level of urbanization, and other factors reflecting aspects of social organization, culture, and lifestyle. (See Appendix 3 for further detail on the nature of some of these relationships.)

A global-level approach can provide general guidance and some broad parameters, and lay out some common concerns and issues that are widely relevant, but more specific and more prescriptive suggestions and “answers” will need to be country specific. Global-level recommendations, for example, would include general guidance, such as the following:

1. The nutrient content of foods for which only expenditures (not quantities) are reported should be imputed from the average cost per nutrient of foods for which quantities and values are reported. The specific imputation methods should be a country-specific decision, taking into account characteristics in food markets, subnational/regional variability in dietary patterns and food markets, and variability in dietary patterns and food markets by socioeconomic status and location. Different qualities of important foods should be taken into account in the food list so that quality-specific prices can be used to impute quantities.

2. Improving data collection on food away from home (FAFH) and on meal participation will require introducing an individualized approach to collecting meal data on what routine meal patterns are, where meals are usually obtained or eaten, and whether or not individual household members have spent 24 continuous hours away from home. In those countries that feel this will require too many additional questions and too much additional interview time, this review has provided evidence that, at minimum, information should be collected about which household members stayed away from home, and for how long. Children’s participation in school feeding programs has also been shown to be important, common, and easy to collect information about.
3. Getting the food list “right” means the number of items in the food list should be roughly 100–125. The exact number and the specific items, however, should depend on country-specific considerations such as the content of the diet and variability in food consumption patterns. For instance, countries in which a very high percentage of the nutrients of interest come from a small number of foods may be able to have a smaller food list.

4. Getting the food list “right” also involves ensuring that significant food sources of essential nutrients are included in the list, as well as vitamin and mineral supplements. The particular food items, however, should be country specific and should include items that are currently important nutrient sources and also ones that are potentially important, including those that are fortifiable or biofortifiable.

5. The recall period should be at most 14 days. The exact number, however, is a country-specific determination that depends on food purchasing habits, food storage routines, dietary patterns, food-specific shelf life, the importance of home food production, and other considerations.

The question remains whether it would be best to begin addressing this reform agenda by working across countries on particular HCES shortcomings or by identifying specific countries that are in particular need or that have common characteristics that we hypothesize condition or directly impact how well HCESs collect food consumption data. This review has drawn upon key findings of the literature to guide the shaping of the agenda toward those topical areas of greatest apparent significance, recognizing that in many cases the external validity of the findings is questionable or at best uncertain. At the very least, there is a need to recognize that both intuition and empirical evidence would suggest that the time is past due for continued discussions of all HCESs simultaneously. It is clear that lessons may be garnered for some HCESs in some countries, but not in others. It is time to start developing more detailed understandings about which lessons are so constrained and in which countries, and what distinguishing features or conditioning factors there are, and why.

This point has been exemplified in this paper in the lengthy discussion of the very different types of issues raised by HCESs that collect information on food consumption, compared with those that collect a combination of purchase and consumption data. It is time to begin categorizing HCESs into typologies that will provide us with more analytically useful ways to study how best to strengthen HCESs by helping to control for some of the many confounding influences, while better ensuring we are comparing apples to apples and oranges to oranges, and accelerating the pace of progress. (See Appendix 3 for some proposed criteria for constructing these typologies.)

**The Potential Importance of Packaging and Sequencing the Studies**

In developing a strategy as to how to proceed with addressing these studies, it is important to recognize that many of the studies overlap in terms of the phenomenon they measure. As a result, how one issue or shortcoming is addressed is likely to influence the magnitude and perhaps even the nature of other issues, thereby affecting the need for and the urgency of other studies. It is important, therefore, to consider how the studies are packaged and to recognize that their sequencing (or possible combination) might be consequential.

The topics of FAFH, meal patterns, and the food list, for instance, overlap: data on FAFH are oftentimes collected by specifying particular food list items as being eaten away from home (Figure 7.1), but it is sometimes addressed by asking about routine meal patterns or the meal patterns of individual household members during the recall period, such as in the India HCES (Figure 7.2). In other instances, FAFH is identified by asking about where meals were acquired and consumed. Clearly there is a need to be cognizant of these overlaps to ensure that double counting is avoided, while built-in checks are made use of and not allowed to simply become redundancies.
This situation suggests that it may make sense to start the complex activity of getting the food list “right” by addressing the more delimited and manageable issue of FAFH. It also encourages us to consider whether one way to address all three of these sets of issues and concerns simultaneously and in a way that would provide an overall analytical framework to help both interviewers and interviewees organize their thoughts, formulate responses, and perform consistency checks would be to ask about meal patterns throughout the recall period, and to do so at an individual household member level, rather than at the household level, which is a much more onerous and cognitively demanding task (as discussed earlier).

**Beauty Is in the Eye of the Beholder: Repurposing Household Consumption and Expenditure Surveys and Anticipating the Degree of Opposition of Existing Stakeholders**

Another pertinent consideration is the extent to which the shortcomings being addressed in the proposed studies will require making changes in the HCES questionnaire or whether they will entail modifications only in how HCES data are processed. Those requiring only changes in how data are processed are low-hanging fruit: they can be implemented immediately by researchers, and—not entailing any changes in HCES instrumentation—they are largely free of the need for negotiations with existing HCES stakeholders. There are two such studies (identified in column D of Table 9.1): an assessment of the appropriateness using the adult male equivalent (AME) and an investigation of alternative nutrient imputation methods. Relative resource and time constraints, and in the case of the AME assessment study, the availability of other data sources (namely, 24-hour recall) remain potential additional considerations that may require choosing between them as well as negotiating with existing stakeholders. In addition, both will require relatively sophisticated analysts, who may be in limited supply, although both studies could be done by a single analyst.

**Table 9.1 A prioritized, global agenda for the key household consumption and expenditure survey food consumption studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Requires changes in the questionnaire</th>
<th>Expected difficulty in obtaining approval of current HCES stakeholders</th>
<th>Amount of additional preliminary work required to design/test revisions</th>
<th>Requires only changes in data processing</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Improving Nutrient Imputation Methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2. Assessing the AME as a proxy for intra-household distribution of food</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>3. Better capturing consumption of food away from home</td>
<td>X</td>
<td>XL</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>4. Modifying the food list to better capture important food sources of essential nutrients</td>
<td>X</td>
<td>XH</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>5. Adjusting food consumption estimates for meal participation</td>
<td>X</td>
<td>XM</td>
<td>X</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>6. Optimizing the recall period</td>
<td>X</td>
<td>XH</td>
<td>X</td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

Source: Authors

Note:  

a H = high; M = medium; L = low. Gauged by anticipated impact on other existing HCES indicators, such as ability to conduct trend analysis, extensiveness of the required changes, and their additional time and resource requirements.  
b Depending on study design, may require 24-hour recall survey data to be available and analyzed. c It is suggested that typologies of HCESs be developed, differentiated by variations in characteristics of HCESs (for example, whether they collect only food consumption data versus a mixture of purchase and consumption data) and country characteristics that influence the relative importance of the studies (for example, level of income, urbanization, importance of food away from home, consumption from own production, and dietary pattern variability). HCES = household consumption and expenditure survey.
Another important consideration in choosing between or rank ordering these two studies is the intended use of the results, which, in turn, may dictate whether the household-level data collected are adequate or the intended application requires developing proxy measures of individual-level measures of nutrition. If only household-level data are of interest, there is no need to assess how well the AME proxies intrahousehold distribution.31

For the four remaining studies, which all entail HCES questionnaire revisions or additions in order to better capture food and nutrition data, an important question is, how difficult is it likely going to be to obtain the consent of existing HCES stakeholders to introduce those changes? For strategic reasons, it is useful to categorize modifications in questionnaires as being of two principal types: those introducing new questions and those revising existing questions. Introducing new questions is likely to be discouraged by existing stakeholders because these multipurpose surveys are already widely regarded as too long, and any net new additions may increase the likelihood of reducing the quality of the data collected as overly burdened respondents seek ways to reduce their time input, including refusing to participate in the survey altogether. That is a concern that is likely to be shared by all existing stakeholders and to result in most stakeholders’ generally discouraging new questions.

In contrast, depending upon the specific questions involved, the modification of existing questions is likely to affect fewer existing stakeholders but to affect them relatively more directly and with potentially greater impacts. Changing existing questions generally results in discontinuities in serial data, undermining the ability to undertake trend analysis and thereby to understand changing behavior patterns and socioeconomic dynamics, which are among the most important and policy-relevant reasons the HCESs are conducted. As a result, generally, it is likely to be easier to introduce new questions than to modify existing ones or, at least, to allow a period of overlap in which a question is asked in both the “old” way and the “new” way, where possible.

The specifics of how best to modify a questionnaire may not be obvious, or they may be contentious. In both of these instances, an additional consideration may be how much preliminary diagnostic work is likely to be needed to be able to arrive at a consensus about how best to design the study. Additional, derivative concerns are the relative technical capacity, time, and resources required to undertake the preliminary studies, and whether or not these capacities and resources are in adequate supply, available, and affordable enough to ensure that they do not preclude the studies’ being undertaken or completed in an adequately rigorous manner such that their results are accepted and the studies are able to realize their anticipated impact. The technical and resource requirements are obvious concerns in the prioritization and sequencing of the six studies proper, as well.

We anticipate that because the inadequate capture of FAFH is such a common and widely acknowledged HCES shortcoming, despite the fact that the issue of how best to go about improving it is not without major technical and political challenges, it will be relatively easy to persuade stakeholders to begin taking steps to ameliorate this common source of underreporting consumption. First steps might consist of identifying general meal patterns (as already discussed and recommended) to capture the number of meals obtained and eaten away from home by individual household members, or they could consist of adding types of meals (breakfast, lunch, dinner, snack) or specific types of foods eaten away from home.

---

31 The AME remains a useful construct, however, with which to adjust household nutrient intake measures to take into account the size and composition of the household in assessing the household’s nutrient intake adequacy.
Ordering and Sequencing Nutrition-Specific Objectives in Repurposing Household Consumption and Expenditure Surveys: An Elusive Global Goal That Is More Manageable at the Individual Country Level

It is important to recognize that different studies and different policies will require using HCESs to derive different types of indicators with different levels of precision and reliability. The least demanding indicators are those based on the consumption of food groups, followed by those based on more specific food items. The more demanding metrics are those that require measuring food quantities, and the most demanding are those requiring measurement of nutrient intakes.

Table 9.2 identifies information needs related to the program cycles of five nutrition interventions: fortification, supplementation, dietary diversification, biofortification, and modification of the food crop mix to improve nutrient availability and intakes. The table includes two different categories of metrics: the more easily collected and reported binary measures of whether a particular food group or a particular food item is eaten or not; and the more difficult to precisely and reliably capture, continuous measures of food quantities and nutrient intakes.

As shown in Table 9.2, the specific information needed and when it is needed varies by program and changes over time, even within a single country. This variability suggests that while we may be able to provide guidance on how to improve the performance of HCESs to better capture food consumption and nutrition information, there is no single best way of doing so; it depends on the country, on the nature of nutrition status, on the perception of the relative importance of different nutrition problems, and on the willingness and ability of policy makers and society to address those problems.

This review has sought to craft a strategic approach to the unfinished global agenda of improving HCESs’ collection of food consumption data. Starting with the priority studies recommended by the 100-country HCES review (Smith, Dupriez, and Troubat 2014), it has focused on a strategic subset of those studies: those dealing most directly and exclusively with the measurement of food. Drawing from the literature, this review has provided a more detailed, circumspect justification as to why these particular studies are needed, while identifying key hypotheses, explaining why these studies are of growing urgency, and demonstrating why now is a propitious time for undertaking them. The review has also identified important study design considerations while pointing out potential challenges to successful implementation stemming from technical capacity as well as economic, administrative, and political considerations, and it often noted that decision points are frequently not unambiguously yes or no, with better or worse likely implications or results. Instead, they are often more equivocal and, even with the aid of additional information, may still require the exercise of considerable judgment in assessing difficult and ineluctable trade-offs.
Table 9.2 Food consumption and nutrient intake information needs during nutrition program cycles

<table>
<thead>
<tr>
<th>Program Phase</th>
<th>Program</th>
<th>Information Need</th>
<th>Binary Measures</th>
<th>Continuous Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Food Group</td>
<td>Food/Crop Item</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Food Quantity</td>
<td>Nutrient Intake</td>
</tr>
<tr>
<td>Problem Identification/Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>What proportion of the population suffers inadequate intake of a micronutrient? (Inadequate Intake Prevalence)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Severity of inadequate intake? (Inadequate Intake Gap)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>How is the problem distributed in the population? (Problem Concentration &amp; Program Targeting Potential)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>What is the relationship of deficiency to inadequate intake?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>What is the level of Dietary Diversity? (Prevalence of Inadequate Dietary Diversity)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>How is Dietary Diversity distributed in the population? (Problem Concentration &amp; Program Targeting Potential)</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Program Intervention Selection and Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Which food crops containing the nutrient are being produced in the country?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Could food crops containing the nutrient be produced in the country?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Are there varieties of the food crop biofortified with the nutrient?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Which food crop varieties are being consumed regularly by large proportions of the population?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>How have these consumption patterns have changed over time?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>What proportion of the population consuming these crops has inadequate intake of a micronutrient that can be added to this crop?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>What is the nutrient intake gap, considering all sources of consumption?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>How is the intake gap distributed in the population?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>What quantities of the potentially biofortified food crop are being consumed?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>How is the consumption of the crop/food vehicle distributed in the population?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Program Fortification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Which foods that are potentially fortifiable with this nutrient are being produced in the country?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>What proportion of this consumption is obtained through the purchase of centrally processed foods (amenable to large-scale fortification)?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>How have these consumption patterns changed over time?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>What proportion of the population purchasing these foods has inadequate intake of a micronutrient that can be added to this food vehicle?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>What is the nutrient intake gap, considering all sources of consumption?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>How is the intake gap distributed in the population?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>What quantities of the potential food vehicle are being consumed?</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>How is the consumption of the crop/food vehicle distributed in the population?</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Program Phase</td>
<td>Program</td>
<td>Information Need</td>
<td>Binary Measures: Food Group</td>
<td>Food Crop Item</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Program Monitoring</td>
<td></td>
<td>24. What proportion of the population is consuming the fortified food/biofortified crop or participating in the supplementation or education program? (Reach)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25. What proportion of the population is consuming the fortified food/biofortified crop or participating in the supplementation or education program? (Coverage)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>26. What proportion of the target population is consuming the fortified food/biofortified crop or participating in the supplementation or education program at a level expected to eliminate the nutrient intake gap? (Effective Coverage)?</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27. What proportion of the target population is consuming the fortified food/biofortified crop or participating in the supplementation at a level likely to exceed the upper level? (Risk of Excessive Intake?)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Impact Evaluation</td>
<td></td>
<td>28. What proportion of the population and target population have achieved significant improvements in micronutrient intake?</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>29. What is the relationship between change in dietary intake and micronutrient deficiency?</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Coates et al. (2012).
### APPENDIX A: RECENT STUDIES WITH NEW APPLICATIONS OF HOUSEHOLD CONSUMPTION AND EXPENDITURE SURVEYS

**Table A.1 More recent food insecurity and nutrition: Related applications of household consumption and expenditure surveys**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conducting subnational food security analysis</td>
<td>Sibrián 2008; Sibrián, Seevalingum, and Mernies 2008; Tanzania, NBS, et al. 2010</td>
</tr>
<tr>
<td>Assessing the diversity of food supplies</td>
<td>Fiedler 2014; Fiedler and Lividini 2015</td>
</tr>
<tr>
<td>Analyzing the nutrient availability of the domestically produced food supply</td>
<td>Fiedler 2014; Fiedler and Lividini 2015, 2016</td>
</tr>
<tr>
<td>Assessing dietary diversity</td>
<td>Rashid and Smith 2011; Fiedler 2014</td>
</tr>
<tr>
<td>Analyzing the relationship between household food expenditure and malnutrition</td>
<td>Campbell et al. 2010</td>
</tr>
<tr>
<td>Identifying and monitoring the prevalence of overweight and obesity</td>
<td>Kolodinsky and Goldstein 2011; Lobato 2009</td>
</tr>
<tr>
<td>Assessing the consumption of snacks and soft drinks by babies in low- and middle-income countries</td>
<td>Huffman et al. 2014</td>
</tr>
<tr>
<td>Identifying the most common food sources of specific nutrients</td>
<td>Bermudez et al. 2012; Fiedler et al. 2013a; Fiedler and Lividini 2016</td>
</tr>
<tr>
<td>Designing and modeling the impact of biofortification programs</td>
<td>Stein, Sachdev, and Qaim 2006, 2008; Stein et al. 2007, 2008; Fiedler et al. 2013; Fiedler and Lividini 2014; Lividini and Fiedler 2015</td>
</tr>
<tr>
<td>Conducting feasibility and cost-benefit analyses of fortification, biofortification, and supplementation program portfolios</td>
<td>Fiedler et al. 2013; Fiedler and Lividini 2014, 2015; Engle-Stone and Brown 2015; Fiedler et al. 2015</td>
</tr>
<tr>
<td>Conducting nutrition epidemiological analysis</td>
<td>Naska et al. 2009, 2015; Psaltopoulou et al. 2004</td>
</tr>
</tbody>
</table>

Source: Authors.
Possible Analytic Studies of Existing HCESs that Can Help to Improve the Accuracy, Reliability, and Efficiency of HCESs’ Collection of Food Consumption Data

- Analyze HCESs with multiple recall periods to better understand the robustness of food-specific consumption measures—for example, to better understand purchase patterns or the significance of shelf life and other food characteristics in affecting how food consumption patterns vary by recall period.
- Analyze the 68th round of India’s consumer expenditure survey (2011/2012) to juxtapose results of the 7-day recall and 30-day recall periods for each of the 87 food items individually, in dietary diversity food groups, and grouped by food characteristic (for example, shelf life / perishability).
- Quantify the significance of each item reported in the food item list in terms of the percentage of households apparently consuming it and the quantity and value of it consumed, in both absolute and relative terms (that is, as a percentage of total weight, total value, and total nutrients).
- Quantify the frequency with which residual food consumption table categories “other x” (for example, “other fruits,” “other vegetables”) are reported by type of food and length of the food list to provide a first approximation of the need for and space available for increasing the specificity of the food list.
- Assess the value-added of the 11 food group questions being asked in the eight Living Standards Measurement Study–Integrated Surveys on Agriculture countries: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda. Develop a consistency analysis of the 11 food groups as reported in these questions versus as reported by individual food item (after mapping the individual food items to the 11 food groups). Test concordance of the two series using Cohen’s kappa statistic for the binary measure of yes/no for consuming at least one item in the food group, and using Lin’s concordance correlation coefficient to test the continuous measures of the quantity or value consumed of the food group.
- Analyze the impact that adjusting for meals eaten away from home has on estimated food consumption levels and the adequacy of nutrient intakes. Candidate databases include those of India, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda.

Country Characteristics That Shape the Structure and Accuracy of Household Consumption and Expenditure Surveys: Level and Patterns of Development, Cultural Traits, and Dietary Habits

Level and Patterns of Development

- Level of development / gross domestic product
  - Greater income, greater diversity of food requires longer food list
  - More infrastructure, greater socioeconomic integration, greater market articulation, more articulated food value chains
    - Longer food list
    - More diverse preparations of similar food items creates need for a more articulated food consumption table
• More accessible food markets, shorter routine shopping intervals make a shorter recall period possible
• More perishable food purchased should be reflected in specificity of food list
• Purchasing food away from home (FAFH) more frequently (meals and snacks) more common and more important in diet, more important to capture

• Percentage of food from own production is inversely related to purchases and FAFH
• Percentage of labor force in agriculture
• Urbanization
  ▪ Correlated with the prevalence of FAFH: greater need to better address FAFH
  ▪ More developed food market: greater specificity in the food list
• Greater mobility and increased likelihood of staying away from home and eating FAFH
• School feeding programs: coverage, participation rates, standards
• Frequency of employer-provided meals

**Culture, Food, and Dietary Considerations**

• Nutrient deficiency: types and severity of nutrient deficiencies should be reflected in the food list with more food entries containing significant amounts of the nutrients of deficiency
• Dietary diversity: lower diversity may mean shorter food list feasible
• Percentage of calories from staple crops: shorter food list may be feasible
• Percentage of calories from starches: shorter food list may be feasible
• Cultural diversity and dietary diversity demands longer food list to capture different types of food, different methods of preparation, different dishes
• Ceremonies and festivals: may need to ask about frequency and nature to control for distorting impacts on household’s apparent food consumption
• Household size: larger households entail greater cognitive burden on key respondent, leading to more recall error, measurement error
• Usual number of meals eaten in a day
REFERENCES


Recent IFPRI Discussion Papers

For earlier discussion papers, please go to www.ifpri.org/publications/discussion_papers.
All discussion papers can be downloaded free of charge.


1554. Implications of slowing growth in emerging market economies for hunger and poverty in rural areas of developing countries. David Laborde and Will Martin, 2016.


