

# Measuring Household Welfare

Short versus long consumption modules

---

Marta Moratti and Luisa Natali

Office of Research Working Paper

WP-2012-No. 4 | October 2012

## INNOCENTI WORKING PAPERS

UNICEF Office of Research Working Papers are intended to disseminate initial research contributions within or relevant to the Centre's programme of work, addressing social, economic and other aspects of the realisation of the human rights of children.

The findings, interpretations and conclusions expressed in this paper are those of the author(s) and do not necessarily reflect the policies or views of UNICEF.

The text has not been edited to official publications standards and UNICEF accepts no responsibility for errors.

Extracts from this publication may be freely reproduced with due acknowledgement.

Requests to utilize larger portions or the full publication should be addressed to the Communication Unit at [florence@unicef.org](mailto:florence@unicef.org).

For readers wishing to cite this document we suggest the following form:

Moratti, M. and L. Natali (2012), 'Measuring Household Welfare: Short versus long consumption modules', *Working Paper 2012-04*, UNICEF Office of Research, Florence.

© 2012 United Nations Children's Fund (UNICEF)

October 2012

ISSN: 1014-7837

## THE UNICEF OFFICE OF RESEARCH

In 1988 the United Nations Children’s Fund (UNICEF) established a research centre to support its advocacy for children worldwide and to identify and research current and future areas of UNICEF’s work. The prime objectives of the Office of Research are to improve international understanding of issues relating to children’s rights and to help facilitate full implementation of the Convention on the Rights of the Child in developing, middle-income and industrialized countries.

The Office aims to set out a comprehensive framework for research and knowledge within the organization, in support of its global programmes and policies. Through strengthening research partnerships with leading academic institutions and development networks in both the North and South, the Office seeks to leverage additional resources and influence in support of efforts towards policy reform in favour of children.

Publications produced by the Office are contributions to a global debate on children and child rights issues and include a wide range of opinions. For that reason, some publications may not necessarily reflect UNICEF policies or approaches on some topics. The views expressed are those of the authors and/or editors and are published in order to stimulate further dialogue on child rights.

The Office collaborates with its host institution in Florence, the Istituto degli Innocenti, in selected areas of work. Core funding is provided by the Government of Italy, while financial support for specific projects is also provided by other governments, international institutions and private sources, including UNICEF National Committees.

Extracts from this publication may be freely reproduced with due acknowledgement. Requests to translate the publication in its entirety should be addressed to: Communications Unit, [florence@unicef.org](mailto:florence@unicef.org).

For further information and to download or order this and other publications, please visit the website at [www.unicef-irc.org](http://www.unicef-irc.org).

### **Correspondence should be addressed to:**

UNICEF Office of Research - Innocenti  
Piazza SS. Annunziata, 12  
50122 Florence, Italy  
Tel: (+39) 055 20 330  
Fax: (+39) 055 2033 220  
[florence@unicef.org](mailto:florence@unicef.org)  
[www.unicef-irc.org](http://www.unicef-irc.org)

## MEASURING HOUSEHOLD WELFARE: SHORT VERSUS LONG CONSUMPTION MODULES

Marta Moratti<sup>a</sup> Luisa Natali<sup>b</sup>

<sup>a</sup> Oxford Policy Management

<sup>b</sup> UNICEF Office of Research

**Abstract.** Consumption expenditure is probably the most common and preferred welfare indicator; however, its measurement is a challenging and time-consuming task. Although short consumption modules have potentially enormous advantage in terms of time and money savings, a recent and comprehensive literature on available experiments comparing short versus long modules is still lacking. The present paper aims at filling this gap trying to draw conclusions in terms of the accuracy of consumption and related poverty and inequality (distributional) estimates based on short modules.

First, the paper briefly reviews the literature on how to accurately measure consumption and how survey design can influence consumption estimates; then, the empirical literature is discussed.

The literature review mainly focuses on studies from the 1990s on developing countries. Available evidence seems to indicate that short modules underestimate consumption with respect to longer ones resulting in lower levels of recorded consumption and therefore less accurate estimates and higher poverty rates.

However, one of the most complete, recent and authoritative studies in the field (Beegle et al., 2010) finds that short modules may actually result in a smaller downward bias compared to the benchmark than other longer consumption modules.

In terms of relative ranking of households, the literature is scant; however, results from rigorous studies indicate that household consumption rankings obtained through short consumption modules are largely consistent with rankings derived from long modules.

A critical review of the available evidence points to a number of factors that hinder the ability to draw firm conclusions; it indicates that there is still room for further investigation and provides some guidance for future field experiments in order for them to reach conclusiveness.

**Keywords:** Measures of welfare, household consumption measurement, survey design, poverty, inequality

**Acknowledgements:** The authors would like to thank Sander Den Boer and Lucian Constantin Mircescu who contributed to the collection of background material.

They are grateful to Patrick Ward for reviewing the paper and providing useful comments; Fahrur Memic for sharing his experience with a short consumption module in BiH MICS3; and to Attila Hancioglu and the Global MICS Team for their assistance and for making it possible to test two short consumption modules in Bangladesh.

Thanks also go to Laura Meucci for her administrative support and to Eve Leckey for her help with the proofs and the layout process.

Finally, the authors would like to thank Chris de Neubourg for his overall support throughout the work.

## TABLE OF CONTENTS

1. Measures of Household Welfare: Why is consumption the most common and preferred welfare indicator?	6
1.1 Asset indices as an alternative measure of welfare	6
1.2 Money-metric measures: Income or consumption?	7
2. How to Measure Consumption?	8
3. How Does Survey Design Influence Consumption Measurement?	10
4. Short Versus Long Consumption Modules – A Literature Review	12
5. Conclusions	24
References	26

## 1. MEASURES OF HOUSEHOLD WELFARE: WHY IS CONSUMPTION THE MOST COMMON AND PREFERRED WELFARE INDICATOR?

Household surveys are an essential source of information on economic and social conditions of households and individuals. Survey data can be used to measure the welfare of households, poverty, and how equally distributed are living standards. Moreover, welfare measures allow to investigate patterns in standards of living across populations and over time.

The focus of the present paper is primarily the use of household surveys for welfare, poverty and inequality analysis. While different indicators exist to perform welfare analysis, consumption has been long favoured by economists as a proxy for living standards. This section analyses in detail, from a theoretical as well as practical point of view, the pros and cons of different indicators often used to measure household welfare.

### 1.1 Asset indices as an alternative measure of welfare

Welfare is usually proxied by measures of consumption or income. However, in recent years, the use of asset-based wealth indices as an alternative metric measure of welfare has become increasingly prominent. Indeed, wealth indices represent the only way to investigate distributional aspects in uniquely detailed large-scale surveys – such as MICS (Multiple Indicator Cluster Surveys) and DHS (Demographic and Health Surveys) – that lack information on income and/or consumption (Howe et al., 2008).

The wealth index has sometimes, and also more recently, been considered a theoretically and practically superior alternative measure of economic status to income and consumption (Rutstein and Johnson, 2004): wealth better reflects long-term welfare as it is less volatile than both income and consumption; it is considered more suitable to analyse multi-dimensional poverty (Filmer and Pritchett, 1999; Filmer and Pritchett, 2001); and finally it is less data intensive and therefore easier to calculate (Sahn and Stifel, 2000; 2003; Azzarri et al., 2006).

However, these features make the wealth index a specific indicator and as such it cannot be comparable to the conventional measures of economic status. Different studies report that the asset index is in fact a generally poor proxy for current household income or expenditure (Filmer and Pritchett, 2001; Montgomery et al., 2000), while it may be a good proxy for long-term or permanent income. Furthermore, there are a number of conceptual and pragmatic reasons that limit the use of asset-based indices as alternative measures of welfare.

First of all, the wealth index provides a relative measure of welfare – namely a household's wealth is measured relative to other households in the sample – but does not quantify the household's current levels of welfare or poverty (Filmer and Pritchett, 2001).

The wealth index – as most commonly constructed – has also been found to have an urban bias and limited discriminatory power at the lower end of the wealth distribution (Filmer and Pritchett, 2001; Rutstein, 2008; Howe et al., 2010).

Moreover, differences in price levels across regions are not taken into account in the asset-based approach (Howe et al., 2008) and the quality of assets is ignored (Moser, 1998); weights on

individual indicators are not grounded theoretically (Filmer and Pritchett, 2001) and the appropriateness of the wealth index is likely to differ across sub-groups of the population. Depending on the purpose of the research, the indicators included in the index might have direct effects on the outcome of interest (Rutstein and Johnson, 2004; Howe et al., 2008).

Finally, substantial concerns emerged with regard to the use of the wealth index for welfare comparison over time and across countries. Although recent studies have proposed methodologies to allow for inter-temporal and intraregional comparisons (Sahn and Stifel, 2000; 2003; Booyesen et al., 2008; Howe et al., 2010), they do not convincingly overcome the observed limitations.

Therefore, for a series of theoretical as well as practical reasons, the wealth index cannot be used as a perfect substitute for income or consumption which, among other considerations, remain the most common and accepted measures of welfare.

### **1.2 Money-metric measures: Income or consumption?**

Researchers have debated intensely on the strengths and weaknesses of different welfare indicators with a quite clear consensus on favouring consumption over income, especially in a developing country context.

In the first place, individuals derive material well-being from the actual consumption of goods and services rather than from the receipt of income per se (Citro and Michael, 1995); therefore consumption seems to better capture the concept of 'standard of living'.

Deaton and Zaidi (2002) argue that consumption better reflects long-term income as it is not closely tied to short-term fluctuations in income and is smoother and less variable than income. Income is more likely to be affected by seasonal patterns resulting either in an underestimation or overestimation of real income. Consumption is more stable especially in agricultural societies as it is smoothed over the seasons, therefore better reflecting (or approximating) the real living standard.

Moreover, although collecting data on consumption is usually very time consuming, the concept of consumption is usually clearer than the concept of income. For this reason it is extremely difficult to accurately measure household income, especially for self-employed households and those working in the informal sectors.

Finally, income is likely to be a more sensitive issue for respondents than consumption (Deaton, 1997): there is some evidence that those who are well-off are less likely to participate in the survey or to respond; this results in an underestimation of income inequality among the population (Korinek et al., 2006).

## 2. HOW TO MEASURE CONSUMPTION?

Measuring consumption is a difficult task. However, good practice techniques and guidelines exist which may be considered when trying to construct an accurate measure of consumption (see Deaton and Zaidi, 2002; and ILO, 2003).

In order to obtain a good measure of welfare, consumption should be comprehensive (Deaton and Grosh, 2000); the questionnaire should cover all components of consumption and all types of consumption.

Collecting information only on a subset of consumption could result in bias: as Deaton and Grosh (2000) put it “the relationship between the part and the whole can vary a great deal from one household to another and from one place or time to another”, therefore the omission of some components could affect the ranking.

Consumption usually includes: 1) food consumption, 2) non-food items (including health, education and other non-food expenditures), 3) housing expenditures (including rent and utilities) and 4) consumer durables.

*Food consumption* comprises food consumed inside the household from a variety of sources (food purchases, self-produced food,<sup>1</sup> food received as gifts, remittances and payments in kind) and food consumed outside the household (restaurants etc.).

*Non-food items* refer to education (such as tuition fees, textbooks, etc.), health (medical care and health expenses) and a wide range of other non-food expenses (such as domestic fuel and power, tobacco products, clothing and footwear, transport, recreation, personal care, miscellaneous goods and services ). A choice however has to be made in terms of the items to include. It is usually recommended to include education expenditures,<sup>2</sup> and to exclude taxes and levies as well as gifts and transfers (Deaton and Zaidi, 2002).

The inclusion of health expenditures is debated. As highlighted by Deaton and Zaidi (2002), it is difficult to measure the increase in welfare coming from health expenditures, as information would be needed on the loss of welfare from illness on one hand and on the increase in welfare from its alleviation on the other. If health expenditures only are accounted for, then differences between two sick people – of whom only one is able to pay for treatment – are missed. The recommendation is to include or exclude these expenditures based on the analysis of the elasticity of health expenditures with respect to total expenditure (inclusion is suggested when the elasticity is high).

The inclusion of lumpy and less frequent expenditures – ceremony-related such as marriages and dowries, births, and funerals – is also an issue when trying to capture consumption. These infrequent outlays are usually not included in the consumption aggregate given their ‘idiosyncratic nature and infrequency’ (Beegle et al., 2010:16). Collecting this kind of expenditure is in fact likely to result in a bias as there will be households that have incurred this type of expenditure in the

---

<sup>1</sup> In order to give a value to food consumption that does not go through the market (i.e. consumption from home production) household surveys need to collect information on prices. There are however a number of theoretical and practical challenges in doing so. For further details see Deaton and Grosh (2000); Deaton and Zaidi (2002).

<sup>2</sup> Deaton and Zaidi (2002) recommend the inclusion of education expenditures although they highlight that 1) they could be considered an investment rather than consumption; 2) they refer to a particular point in life time.

reference period and others who did not - although they might have spent on these same items in a previous period (Deaton and Zaidi,2002).

*Housing expenditure* comprises actual rent or rental equivalence value, house repair, decoration and so on. Rents can sometimes be observed directly (i.e. households that rent their dwelling); for households that do not report rents (i.e. owner occupiers), respondents can be asked to provide an estimate (or rental equivalent); another approach is to predict rental payments through the use of imputation models (i.e. by hedonic regressions). Both procedures, however, work well only where an active rental market is in place. Moreover, the credibility of imputation models can be easily compromised when only a small or unrepresentative part of the population rents (Deaton and Zaidi,2002; Deaton and Grosh, 2000).

Finally, another important group of items to consider is *consumer durables*. When dealing with durable goods (such as home, vehicles, washing machine, computers, etc.) what should be computed is not the expenditure itself but the flow of services that they yield. However, in order to compute this flow of services for durable goods, information is needed on the age of each durable good as well as on its original and current value; in practice, estimating the value of service flows also involves crucial assumptions such as definition of durable good, depreciation rate of different items and so on (Deaton and Zaidi,2002; Deaton and Grosh, 2000).

### 3. HOW DOES SURVEY DESIGN INFLUENCE CONSUMPTION MEASUREMENT?

This section sums up the most important issues that should be taken into account when designing a consumption questionnaire. First, a decision should be made on the **method of data capture** – namely on whether to use a *diary versus a recall approach*. The use of diaries is considered to reduce memory lapse as – in principle – this allows for items to be recorded immediately after purchase; however, it implies that the person filling in the diary should be literate (Deaton and Grosh, 2000).

Second, the **reference period** should be defined. There are no definite guidelines on the optimal period for which consumption should be reported; the reference period might range from three days up to one year. There are two main types of reporting error related to the choice of the reference period. On one hand, it has been shown that as the recall period increases, progressive forgetting leads to an increasing underestimation in consumption (recall bias/errors, omitted expenditures). On the other hand, some studies have found evidence of telescoping errors as respondents include expenditures that happened outside the recall period (upward bias) (Deaton and Grosh, 2000).<sup>3</sup> Even when no reporting error is made, another issue is that the use of short recall periods (i.e. one day) has been found to increase the likelihood of zero purchases for some items as there may be people who spend nothing on a given day. This however is not indicative of the true extent of their welfare and, when only single visits are made, this approach results in low expenditure in some households (Deaton, 1997). In general, in an attempt to balance potential recall errors from long reference periods with potential variance from short periods, most surveys apply the following general guidelines: short periods better suit high frequency purchases and longer periods better record rarely-purchased items (Deaton, 1997; Deaton and Grosh, 2000).

It should however be highlighted that expenditures for any given household vary from one reporting period to another; this variation is explained by intra-household components (variation for each household over the year such as seasonal patterns but also random non-seasonal fluctuations – short-term idiosyncratic shocks) and inter-household components (genuine differences in annual expenditure across households). However, even with a fairly long reference period (for instance, two weeks), the questionnaire is not able to capture seasonality or other fluctuations in consumption over the year (Deaton, 1997).

Third, a choice should be made concerning the **respondent**. It should be considered whether to interview a single respondent – usually the person most knowledgeable about consumption expenditures of the household – or multiple respondents. Ideally, multiple respondents should be interviewed as the use of a single respondent might result in a loss of accuracy in capturing individual consumption by household members that the survey respondent does not observe directly: however, collecting information on individual consumption (multiple respondents) is a far more complex and expensive process than gathering consumption at the household level (single respondent) (Deaton and Grosh, 2000; Beegle et al., 2010).

---

<sup>3</sup> Neter (1970) distinguishes between sampling and non-sampling related errors. Among the latter category the following can be distinguished: recall errors (related to memory); telescoping; respondents' fatigue; 'prestige' errors (misreporting due to various social pressures); conditioning effects (from being in the survey); respondent effects (where the identity of the respondent affects the answer that they give); interviewer effects and effects associated with the design of the instrument.

Finally, the ***degree of commodity detail*** can vary greatly and ranges from less than 20 to over 400 items. The level of disaggregation is an important determinant of the cost as well as of the accuracy of the consumption module. There is also some evidence that the more categories of consumption respondents are asked to report, the higher the consumption estimate. The organization of the module could be done in different ways by grouping items: by the type of the item, by the place of purchase, by recall period or by the kind of follow-up questions to be asked (Deaton and Grosh, 2000). A decision should also be made on how to report values: actual amount (continuous variable) versus specified intervals<sup>4</sup> (categorical variable). The order, or the sequence of questions, does not seem to influence the answers given by respondents.

These issues are of paramount importance as they are considered to determine the accuracy of the consumption estimates.<sup>5</sup> In addition, the design will affect the costs and the time required to complete the interview/questionnaire.

It is therefore extremely important to understand how consumption estimates are influenced by survey design. The following section reviews in detail the literature on the degree of commodity detail in consumption surveys.

---

<sup>4</sup> Making the respondent choose from an expenditure list of pre-selected categorical responses and intervals increases the risk of measurement error (Sulla and Tiongson, 2008). Household consumption estimates are likely to be biased given no information on the distribution of consumption within these intervals is available.

<sup>5</sup> Also, these issues are closely interlinked and should not be taken separately. For instance, the survey design defines the possibility of multiple visits and therefore influences the decision on whether to use a diary versus a recall approach. The method of capture used also interacts with the choice between recall periods. Moreover, when designing a consumption module to be included in an existing household survey it is important to make sure that other parts of the survey do not already collect some of the information needed (Deaton and Grosh, 2000).

#### 4. SHORT VERSUS LONG CONSUMPTION MODULES – A LITERATURE REVIEW

The advantages of a short consumption module are potentially enormous. As Sulla and Tiongson (2008) summarize, there are several reasons why a condensed questionnaire could result in significant monetary savings (Hentschel, 2004). The interview time is expected to be shortened. Furthermore, lower respondent fatigue should positively affect the response rates (lower non-response) (Sulla and Tiongson, 2008; Beegle et al., 2010). Finally, adding a short consumption module in already existing surveys (general-purpose, health focused, etc.) could significantly increase its usage for policy analysis and research purposes (Browning et al., 2003) and could provide a more viable option for war-afflicted countries (Mullen, 2006).

A strand of the literature has compared short consumption/expenditure<sup>6</sup> modules with longer and more detailed ones in order to test the accuracy of the resulting estimates. However, the literature is still limited and few field experiments are available in developing countries. The literature review mainly focuses on studies from the 1990s and does not include developed countries studies.

According to a number of studies (Jolliffe, 2001; Pradhan, 2001; Steele, 1998; Lanjouw and Ravallion 1996; The Statistical Institute and Planning Institute of Jamaica, 1994; Beegle et al., 2010), abbreviated expenditures/consumption modules are systematically biased downwards, resulting in lower levels of recorded expenditures/consumption and therefore less accurate estimates and higher absolute poverty rates. The underlying idea is that a longer and more detailed list of commodities prompts respondents to recall more precisely their expenditure behaviour (Beegle et al., 2010).

Jolliffe (2001) uses data from a between-groups design experiment to test whether the length of consumption questionnaires affects consumption and poverty estimates.

Two consumption questionnaires – designed to elicit the same full definition of consumption and differing only in the extent to which consumption items are aggregated – were integrated within the national household survey of El Salvador (EHPM - Encuesta de Hogares de Propósitos Múltiples de El Salvador) in 1994 and administered to two non-overlapping samples. A longer and more detailed questionnaire – covering 72 food items and 25 non-food items – was canvassed on around a quarter of the 4,229 households interviewed whereas the remaining households were administered a shorter module asking about the consumption of 18 categories of food items and 6 non-food items. The recall period for the initial prompting questions varies across questionnaires.

Jolliffe's analysis indicates that although the 2 sub-samples are identical in all respects,<sup>7</sup> they vary in terms of consumption levels. The estimate of mean per capita household consumption from the long questionnaire is 31 per cent higher than the one from the short module; the same pattern is found when examining the food and the non-food components of total household consumption (respectively 20 per cent and 88 per cent higher) and when looking at average consumption levels by decile (between 25 per cent and 45 per cent higher).

---

<sup>6</sup> The review focuses not only on consumption modules but also on expenditure modules.

<sup>7</sup> The author tests for the similarity of the two sub-samples in characteristics that are expected to be correlated with consumption: this test ensures that those households that received the short questionnaire are similar to those that received the long questionnaire.

The stochastic dominance analysis shows that the consumption measure from the long questionnaire is significantly greater than the one resulting from the short questionnaire<sup>8</sup> whether measured as total, food or non-food household consumption. This result entails that the short questionnaire will generate higher estimates of absolute poverty (incidence and depth of poverty) with respect to the longer one.

Jolliffe (2001) also computes the 3 Foster-Greer-Thorbecke (FTC) poverty indices (headcount, poverty gap, and squared poverty gap) for each module using absolute (reference and severe) poverty lines. Absolute poverty<sup>9</sup> derived from the short questionnaire is statistically significantly higher than the one from the long module.

Another study, Pradhan (2001), uses the Indonesian national socio-economic survey (Susenas) data from 1993, 1996 and 1999 to test whether the level of aggregation in consumption questionnaires affects the quality of consumption data. Every three years in fact, the Susenas collects – apart from various socio-economic indicators – data on household consumption using both a short (*core*) and a long (*module*) questionnaire. The households sampled are around 200,000 every round; the *module* (long) consumption questionnaire is administered to around a third of the total sample, whereas the remaining households are administered only the *core* consumption questionnaire.

The module questionnaire collects the expenditures and value of home production for 218 food items with a reference period of one week and the value of consumption for 102 non-food items using two different reference periods (one month and one year).

The items from the module questionnaire can be aggregated into 15 food and 8 non-food broad categories. The *core* questionnaire collects the value of consumption only for these 15 commodities.<sup>10</sup> The reference periods are the same as in the module.

The author compares the consumption measures obtained from the two samples for the three years: the differences in the means between the two modules indicate that the higher level of aggregation of the short questionnaire results in an underestimation in average consumption by between 12 per cent and 20 per cent (depending on the year<sup>11</sup>) relative to the longer and more disaggregated module. Non-food expenditures are more severely underestimated (between 23.8 per cent and 30.1 per cent) whereas the degree of underestimation in food consumption ranges between 3.5 per cent and 11.4 per cent. The differences in mean consumption between the *core* and the *module* questionnaires increase by quartile: among the wealthier groups the non-food share in consumption is indeed higher.

Pradhan (2001) also shows that the divergence in consumption estimates is correlated with the level of total expenditure. Using a pseudo cross-section approach to eliminate the effect of measurement error it results that using an abbreviated consumption module not only yields lower estimates but the degree of underestimation also increases with consumption; a one per cent increase in average consumption increases the degree of consumption underestimation by about 0.4 percentage point.

---

<sup>8</sup> The long questionnaire sample first and second-order stochastically dominates the short questionnaire.

<sup>9</sup> In terms of both the reference and severe poverty line.

<sup>10</sup> The value of home production is not reported in the short questionnaire.

<sup>11</sup> The year that reports the highest underestimation is 1996 which is also the year in which consumption peaked.

Pradhan also carries out comparisons of consumption means at a more disaggregated level.<sup>12</sup> Among food categories, high underestimations are found for vegetables, fruit and prepared foods (which accounts for 15.6 per cent of the total difference in consumption). However, most of the difference in total consumption (68 per cent) is driven by differences in non-food consumption. In particular, for housing, ‘miscellaneous goods and services’ and durable goods, the average *module* consumption turns out to be higher than that of the *core* with an underestimation of 27, 54.6 and 45.5 per cent respectively. However, housing contributes the most to the total difference in consumption (26 per cent against 12 and 18 per cent for durables and ‘miscellaneous goods and services’ respectively). Pradhan also reports overestimation of education expenditures (around 17 per cent). Overall, findings are consistent across the reference periods used for non-food (one month and one year).<sup>13</sup>

Pradhan also investigates the consequences of using a more aggregated consumption questionnaire in terms of poverty measurement. Absolute poverty estimates derived from the short consumption module are higher than those based on the long one also when constructing an adjusted poverty line for the *core* questionnaire. Therefore, even when the author controls for the systematic underestimation in the short consumption measure, higher measurement error of the short consumption module relative to the long one still results in higher poverty estimates.

The Statistical Institute and Planning Institute of Jamaica ran an experiment in 1994. In the JSLC (Jamaica Survey of Living Conditions) two shortened versions of the standard household consumption expenditure module were tested. The experimental consumption modules were: a shortened item-by-item consumption module and a point-of-purchase module.

The *shortened item-by-item* consumption module was derived from the standard consumption module by aggregating items within categories. The total number of items<sup>14</sup> was reduced from 119 in the standard module to 37 items; daily expenditures from 6 to 4; food expenditure from 55 to 13 (food expenses included home production and food received as gift); consumption expenditure from 49 to 15; and non-consumption expenditure from 9 to 5. The same reference period was applied to the shortened and standard modules.<sup>15</sup>

The *point-of-purchase* module collects information on the amount spent at different shops, rather than the item-wise expenditure.<sup>16</sup> The questions on home production and gifts are omitted as

---

<sup>12</sup> For 1996, the year in which the difference was the largest.

<sup>13</sup> Non-food consumption is underestimated slightly more (30.1 per cent) when using a one-year reference period with respect to the one month reference period (27.8 per cent).

<sup>14</sup> Commodity groups covered are: food and beverages; fuel and household supplies; household operational supplies; other housing expenses; housing durable goods; personal care; health care; clothing and footwear; transportation; education; recreation; miscellaneous consumption.

<sup>15</sup> Past 7 days for daily expenses; past 7 days and past 30 days for food consumption; past 30 days and past 12 months for non-food consumption.

<sup>16</sup> The point-of-purchase module was divided into 4 blocks:

*Block m1:* meals consumed away from home in the last 7 days and last 30 days; purchases in the last 7 days and last 30 days at all supermarkets and other self-service establishments, green groceries, meat shops, liquor/beverage shops, bakeries, market/street vendors, vegetable and fruit shops, other retail stores dealing in food, beverages and tobacco and household supplies, wood/charcoal vendors, petrol/gasoline retailers and kerosene vendors;

*Block m2:* cooking gas;

- *Block m3:* purchases in the last 30 days and last 12 months made at textiles and apparel shops, general purpose stores, pharmacies and drug stores, furniture, furnishings etc. stores, household appliance stores, market/arcade vendors, book, stationery and newspaper vendors, sports goods and other retail stores;
- *Block m4:* payments during the last 30 days and last 12 months for motor car expenses, educational services, personal services, repair services, and other services.

accounting for a relatively small part of total consumption.<sup>17</sup> The reference period in this case is either the past 7 days and 30 days for more frequent purchases, or 30 days and last 12 months for non-frequent purchases.

The point-of-purchase module was canvassed on half of the households also covered by the standard consumption module<sup>18</sup> (that is 1,248 dwellings out of total 2,496 sample). However, the shortened item-by-item consumption module was administered to a different sample (the other half – 1,248 households) which was not administered the standard consumption module.<sup>19</sup>

The analysis carried out is basic and limited to comparisons in mean consumption. On one hand, findings relative to the shortened item-by-item consumption module indicate that the estimates of mean per capita consumption are around 20 per cent lower than those derived from the standard consumption modules.<sup>20</sup> The shares of each commodity group over total consumption were also all lower except for household operational expenses and health care. Most of the divergence in mean per capita consumption is accounted for by the ‘food and beverages’, ‘clothing and footwear’ and ‘transportation’ group.

Findings relative to the point-of-purchase module show that the estimates of mean per capita consumption are similar to those from the standard consumption module<sup>21</sup> although home production and gifts consumption had been excluded.<sup>22</sup> There are however substantial differences when comparing estimates at the household level for a considerable number of households. Although the study represents an interesting pilot/field experiment there is no in-depth analysis of the results.

Two more studies (Steele, 1998; Lanjouw and Ravallion, 1996;) find that more detailed questionnaires result in higher mean consumption estimates. As in the previous cases, these studies differ, however, in that they do not compare long questionnaires with aggregated ones which condense items from the long module into broad comprehensive categories. Rather they compare consumption questionnaires that differ in terms of their length but that do not necessarily aim at covering exactly the same comprehensive definition of consumption.

In 1993, in Ecuador, two food consumption modules were piloted. Respectively these included questions on 122 and 72 food items. The analysis conducted by Steele (1998) indicated that the ratio of food expenditures for the long to the short module was 1.67 (as cited in Beegle et al., 2010; Deaton and Grosh, 2000).

Another Ecuadorian experience seems to support the finding that the inclusion of extra items in consumption modules result in higher consumption estimates and lower poverty estimates. This study however differs from the previous ones as it is a longitudinal study (Lanjouw and Lanjouw,

---

<sup>17</sup> The questions on housing expenses remained as a part of the housing module.

<sup>18</sup> The test was based on the assumption that administering the standard consumption module and the experimental consumption module based on the point of purchase would not have affected the reliability of either estimate. The standard and experimental point-of-purchase modules were administered in a different (but systematic) order in different households so as to be able to study whether this results in any bias.

<sup>19</sup> The shortened item-by-item consumption module – as derived from the same pattern of commodity groupings as the standard modules though with further aggregation of the items – was thought to influence the reported consumption figures.

<sup>20</sup> 20.8% lower in Jamaica, 22.2% lower in KMA (Kingston Metropolitan Area), 20.3% in other towns, 19% for rural areas. The differences in means are statistically significant.

<sup>21</sup> None of the differences was statistically significant.

<sup>22</sup> The order of canvassing the modules had no effect on the estimates.

2001). In 1994 and 1995, two LSMS surveys were carried out in Ecuador. The definition of the consumption aggregate however had been modified over the years: the food consumption module in 1994 consisted of 73 items compared to 94 in 1995; several non-food items had also been added in 1995. The incidence of poverty passed from 52 to 45 per cent over the period. This drop in poverty appeared startling given the short period of time under review, the sluggish economic growth registered and the absence of new policies aimed at poverty reduction. Lanjouw and Lanjouw (2001) elaborated a technique<sup>23</sup> to produce poverty estimates comparable across surveys and over time and robust to changes in survey design (i.e. changes in the consumption definition); once they applied it to Ecuador, they found that the proportion of people living below the poverty line seemed to have actually increased from 52 per cent to 56 per cent over the period 1994-5. Therefore, findings showed that the apparent poverty decline was the result of modifications in survey design, namely the extension of the item list.

A recent study (Beegle et al., 2010) has distinguished itself from the available strand of literature, comparing estimates from different consumption questionnaires – and not only short ones – with a benchmark. This authoritative study confirms that shorter consumption modules bias consumption estimates downwards; however what seems particularly interesting is that one of the two short modules tested (the subset module) is found to result in a downward bias smaller than that of other longer consumption modules.

Beegle et al. (2010), during a survey experiment called SHWALITA (the Survey of Household Welfare and Labour in Tanzania), randomly assigned to 4,000 households in Tanzania eight alternative consumption questionnaires (five recall and three diaries) which differed greatly in terms of survey design.<sup>24</sup> The authors then tested the sensitivity of consumption estimates and related poverty and inequality (distributional) measures to different survey designs.

This study compares estimates from different consumption questionnaires – not only short modules – with a benchmark. This is one of the two reasons that make this paper stand out in the literature. The second reason is related to the type of benchmark used; in fact, unlike the majority of papers where the benchmark used for comparison was a long questionnaire, in this study it is a frequently-supervised personal diary,<sup>25</sup> which it is considered will result in consumption estimates that are the closest to actual consumption.<sup>26</sup>

Our review of Beegle et al. (2010) will focus mainly on the comparisons of two short consumption modules to the benchmark. The two short recall modules – a ‘subset’ (module 3) and a ‘collapsed’ one (module 4) – collect information on food and non-food expenditures.

---

<sup>23</sup> Further details on the methodological aspects of the technique used can be found in Lanjouw and Lanjouw (2001).

<sup>24</sup> Differences are reported not only in the number of items in the recall list but also in the method of data capture, respondent interview, length of reference period, and nature of the cognitive task for the respondent. The consumption modules tested are: 1) a long list (58 food items) 14-day recall; 2) a long list (58 food items) 7-day recall; 3) a subset list (17 food items; subset of 58 foods) 7-day recall; 4) a collapsed list (11 food items covering universal food categories) 7-day; 5) a long list (58 food items) usual 12-month recall (these five recall modules do not differ in terms of non-food section); 6) household diary with frequent visits, 14-day diary; 7) household diary with infrequent visits, 14-day diary; 8) the benchmark is a personal diary with frequent visits (14-day diary).

<sup>25</sup> Beegle et al. (2010) use a 14-day personal diary with frequent visits as a benchmark consumption expenditures module. The personal diary is kept by each adult household member whereas child consumption expenditures are recorded in the diaries of the adults who are most knowledgeable about children’s daily activities. A consumption measure is obtained by adding up everything brought into the household through harvests (own production), purchases, gifts (or other sources), and stock reductions and subtracting everything that left the household through sales, gifts and stock increases. In order to avoid double counting there is intensive supervision: interviewers are trained to double-check for similar items recorded on the same day by different household members.

<sup>26</sup> A frequently supervised diary is meant to account for personal outside-the-house consumption as well as minimizing recall and telescoping error due to the frequent supervision.

In the 'subset' module, food expenditures (inside the household) include a subset list of 17 food items. The 17 food items included are those that constitute, on average, 77 per cent of food consumption expenditure in Tanzania based on the previous Household Budget Survey and are therefore the most frequently consumed foods. For each of the 17 items, the respondent is asked to report the quantity 1) consumed, 2) purchased, 3) coming from own production and 4) from gifts and other sources in the past 7 days; whereas the amount spent is collected only for purchases.<sup>27</sup>

Food expenditures (inside the household) in module 4 are a collapsed list where the 58 food items from a longer module are aggregated into 11 comprehensive categories.<sup>28</sup> Respondents are asked to report whether or not they have purchased, consumed from own harvest or gifts any of the category items; the respondent is then asked to provide a monetary value (for purchases, consumption from home production and gifts including consumption from stock). The reference period used here is the same as in the 'subset' module (7 days). Both modules also include an identical short section on food expenditures outside the household.<sup>29</sup>

The non-food consumption section does not differ across the two modules under review. Non-food items are divided into two groups based on the frequency of purchase. Frequently purchased items (8 categories) were collected by 14-day recall whereas non frequent purchases are collected by one month (3 categories) or 12 month (21 categories).

In order to compare the consumption estimates from module 3 with other modules (and the benchmark), the authors scale up food expenditures for that module. Reference periods are the same across the 2 short consumption modules.

The authors first compare means and medians of (total and component<sup>30</sup>) per capita consumption by module. In general, all 7 modules underestimate consumption, and in particular food consumption, relative to the benchmark.<sup>31</sup> In particular, module 4 (7-day collapsed list) records the lowest mean and median for consumption. However, once rescaled, module three (the subset list) has the highest mean total consumption after the benchmark.<sup>32</sup> This is an interesting result as it actually shows that a short consumption module underestimates consumption compared to the frequently supervised personal diary benchmark but does so less than other long recall modules.

Regressing consumption (total or by component) on dummy variables that indicate the module type (the benchmark module is the reference - or left-out - category), the findings above are confirmed.<sup>33</sup> The results indicate that module 3 (the subset list questionnaire) underestimates by only 7 per cent (the minimum across all modules) whereas module 4 (the collapsed list) results in 28 per cent less consumption (the highest underestimation across modules) relative to the

---

<sup>27</sup> This is possible as SWALITA also collects price data.

<sup>28</sup> The 11 categories are the following: Cereals and cereal products; Tubers and plantains; Sugar and sweets; Pulses; Nuts and seeds; Vegetables; Fruits; Meat, meat products, fish; Milk and milk products (excl. fats); Oil, fats, spices; Beverages.

<sup>29</sup> Respondents report a monetary value of 7 categories eaten/drunk outside the household (consumption of guests is excluded).

<sup>30</sup> Total consumption; food consumption; non-food consumption (frequent; non-frequent).

<sup>31</sup> The non-food expenditure section is identical across the five recall modules in terms of wording and structure.

<sup>32</sup> The 7-day long list recall (module 2) has the highest mean (total) consumption, higher even than the personal diary. However, the subset list results in a higher mean consumption than the following modules: long 14-day, long usual 12-month, household diary frequent, household diary infrequent, and collapse 7-day.

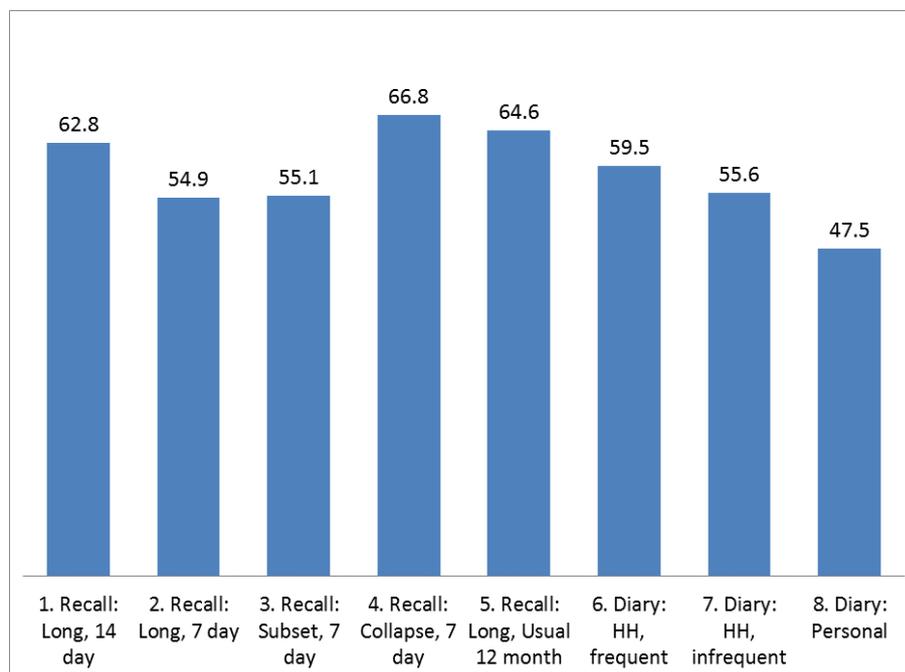
<sup>33</sup> The coefficients can be interpreted as the per cent deviations in mean consumption from the benchmark (example: mean consumption collected with module X is n% lower than mean consumption collected using the benchmark module).

frequently supervised personal diary benchmark. A similar pattern (and magnitude of impact) is recorded for food consumption.

The analysis is then performed for a homogenous subsample (3 recall modules with the same reference period differing only in the length of the food section) rather than pooling all the module types together, and the long 7-day recall module (58 food items) is used as the reference (left-out) category. The findings, in line with the previous results, indicate that total consumption as measured by the collapsed list module is 32 per cent lower than the reference (left-out) category and this is driven from an underestimation of food consumption; however, the subset module – once scaled – only slightly underestimates consumption (by 6 per cent) with respect to the full detailed listing.

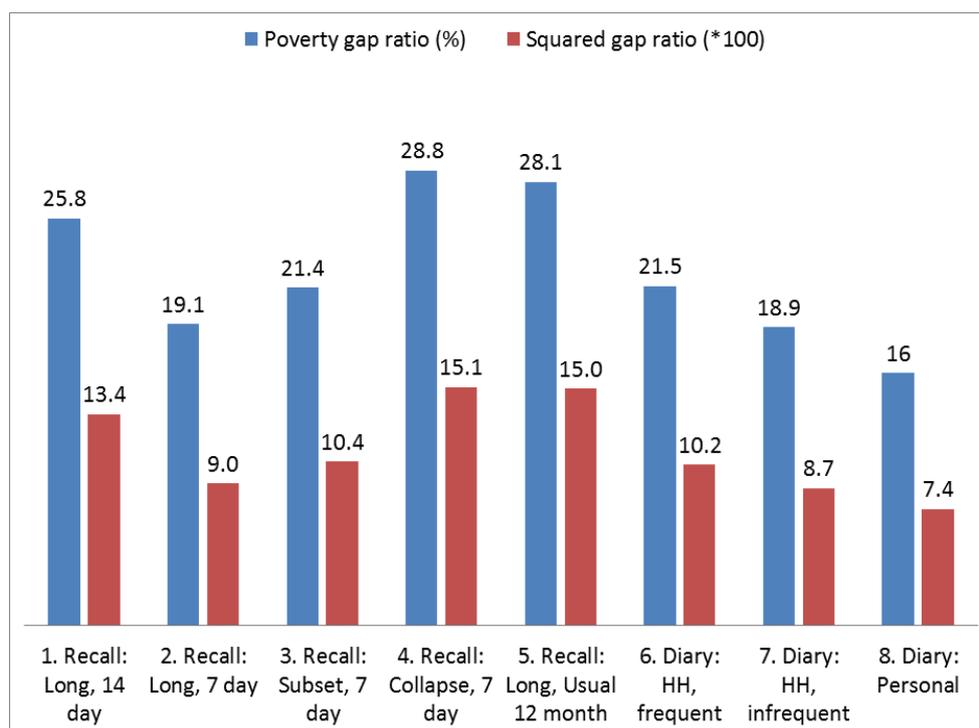
The authors finally investigate the implications for poverty analysis. They compute the poverty headcount rate (an absolute poverty line is fixed) for each module: the benchmark provides the lowest level of poverty followed by module 2 (long list – 58 food items) and module 3 (subset). Module 4 yields a much higher (and the highest overall) poverty headcount (see Fig. 1). A similar pattern is found when comparing two other poverty measures (the poverty gap ratio and poverty squared gap ratio) across modules: module four diverges most from the benchmark (see Fig. 2).

Figure 1: Poverty headcount rate (%) at \$1.25/person/day by module type



Source: Beegle et al., (2010)

Figure 2: Poverty gap ratio (%) and squared gap ratio (\*100) by module type



Source: Beegle et al., (2010)

Overall, the collapsed module does not seem to perform well with respect to the benchmark and the authors do not suggest utilizing this option. However, it could be argued that this result may be driven or biased by the nature of the consumption module questions rather than by its aggregation/length only. Instead of asking – as most modules do – for the monetary value for purchases, home production and gifts separately, the collapsed module asks the respondent to indicate whether or not any of these categories was consumed in the household (yes or no answer) and only then to provide the overall monetary value for all the above-mentioned categories; this is likely to require a different cognitive effort. On the other hand the subset list, once rescaled, is considered ‘a suitable substitute to longer list recall modules’.<sup>34</sup>

Two studies support the idea that short consumption modules might result in a minimal loss of accuracy (WB, 1992; The Statistical Institute and Planning Institute of Jamaica, 1994 – point-of-purchase module). As already seen, in the experiment run by the Statistical Institute and Planning Institute of Jamaica, the point-of-purchase module reports similar means and variances to the standard consumption module; the analysis undertaken however is very simplified and limited in scope.

The core/module design of the Susenas – already reviewed in Pradhan (2001) above – had initially been field tested in 1991 (before its formal introduction into Susenas) in three provinces of Indonesia in order to understand how the short consumption module performed with respect to

<sup>34</sup> “On the other hand, the other short commodity list design (the subset list module 3), when scaled up based on reference data, performs very close to the longer list form and may be a suitable substitute to longer list recall modules. The gain from such a reduction in list length, however, is slight – less than 10 minutes of interview time – and researchers would have to decide whether the loss of additional detail in consumption information is worth the moderate reductions in interview time.” (Beegle et al., 2010:31).

the long questionnaire. 8000 households were randomly assigned either a short or a long questionnaire in two provinces (WB, 1992,<sup>35</sup> annex 4.2 as cited in Pradhan, 2001).

The long consumption module included 218 food items and 102 non-food items that were reduced to 15 and 8 respectively in the short questionnaire (as cited in Deaton and Grosh, 2000).

Total food consumption was very similar across the short and the long questionnaire, both in terms of mean (the difference in mean food consumption was less than 1 per cent) and of distribution (as cited in Pradhan, 2001; Deaton and Grosh, 2000). The mean non-food consumption – based on a one-month reference period as collected by the core questionnaire – was underestimated by less than 15 per cent with respect to the module one; the divergence was mainly driven by the housing and “goods and services” categories.<sup>36</sup>

Deaton and Grosh (2000) report another short questionnaire experiment in a developing country that seemed to yield reasonably accurate consumption data. However, as it was not possible to retrieve the original reference, the results reported here are based on Deaton and Grosh (2000).<sup>37</sup> In a small-scale field experiment of 44 households in 2 villages in West Bengal (Bhattacharya, 1963 as cited in Deaton and Grosh), it was found that consumption estimates using broad groups/items were only slightly (but not significantly) lower than detailed questionnaires. A single question produced 25-30 per cent lower estimates although still highly correlated (0.98) with the estimates from the full list.

*To recap (results in terms of consumption estimates and absolute poverty), shorter modules tend to bias consumption downwards, and non-food consumption is usually underestimated more than food consumption. There is also some indication that ‘using a high level of aggregation yields a lower consumption measure and the fraction of underestimation increases as consumption rises’ (Pradhan, 2001): this supports the idea that short consumption modules may be more suitable for low income countries rather than for advanced economies. Higher absolute poverty rates usually follow from an underestimation in consumption estimates. However, some recent evidence indicates that some short consumption modules underestimate less than longer recall consumption modules when compared to a benchmark.*

Even less evidence is available in terms of the impact of short versus long consumption modules on distributionally sensitive measures. Studies have compared relative poverty estimates, inequality measures (such as Gini index, etc.) and, more generally, the ranking. Available evidence is scarce as many of the studies reviewed in the previous section do not provide ranking or inequality analysis. However, the most sound and authoritative studies indicate that short consumption modules are good at capturing the relative ranking of households.

Jolliffe (2001) computes the 3 Foster-Greer-Thorbecke (FTC) poverty indices (headcount, poverty gap, and squared poverty gap) for each module using relative (half of median and half of mean)

---

<sup>35</sup> The authors did not have access to the primary source and are referring mainly to Pradhan (2001) and Deaton and Grosh (2000). Information on the study is sometimes contrasting across the two sources: Deaton and Grosh report, for instance, that *both* a short and a long questionnaire were administered to the same household. However, as indicated in the following footnote, this biased the results and the estimate differences obtained are therefore not reliable.

<sup>36</sup> Pradhan (2001) also reports that in the third province, where households had been administered both the Core and the Module questionnaire, a simple linear regression yielded an R squared of 85 per cent. However, it should be noticed that administering both the long and short consumption modules to the same household is likely to create a bias.

<sup>37</sup> Moreover, as stated clearly at the beginning of the section, the review of the literature focuses on studies from the 1990s.

poverty lines. The level of *relative poverty*<sup>38</sup> does not significantly differ, from a statistical point of view, across modules suggesting that the distributions have similar shapes. This finding seems to be supported by the comparisons of Gini inequality indices as there are no statistically significant differences in inequality estimates (at the national, regional and urban/rural level) between the long and the short questionnaires. However, the analysis of the geographic distribution of poor persons (in relative terms) shows statistically significant differences across the 2 sub-samples.<sup>39</sup>

Pradhan (2001), using the Indonesian national socio-economic survey (Susenas) data, investigates the consequences of using a more aggregated consumption questionnaire in terms of ranking. He generates concentration curves for a number of socio-economic indicators<sup>40</sup> using separately the long and short consumption measures. The curves lie very close to each other indicating that the rank that an individual holds in the consumption distribution is similar across both consumption measures<sup>41</sup> (*core and module*).

Pradhan also finds that the Gini coefficient is slightly higher for the longer questionnaire when computed based on total or non-food consumption, whereas the short questionnaire provides a lower inequality measure of food consumption (Gini).<sup>42</sup>

Beegle et al. (2010) investigate how household characteristics affect reporting, interacting the module type dummy with the household characteristic under review (within the same regression framework discussed above). The subset module (module 3) reports significantly less consumption than does the benchmark as household size increases and for households with more adult members; like other recall modules, the subset and collapsed list module suffer from recall error.<sup>43</sup> Asset-poor households seem to underreport compared to the benchmark whereas asset rich households report the same or greater consumption than the benchmark. This could result in overstating consumption inequality. However, the authors also analyze the differences in distributional measures: they report Gini coefficients for each module over the four consumption categories (total, food, non-food: frequent and non-frequent) and find that the subset and collapsed list (modules 3 and 4) generally yield a more equal distribution.<sup>44</sup> This is probably a result of omissions; if these are higher among wealthier groups then it will result in a truncated distribution at the upper end.

---

<sup>38</sup> When the poverty line is set either at half the mean or half the median.

<sup>39</sup> The short questionnaire indicates that the large majority of poor persons are urban residents (59%), whereas the data from the long questionnaire indicate that the majority of poor persons reside in rural areas (52%). Moreover, the short questionnaire indicates that the highest incidence of poverty is in the Eastern region whereas the long questionnaire indicates that the highest incidence of poverty is in the Western region of El Salvador. Therefore, although relative poverty rates are similar across modules, further analysis suggests that the two consumption modules may result in different rankings.

<sup>40</sup> The socio-economic indicators are: 'school enrolment for children in the junior and senior secondary age group, whether the respondent sought medical care in the past month and whether young children received a complete set of vaccinations'.

<sup>41</sup> 'If the structural underestimation is a monotonically increasing function of Module consumption (as found in Figure 3), it will not change the ranking. The measurement error causes a reclassification, resulting in a weaker relation between consumption and benefits. Since the underestimation is higher for the rich, the effect of the measurement error will be less for the higher income groups.' (Pradhan, 2001: 20)

<sup>42</sup> Pradhan also investigates the consequences of using a more aggregated consumption questionnaire in terms of elasticity of a non-consumption welfare measure with respect to per capita consumption ('gradient' analysis). Higher consumption elasticities are obtained in the gradient analysis when using the short questionnaire ('because the short consumption measure rises more slowly than the long as true consumption rises, the estimate of  $\gamma$  will be an overestimate' Pradhan, 2001: 21); however, the overestimation of the short relative to the long questionnaire varies considerably according to the (non-consumption) welfare measure used (ratio between 0.60 and 0.92).

<sup>43</sup> Beegle et al. (2010: 21) explain the finding: underestimation 'may be due to increased cognitive demands of one respondent asked to recall the consumption of an entire household as the number of members increase. Alternatively the relative importance of out-of-household consumption may increase with household size, particularly for frequent non-food expenditure, which is often privately consumed (e.g., cigarettes, cell-phone top-ups, bus and taxi fares) and this consumption is systematically missed by reliance on a single household respondent.'

<sup>44</sup> 'For both module types, this may not be surprising since the relative lack of prompts among the aggregate categories is expected to yield compressed consumption measures due to omissions. If the diversity of consumption goods is greater among wealthier households, this error type will lead to truncated distributions at the upper end.' (Beegle et al., 2010: 24)

Finally, the authors also try to investigate the implications for poverty analysis in terms of relative ranking of the households. The authors regress a poverty indicator (binary dependent variable as well as continuous) on module type, a household characteristic under review and interaction term between poverty indicator and the household characteristic (within the same regression framework as discussed above). Results suggest that although different module types yield highly different poverty measures, the characteristics of poor households do not vary across modules. The authors conclude saying that: “these modules, as virtually all other modules, yield household consumption rankings that are remarkably stable with respect to key household characteristics, indicating that analysis that focuses on the determinants of household ranking within the consumption distribution will be largely consistent regardless of which module is used.” (Beegle et al., 2010: 31)

A different but relevant study is provided by Morris et al. (2000). The authors provide an approach to identify a short list of expenditure items that, when aggregated, correlates highly with total expenditures. The aim of the paper is therefore not that of estimating correctly/accurately the level/value of expenditures but rather to find a proxy that mirrors as closely as possible true consumption and preserves the household ranking.

Using this technique, the authors are able to proxy household expenditures without significantly expanding the length of the questionnaire. Two national representative surveys of the rural areas of Côte d’Ivoire for 1986 and 1989 are used. The consumption module includes questions on 34 food items (both in terms of purchases and home production) with a 2-week recall period and 39 non-food items with daily and annual recall period.

Once the total household expenditure (annualized) for each household is computed, the max\_r procedure is used to capture the 10 items that better mirror total household expenditure and its distribution (better ‘preserving the relationships between households’ Morris et al., 2000: 386)<sup>45</sup>. The algorithm maximises the correlation  $r$  between the proxy measure (the sum of 10 selected expenditure items) and the true measure (total household expenditures).

In rural areas of Côte d’Ivoire, the sum of the 10 selected expenditure items was correlated with the true measure (total household consumption expenditure items) at the  $r = 0.74$  level in 1986 (development data set). The reduced measure (the sum of the 10 expenditure) was also found to correlate highly ( $r=0.72$  level) with total household consumption expenditure in 1988 (the second independent data set or validation data set).

Although this technique could result in time savings in the field, the authors highlight that it might not always be possible to identify such a relationship in every case. Moreover, in order to use this methodology, a recent household consumption module should be available.

As with distributional measures, it is difficult to conclude anything on the relative time savings of a short consumption module compared to a longer and more detailed one. According to the 1992 World Bank study in Indonesia, the core questionnaire – including not only the short consumption questionnaire but also the socio-economic questions – resulted in time savings compared to the

---

<sup>45</sup> This procedure was originally used by Mark et al. (1996) to determine which food items should be included in surveys designed to ascertain the effect of nutrients on disease risk.

module questionnaire: the interviewing time for the core was 52 compared to 82 minutes for the module.

The Statistical Institute and Planning Institute of Jamaica report the interviewing time for the standard JSLC (Jamaica Survey of Living Conditions) questionnaire and the two shortened versions of the standard household consumption expenditure module. The interviewing time for the standard module was 50 minutes compared to the shortened item-by-item module that took on average 37 minutes, and the point-of-purchase one which had an average duration of 29 minutes.

Beegle et al. (2010) in their analysis also consider the average time required to interview one household using different module types. In general, the subset and collapsed modules recall result in an average time saving of 8-9 minutes compared to the recall module that lists 58 foods which requires around 50 minutes.

## 5. CONCLUSIONS

Consumption expenditure is probably the most common and preferred welfare indicator; however, its measurement is a challenging and time-consuming task.

Although short consumption modules have potentially enormous advantages in terms of time and money savings, a recent and comprehensive literature on available experiments comparing short and long modules is still lacking for developing countries. The present paper aims at filling this gap.

A critical review of the available evidence points to a number of factors hindering the ability to draw firm conclusions. *First*, experiments have been scarce and mostly run in single countries in different regions at different points in time. The findings of each paper are dependent on the specific setting and are therefore difficult to generalize. Countries differ not only in terms of their living standards but also in terms of the share of consumption from home production, the balance between inside and outside food consumption, diversity of consumption choices and dietary diversity, and household structures (Beegle et al., 2010); all factors that have a bearing on results.

*Second*, different methodologies have been adopted that are not always easily comparable across studies; some analyses are basic and more superficial whereas others provide a more in depth and extensive investigation of the issue at stake.

*Third*, there is a difference – often overlooked in the literature – between subset and collapsed questionnaires; the latter is aimed at measuring the same definition of consumption by asking about consumption in a number of comprehensive broad categories, while the former uses a reduced number of single items (not comprehensive but expected to mirror total consumption). More generally, in field experiments short and long consumption modules do not always capture the same comprehensive definition of consumption.

*Fourth*, the use of different recall periods between the modules tested acts as a confounding factor.

*Last*, the use of different benchmark modules constitutes another confounding factor in these studies: sometimes personal diaries are used, other times long recall questionnaires and so on; studies rarely provide a rationale on how the long questionnaire has been or should be selected as a benchmark and seem to implicitly assume that any long questionnaire is closer to actual consumption.

The review of the literature, and the above-mentioned factors in particular, indicate that there is room for further investigation. In order to be conclusive, future field experiments should compare short consumption modules – clearly distinguishing between collapsed and subset modules – to a well-defined long benchmark, chosen for its ability to produce consumption estimates that are the closest to actual consumption. Survey design should be the same across the short and long consumption modules: the method of data capture, the level of respondent, as well as the length of the reference period, should be consistent in order to isolate the impact of the length or level of detail in consumption questionnaires on the accuracy of the resulting estimates. It is also fundamental that short and long consumption modules capture the same comprehensive definition of consumption (therefore including exactly the same consumption components).

Several/multiple methodologies (quantitative analysis) should be applied and, where funding allows, different contexts should be explored.

In general, available evidence from developing countries seems to indicate that short modules underestimate consumption with respect to longer ones, resulting in lower levels of recorded expenditures/consumption and therefore less accurate estimates and higher poverty rates. However, one of the most complete, recent and authoritative studies in the field (Beegle et al., 2010), confirms that shorter consumption modules bias consumption estimates downwards although they find that some short modules may actually result in a smaller downward bias, compared to the benchmark, than longer consumption modules.

In particular, they find that the 'subset' module, which includes 17 food items that constitute, on average, 77 per cent of food consumption expenditure, "performs very close to the longer list form and may be a suitable substitute to longer list recall modules" (Beegle et al., 2010:31). Also, the 'subset' module performs much better than the alternative short module, defined as 'collapsed' module, where the 58 food items from a longer module are aggregated into 11 comprehensive categories.

This result has important implications for the design of the short consumption module as it suggests that 'subset' modules, which identify the most common items consumed and are subsequently scaled up, provide more reliable information than modules that use item categories.

In terms of relative ranking of households, results from rigorous studies indicate that household consumption rankings obtained through short consumption modules are largely consistent with rankings derived from long modules.

## REFERENCES

- Azzarri, C., G. Carletto, B. Davis, and A. Zezza (2006). 'Monitoring Poverty Without Consumption Data: An application using the Albania Panel Survey', *Eastern European Economics*, 44(1): 59-82.
- Beegle, K., J. De Weerd, J. Friedman, and J. Gibson (2010). 'Methods of Household Consumption Measurement through Survey: Experimental Results from Tanzania', *World Bank Working Paper Series 5501*.
- Bhattacharyya, N. (1963). "On the Effects of Itemization in the Family Budget Schedule." Indian Statistical Institute, Calcutta.
- Booyesen, F., M. Von Maltitz, S. Van Der Berg, R. Burger, and G. Du Rand (2008). 'Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African countries', *World Development*, 36 (6): 1113-1130.
- Browning, M., Crossley, T. F. and G. Weber (2003). "Asking Consumption Questions in General Purpose Surveys," *Economic Journal*, Vol. 113, No. 491 : pp. 540-67.
- Citro, C. F. and R. T. Michael (1995). *Measuring Poverty: A New Approach*. Washington, DC: National Academy Press.
- Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. The Johns Hopkins University Press, Baltimore and London.
- Deaton, A. and M. Grosh (2000). 'Consumption', Chapter 5 in M. Grosh and P. Glewwe (eds.), *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience*, Washington, DC: World Bank.
- Deaton, A. and S. Zaidi. (2002). 'Guidelines for Constructing Consumption Aggregates for Welfare Analysis', *World Bank LSMS Working Paper* 135.
- Filmer, D. and L. Pritchett (1999). 'The Effect of Household Wealth on Educational Attainment: Evidence from 35 countries', *Population and Development Review* 25 (1): 85-120.
- Filmer, D. and L.H. Pritchett. (2001). 'Estimating Wealth Effect Without Expenditure Data or Tears: An application to educational enrollments in states of India', *Demography* 38 (1): 115-32.
- Hentschel, J. (2004). "Using Rapid City Surveys to Inform Municipal Social Policy: An Application in Cali, Colombia," Policy Research Working Paper Series No. 3369 (Washington: The World Bank).
- Howe, L.D., J.R. Hargreaves, and S.R. Huttly (2008). 'Issues in the Construction of Wealth Indices for the Measurement of Socio-economic Position in Low-income Countries', *Emerging Themes in Epidemiology* 2008 (5): 3.
- Howe, L.D., J.R. Hargreaves, G.B. Ploubidis, B.L. De Stavola and S.R. Huttly (2010). Subjective Measures of Socio-economic Position and the Wealth Index: A comparative analysis. *Health Policy and Planning* May;26(3):223-32.
- ILO (2003). *Household income and expenditure statistics*. Report prepared for the 'Seventeenth International Conference of Labour Statisticians', Geneva, 24 November-3 December 2003.
- Jolliffe, D. (2001). 'Measuring Absolute and Relative Poverty: The sensitivity of estimated household consumption to survey design', *Journal of Economic and Social Measurement*, vol. 27(1/2): 1-23.

- Korinek, A., J. Mistiaen and M. Ravallion (2006). "Survey nonresponse and the distribution of income," *Journal of Economic Inequality*, Springer, vol. 4(1), pages 33-55, April.
- Lanjouw, J. and P. Lanjouw. (2001). 'How to Compare Apples and Oranges: Poverty measurement based on different definitions of consumption', *Review of Income and Wealth* 47 (1): 25-42.
- Lanjouw, J. and M. Ravallion (1996). 'How Should we Assess Poverty Using Data from Different Surveys', *Poverty Lines Newsletter*. The World Bank: September 1996.
- Mark S.D., D.G. Thomas, and A. Decarli (1996). 'Measurement of Exposure to Nutrients: An approach to the selection of informative foods, *American Journal of Epidemiology*, 143: 514-21.
- Montgomery, M., M. Gragnolati, K. Burke, and E. Paredes (2000). Measuring Living Standards with Proxy Variables, *Demography* 37 (2): 155-174.
- Morris, S., C. Carletto, J. Hoddinott, and L. Christiaensen (2000). 'Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa', *Journal of Epidemiology and Community Health*, 54 (5):381-387.
- Moser, C. (1998). 'The Asset Vulnerability Framework: Reassessing urban poverty reduction strategies', *World Development* 26 (1): 1-19.
- Mullen, P. D. (2006). "Measuring Socio-economic Patterns in a Chronic Conflict Situation: Rapid Assessments and a Household Survey in Southern Sudan," *Journal of African Economies*, Vol. 15, No. 3: pp. 470-503
- Neter, J. (1970). "Measurement Errors in Reports of Consumer Expenditures" *Journal of Marketing Research* 7:11-25.
- Pradhan, M. (2001). 'Welfare Analysis with a Proxy Consumption Measure: Evidence from a repeated experiment in Indonesia', Economic and Social Institute, Free University, Amsterdam.
- Rutstein S.O. and K. Johnson (2004). 'The DHS Wealth Index', *DHS Comparative Reports 6*, MEASURE DHS+ Calverton, MD.
- Rutstein, S.O. (2008). 'The DHS Wealth Index: Approaches for rural and urban areas', *DHS Working Papers 60*, Measure DHS, Calverton, MD.
- Sahn, D.E. and D. Stifel (2000). 'Poverty Comparisons over Time and Across Countries in Africa', *World Development* 28 (12): 2123-2155.
- Sahn, D.E. and D. Stifel (2003). 'Exploring Alternative Measures of Welfare in the Absence of Expenditure Data', *Review of Income and Wealth* 49 (4): 463-489.
- Statistical Institute and Planning Institute of Jamaica (1994). *Survey of Living Conditions*, 1994. Planning Institute of Jamaica, Kingston.
- Steele, D. (1998). 'Ecuador Consumption Items'. World Bank. Mimeo.
- Sulla, V. and E.R. Tiongson (2008). 'Consumption Aggregates In Bosnia Herzegovina: A comparative study of survey instruments', Poverty Reduction and Management Unit, Europe and Central Asia Region; The World Bank (Preliminary draft).
- World Bank (1992). 'Indonesia: Public expenditures, prices and the poor', *Report 11293-IND*, Indonesia Resident Mission, Jakarta.

**UNICEF**  
**Office of Research**

UNICEF Office of Research  
Piazza SS. Annunziata, 12  
50122 Florence, Italy  
Tel: +39 055 20330  
Fax: +39 055 2033 220  
[florence@unicef.org](mailto:florence@unicef.org)  
[www.unicef-irc.org](http://www.unicef-irc.org)  
© The United Nations Children's Fund (UNICEF)