Impact of household budget surveys' design on the measurement of food consumption

Piero Conforti, Nathalie Troubat, Klaus Grunberger[±]
FAO of the UN, Statistics Division – Viale delle Terme di Caracalla, 00153 Rome, Italy

Abstract

Surveys characteristics affect the quality of the measurement of food consumption that habitually occurs within the household. It is therefore important to develop guidelines or recommendations on best practices to be adopted when designing a survey to collect food data. This paper presents results on the impact of survey design on food consumption measurement from a panel of about seventy surveys. Results show that food consumption measurement is very sensitive to survey design. Collecting information on acquisitions rather than food consumption increases the estimated level and variability of food consumption. Differences may imply totally different results in certain countries, and seriously bias the planning of food security and nutrition intervention programs.

Keywords: Household budget surveys, survey design, food consumption, food acquisition

1. Background

Food data are collected in National Household Budget Surveys (NHBS), mainly to assess poverty and update consumer price index. In the absence of better information on individual food consumption, these data have also been used to conduct analysis on food security and nutrition. As NHBS were not originally designed to analyze food security and nutrition, their food data collection questionnaires do not always gather the information that is necessary for such purpose.

According to a recent joint study conducted by FAO, the International Household Survey Network and the World Bank (Dupriez et al, 2014), out of about 100 surveys assessed only 13% were judged to be reliable for analyzing food and nutrition security. Various criteria were considered to assess the suitability of food consumption data for food security analysis: the reference period during consumption was reported, the time frame of the survey, the mode of data capture (recall or dairy), whether or not food collected refers to acquisition or actual consumption, the comprehensiveness and specific of the food list, the quality of the information on food away from home, and seasonality. Each of these characteristics affects the quality of the measurement of food consumption that habitually occurs within the household. It is therefore important to develop guidelines or recommendations on best practices to be adopted when designing a survey to collect food data.

To draw such guidelines, it is necessary to develop a good understanding of the impact of survey design on food consumption measurement. The impact can be assessed with two different approaches. First, by piloting surveys that collect food data using different designs. Second, by analysing existing surveys that collect food data from different designs. The first option is more accurate, but also more time- and resource- consuming than the second.

This paper will present results on the impact of survey design on food consumption measurement from a panel of about seventy surveys. Food consumption is measured in terms of calories, as Dietary Energy Consumption (DEC). This is one of the key background variables employed by FAO in the measurement of undernourishment. The paper will also

discuss case studies that have looked at the impact of survey design, notably the one by Beegle et al. (2012) on Tanzania, and a more recent study conducted on the Mongolia 2007-08 Socio Economic Survey (Troubat and Grunberger, 2015).

Next section describes the general characteristics of the Household Budget Surveys, focusing on those employed in this study. Section 3 presents the sample of surveys employed in the analysis while results are presented in section 4, considering separately the impact of survey design on the level and variability of the DEC. Section 5 discusses the result of the case studies, while section 6 offers concluding remarks.

2. Types of National Household Budget Surveys

As mentioned, data on food "consumption" is collected in the vast majority of the National Household Budget Surveys (NHBSs). However, the data that are effectively collected may differ significantly across surveys. The term "food consumption", to begin with, lend itself to multiple meanings. Economists, for instance, use this term to designate the purchase of goods in a market, while they tend to disregard the destiny of what was purchased and have more interest in expenditure than in physical quantities. Nutritionists refer to consumption as the ingestion of a food, possibly net of unusable biological parts, and take no interest in purchases, if not as proxies of ingestion.

Food security analysis needs to look at both sides, as they concur to define what were presented as different dimensions of food security[1]. Food expenditure, together with the other means to acquire food such as transfers, gifts and own-production is a variable of interest from the point of view of the ability of households and individuals to access the necessary amount of food. What is ingested, at the same time, concurs to determine the physical well-being of each individual, through the characteristics of the diet. Expenditures and acquisitions in general are frequently – but not exclusively – centralized at the household level, at least for the food that is prepared and consumed in the house, while consumption away from home may be more frequently individual. Ingestion, instead, is exclusively an individual fact. Hence Household Surveys tend to be used primarily to collect data on purchases and other forms of acquisition, while ingestion is typically collected in individual-level surveys.

However, things are not so clear-cut in practice. If the survey is based on a very efficient sample, conveniently spread across time and space and representative of the relevant population strata, one may argue that stock changes – ie the difference between food purchased or else acquired and food physically consumed -- should become negligible, as they should cancel out in the average. In fact a lot of information on physical amount consumed can be derived from Household Budget Surveys, which tend to rely on wider samples and be implemented with a higher frequency compared to individual ingestion surveys. Hence food security analysis benefits from the ability of these surveys to look across the entire process that goes from the acquisition to the ingestion of food. The key for using the information on physical amount embedded in Household Surveys is in their design. In fact, the ambiguity of the term "consumption" concurs to determine different approached to the way this variable is collected in the questionnaire of Household Budget Surveys.

[1] See FAO, IFAD and WFP (2013).

In recent years, the FAO Statistic Division has accessed and processed more than 100 surveys NHBS in the framework of the global monitoring of undernourishment[2]. To estimate the proportion of undernourished persons in a population – or Prevalence of Undernourishment (PoU) -- FAO uses a parametric approach in which the distribution of food consumption is compared with a minimum threshold. The distribution is characterized starting from food consumption data collected in NHBSs, translated in calories.

There are four different approaches to collecting food "consumption" data in the surveys analysed, based on how the enumeration is conducted.

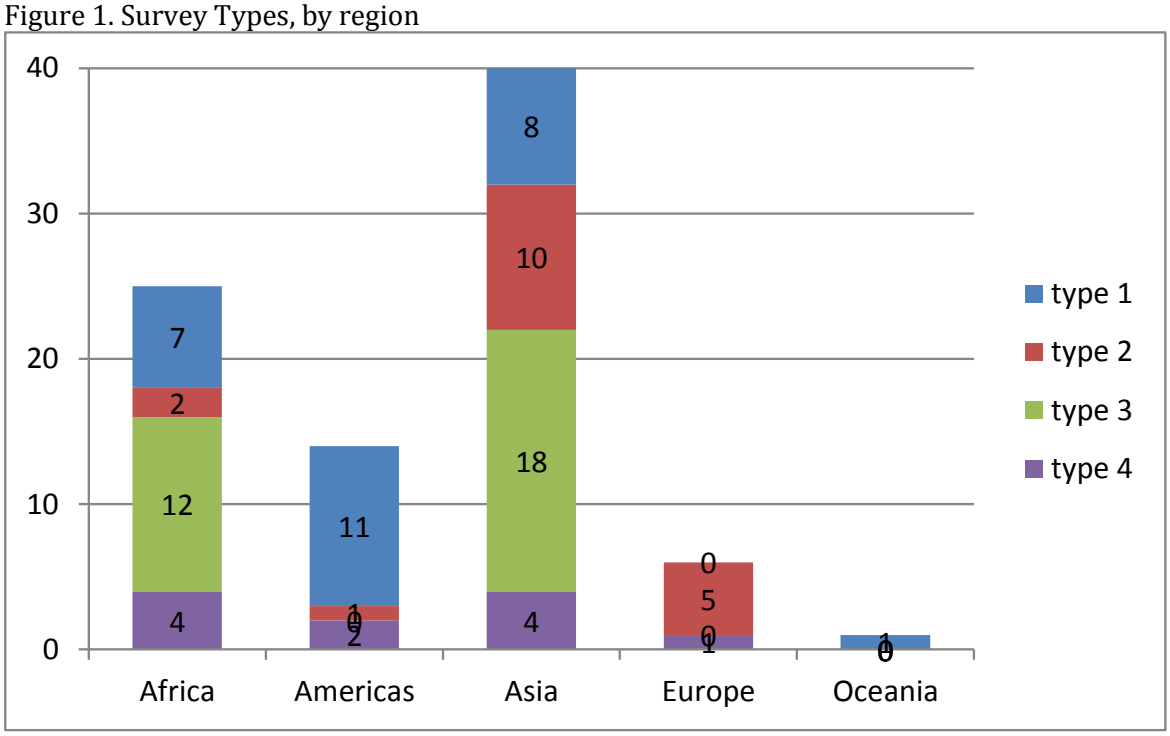
- Type 1: Acquisition for consumption. Households report on food acquired through purchases, own production and in-kind transfers over a past period. Food physically consumed is not reported.
- Type 2: Combination of acquisition and physical consumption. Households report on acquisitions only for purchases, and separately on physical consumption from own-production and transfers only. Information on acquisition and consumption is not comparable.
- Type 3: Physical consumption and sources of acquisition. Households report food they physically consumed and whether that food was purchased, own-produced or received as a transfer.
- Type 4: Both acquisitions and physical consumption for all sources. Quantities and values of food consumption are consistent[3].

Each of these approaches has its limitations. Examples of type 1 surveys are the 2008 *Encuesta Nacional de Ingresos y Gastos de los Hogares* of Mexico, or the 2009 Household Income Survey of Georgia. Type 1 surveys sometimes collect only monetary values, whose conversion into quantities is not straightforward. One exemple is the Questionnaire des Indicateurs Unifiés du Bien Être 2003 of Burkina Faso. Surveys of type 2 collect values of purchases and quantities of other forms of acquisition in different modules, whose information cannot always be reconciled due to several reasons, such as different description of items. Examples include the 2005-06 Ghana Living Standard Survey or the 2011 Household Budget Survey of Azerbaijan. Surveys of the third types often do not report information on food consumed away from home, which needs to be somehow imputed. Examples include the 2006-07 Socio Economic Survey of India or the 2010-11 Uganda National Household Survey. Finally, surveys of the fourth type are the most comprehensive, but they tend to be expensive and complex to implement. An example is the survey conducted in Brazil in 2008; another is the survey conducted in Cape Verde in 2001-02.

[2] The results of this monitoring exercise are published in the *State of Food Insecurity in the World* reports series. Examples are FAO, IFAD and WFP (2013, 2014).

[3] This type of survey can in fact provide consistent information on both consumption and acquisition using different survey organizations. We can further distinguish three sub-types: 1. acquisition from purchases and consumption from all sources of acquisition, so that a total value of consumption can be built using purchasing prices (eg Armenia); 2. acquisitions from all sources on a large sample and intake from all sources on a smaller sample (eg Cape Verde 2001, Brazil 2007); 3. acquisitions from all sources and consumption from all sources on the same sample (eg Mongolia 2008).

Figure 1. Survey Types, by region



Source: own calculation

Figure 1 shows the distribution of 86 surveys reviewed by type and geographical location. Type three – that is, surveys mostly based on consumption -- accounts for 35%, followed by type 1, which accounts for 31%.

Within the four survey types highlighted above, there are a number of other surveys characteristic that can affect the measurement of food consumption. One is the number of food items included in the survey. While in principle a longer and more detailed list of items should allow a higher accuracy, this may not be the case, as administering the questionnaire may become more demanding for the respondent and the enumerator. Hence results may show a decreasing accuracy beyond a certain length.

Another key feature of the surveys is the recall period, which is the length of the period for which respondents are requested to report. This period can vary from one day in the case of a single-day diary, all the way up to 365 days for one-year recall surveys. However, a diary can be filled repeatedly for a given period of time – usually from 7 days to 30 days. Moreover, the length of the overall period in which the survey takes place also affects the type of results that can be obtained, for two reasons. First, seasonality in consumption can only be captured if a whole 12-month period is adopted; and, secondly, if the sampling is well designed and distributed over a long period of time, there are higher chances for the difference between acquisitions and consumptions to be averaged out.

3. The sample of surveys analysed

For the work presented in this paper we could only use a sub-set of all the surveys reviewed above, as we had to select those surveys for which elementary information is available in

order to assess the impact of the different design on the food consumption measurement. We could use 77 surveys out of the total sample of 86 reviewed above.

For these 77 surveys, panel a of Figure 2 shows the distribution of surveys by type and sub-region. Type one above – which corresponds to surveys based on acquisitions -- accounts for 35%, and is the most frequent type of enumeration. This is followed by type 3, which accounts for 30% and is the group of surveys focused on consumption. As a result we have, 16 surveys belong to type 2, while 11 surveys belong to type 4. In Latin America the majority of surveys are of type 1, that is, they only collect data on acquisitions, while in Southern Asia most surveys belong to type 3, that is, they focus on consumption only. In Eastern Europe and Central Asia type 2 approach is more prevalent.

Figure 2 – panel a.
Distribution of surveys in our sample

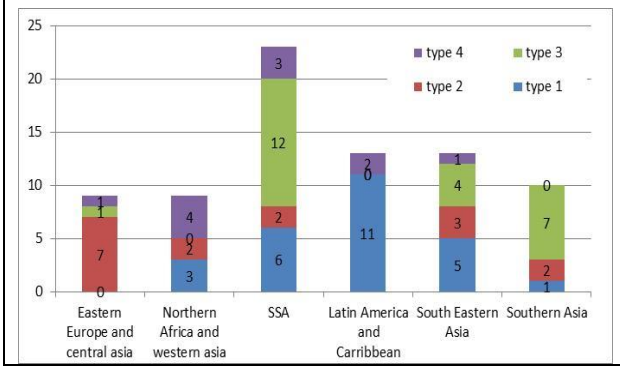
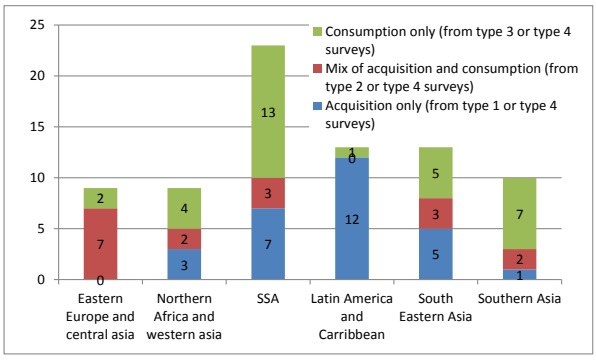


Figure 2, panel b.
Distribution of surveys by type of information analysed



Source: own calculation

In order to assess the impact of survey design on the measurement of food consumption we had to process similar information across the different survey types. This means that for Surveys of Type 3 and 4 we could consider the quantity consumed; for surveys of types 2 and 4 we could consider both information on acquisition and consumption; and for surveys of types 1 and 4 we could consider information on the quantities acquired. This led to the re-grouping surveys as per panel b of Figure 2: we have ”consumption only”, “acquisition only” and “mixed acquisition and consumption” type of information. The survey designs used to collect these types of information are compared in the following sections.

As mentioned, the acquisition vs consumption approach is not the only important element of a survey design that can affect the quality of the data collected. Other relevant characteristics include the use of recall interviews or food diaries; the ability of surveys to capture seasonal variations of the consumption and the effectiveness in capturing food away from home (FAFH) and own consumption.

In the 77 survey sample, diaries and recall interview enjoy similar popularity, as 45% of the surveys are collecting data via diaries while 55% via recall interviews. The majority, 56 surveys, collect data over 12 months, hence providing measurements which are robust to seasonal variation in food consumption and acquisition. Most of the surveys capture FAFH (75 surveys) and own consumption (74 surveys), which are important food sources besides purchases.

4. Cross-survey Analysis

Survey design and the level of the Dietary Energy Consumption

We want first to verify whether, in the selected sample of surveys, the level of the Dietary energy Consumption (DEC) varies with the type of survey design.

Overall, the mean DEC is 2240 kcal/person/day. In the various panels of Table 1 we show differences in mean DEC according the survey design. Between the different types of information analysed, consumption only have the lowest DEC and mix of acquisition and consumption have the highest DEC. The difference between consumption only and acquisition only is not significant (p-value=0.31). Considering mix forms of data collection and acquisition, the difference increases to 332 kcal/person/day and becomes statistically significant. Same result when we consider difference between mix form and consumption. Surveys with diaries have a 81 kcal/person/day higher DEC than recall interviews, and the difference is not significant (p-value=0.28).

The number of items in the food lists is widely variable, from a minimum of 25 items to a maximum of over 5000 in the Brazilian 2008 survey.

Table 1: Dietary Energy Consumption (DEC) from different survey designs

	N	Mean DEC (Adept)	Std. Dev.	Min	Max
All surveys	77	2240	305	1687	3141
A: surveys					
Acquisition only	32	2212	259	1687	2837
Mix of acquisition and consumption	17	2534	356	2007	3141
Consumption only	28	2102	222	1785	2617
Mean difference between acquisition and consumption surveys	- 65 (95% CI= [-192 62]; p-value=0.31)				
Mean difference between mixed and acquisition only	332 (95% CI= [129 534]: p-value=0.002)				
Mean difference between mixed and consumption only	397 (95% CI= [199 594]; p-value=0.000)				
B: Diary vs recall:					
Diary	35	2283	370	1687	3141
Recall interviews	42	2202	249	1799	2841
Mean difference	81 (95% CI= [-66 228]; p-value=0.28)				
C: Length of the reference period of the diary:					
Diary 2-7 days	13	2256	330	1867	3141
Diary 8-15 days	11	2516	427	1687	2952
Diary 30 days	11	2080	217	1785	2571
D: Recall period of interviews					
Recall 2-7 days	18	2232	271	1799	2837
Recall 8-15 days	11	2178	163	1949	2412
Recall 30 days	5	2143	316	1900	2753
Recall six months to 1 year	8	2214	286	2007	2841
E: Record of Food Away From Home:					
FAFH	74	2248	312	1687	3141
No FAFH	3	2000	179	1821	2180

Mean difference	248 (95% CI= [-114 611]; p-value=0.176)				
F: Record for own consumption:					
Own consumption	72	2260	309	1687	3141
No Own consumption	5	1928	131	1805	2118
Mean difference	332 (95% CI= [54 610]; p-value=0.02)				
G: Region					
Eastern EU and Central Asia	9	2719	217	2481	3141
North Africa and Western Asia	9	2386	379	1785	2952
SSA	23	2106	191	1687	2461
Latin America	13	2261	228	1867	2837
South-Eastern Asia	13	2046	191	1805	2571
Southern Asia	10	2200	294	1936	2841

Source: own calculations

Differences with respect to the reference period are reported separately for diaries and recall surveys. This variable needs to be handled separately for recall and diary surveys, because the reference period is supposed to affect recall interviews differently than diaries. In case of the recall interviews, the reference period is often related to recall error. For diaries, an increasing diary length might increase respondents' fatigue in filling the diary. For diaries a reference period of 8-15 days have the highest mean DEC. For recall interviews the highest DEC is observed for the shortest or longest reference period but differences between DEC are not significant for all recall periods. The lowest mean DEC are provided by surveys with the maximum reference period for diaries which is 30 days. This finding suggests a bell-shaped relationship between DEC and reference period for diaries.

Surveys which collect information on FAFH show a higher DEC than surveys which do not. However, we observe only 3 surveys without FAFH, which are too few to make a meaningful comparison. Also the fact that surveys capture own consumption increases the DEC, where the difference of 332 kcal is weakly statistically significant (p-value = 0.08).

Panel G of Table 1 reports the differences in DEC between geographical regions. These are remarkable. As discussed above, surveys in the same geographical region are often based on similar designs. Mean comparisons may therefore confound differences between world regions with survey characteristics. One way to control for this confounding factor is to estimate the effect of survey design with multiple regressions.

The comparison of simple averages, run so far, might be misleading. More reliable information can be obtained by considering the impact of survey design in a multiple regression, and controlling for the characteristics of the country in which the survey takes place.

To this end we estimate the following equation:

$$\ln(DEC_i) = c + \beta S_i + \gamma C_i + \epsilon_i \qquad (1)$$

where the index *i* refers to a specific survey. As dependent variable we use the logarithm of DEC. S_i contains a vector of variables related to survey design, while C_i is a vector of country characteristics. In particular, S_i includes variables explaining the features of the information analysed (consumption only is the reference category) a dummy variable indicating whether the survey is based on a recall interview (diaries are the reference category), the logarithm of number of in house food items included in the survey and a

dummy indicating whether own consumption is captured in the survey.[4] Dummy variables are also included for the major regions, along with GDP per capita.

Results (Table 2) suggest that the level of DEC depends on whether acquisition, consumption or a mix of acquisition and consumption is analysed. The regression also shows no difference between surveys based on diaries and those based on recall interviews. Rather, the number of items seems to exert a negative influence on the level of the DEC, albeit with a very small impact. The inclusion of own consumption too, seems to have no impact.

The level of the DEC seems not to vary significantly with GDP per capita, which is surprising. However, the regional dummy variables seem to capture most of the income effect. Differences between regions are remarkable. If these variables are omitted, the effect of GDP becomes significantly positive[5].

Equation 1 was re-estimated for the subsets of surveys based on diary and for recall surveys. Moreover, a term referred to the length of the reference period was added to the specification in column (1) of Table 2. Comparing columns (2) and (3), it is evident that when diaries are employed, the DEC from acquisition only of data is about 30% higher than the DEC from consumption only type of data. In the case of recall surveys, this difference becomes negligible.

Table 2: Regression result for the level of the Dietary Energy Consumption (DEC)

	(1) All	(2) Diary	(3) Recall	(4) All FAFH
Acquisition only	0.07'	0.24**	-0.01	0.07'
	(0.04)	(0.06)	(0.06)	(0.04)
Mixed Acquisition and consum.	0.14**	0.38**	0.11*	0.16**
	(0.03)	(0.07)	(0.04)	(0.04)
Recall	0.06			0.06
	(0.04)			(0.04)
ln(# items)	-0.03	-0.02	-0.05'	-0.02
	(0.02)	(0.03)	(0.03)	(0.02)
No own consumption	-0.10		-0.05	-0.08
	(0.07)		(0.07)	(0.11)
ln(GDP)	-0.01	0.04	-0.01	0.00
	(0.01)	(0.02)	(0.01)	(0.01)
Ln(reference period)	-0.04*	0.03	-0.04*	-0.04*
	(0.02)	(0.04)	(0.02)	(0.02)
FAFH ratio				0.00
				(0.00)
North Africa and Western Asia	-0.05	0.06	-0.18'	-0.05
	(0.05)	(0.06)	(0.10)	(0.05)

[4] We do not include a dummy for FAFH, because of the low number of surveys where FAFH is not measured.
[5] Results are not reported, but they are available upon request. When regional dummies are removed, one additional effect is that the parameter of acquisition-only surveys is no longer significant. The reason is that in one region, Latin America and the Caribbean, most surveys are of the acquisition-only type.

SSA	-0.22**	-0.11	-0.25**	-0.20**
	(0.05)	(0.08)	(0.06)	(0.05)
Latin America	-0.14*	-0.07	-0.09	-0.09
	(0.06)	(0.08)	(0.08)	(0.07)
South-Eastern Asia	-0.20**	-0.07	-0.27**	-0.18**
	(0.06)	(0.08)	(0.07)	(0.06)
Southern Asia	-0.19**		-0.22**	-0.18**
	(0.05)		(0.06)	(0.06)
Constant	8.09**	7.19**	8.28**	8.04**
	(0.13)	(0.29)	(0.14)	(0.14)
N	77	35	42	74
R^2	0.608	0.793	0.628	0.608

Source: own calculations
Note: Standard errors (in parenthesis) are clustered at country level, with ' p<0.10, * p<0.05 and ** p<0.01. In case that the coefficient has not been estimated due to multicollinearity, field is filled with a single dot. Surveys who do not capture FAFH are omitted from regressions in column (4).

This may indicate that stocks are captured effectively by diaries, while they are neglected by recall surveys, even if the survey is designed to distinguish them. The data collection method, in other words, would have a stronger impact on the result compared to the enumeration method.

Both diary-based and recall-based surveys show a negative correlation between the DEC and the length of the food item list, but the impact seems to be stronger for recall-based surveys. These also show decreasing DEC with the length of the reference period. This indicates that recall bias might be more prevalent when interviews ask to recall food consumption of a longer period. The length of the reference period seems to have a positive impact on DEC from diaries but the coefficient is not significant and fails in capturing the bell shape relationship between DEC and the reference period.

One problematic area in the measurement of food consumption in household surveys is FAFH. As it is difficult to recall and record the physical characteristics of foods consumed away from home, the expectation is that FAFH is under reported in Household Surveys. To verify this assumption, we add the ratio of calories from FAFH as a covariate to equation 1. Results (Table 2, column (4)) suggest no impact of the ratio of FAFH on the DEC.

Survey design and the variability of the DEC

A priori, we expect that food data collected in type 1 surveys to show higher variability compared to those collected in type 3 and 4 surveys. Indeed, Coefficient of Variation (CV) are highest for acquisition surveys. The difference between consumption and acquisition is weakly significant (Table 3, panel A)[6].

[6] For the analysis of the impact on the variability of the DEC we had to exclude two surveys from the sample as the distribution of this variable across households was not available.

Table 3: Coefficient of Variation of the Dietary Energy Consumption, by survey design

	CV	
	N	**Mean CV (in %)**
All surveys	75	58
A: surveys		
Acquisition only (from type 1 and type 4)	29	76
Mix of acquisition and consumption (from type 2 and type 4)	16	38
Consumption only (from type 3 and type 4)	31	52
Mean difference acquisition - consumption	24 (95%-CI=[0.7 49]; p=0.043)	
Mean difference mix of Acq. & Cons. – Acquisition only	-39 (95%-CI=[-60 -17]; p=0.00)	
Mean difference mix of Acq. & Cons. – Consumption only	-14 (95%-CI=[-34 5]: p=0.14	
B: Seasonality control:		
Seasonality control	59	60
No seasonality control	16	50
Mean difference	11 (95%-CI=[-3 24]; p=0.123)	
C: Seasonality control – Acquisition surveys only:		
Seasonality control	22	82
No seasonality control	6	54
Mean difference	28 (95%-CI=[2 54]; p=0.04)	
D: Seasonality control – Consumption surveys only:		
Seasonality control	24	53
No seasonality control	7	48
Mean difference	5 (95%-CI=[-28 38]; p=0.76)	

Source: own calculation

Panel B of table 3 shows the differences in CV between surveys that account for seasonality and those that do not. The former show 17% higher CV, indicating that DEC is seasonal. Again, Panels C and D show the impact of seasonality for acquisition and consumption surveys respectively. The smaller difference for consumption surveys (5%; see panel D) might be explained by the fact that acquisition varies more strongly over the year due to the harvest season than consumption.

As for the level of the DEC, to obtain more reliable information we performed a multiple regression based on the following equation:

$$\ln(CV_i) = c + \beta S_i + \gamma C_i + \epsilon_i \qquad (2)$$

where the index *i* are survey. As dependent variable we use the logarithm of the CVs. As in equation 1, S_i is a vector of variables related to survey design, and C_i is vector of country characteristics. Covariates in S_i differ from the equation (1), as the CV is likely to be affected by different variables. In particular, S_i contains variables accounting for the survey design – consumption-only is the reference group -- and dummy variables identifying surveys that capture seasonality, those based on recall interviews -- diary-based is the reference group -- and those that do not collect information on own consumption. Variables in C_i include country GDP per capita and dummy variables for the regions.

Results (Table 4) suggests that CVs of acquisition surveys are, on average, 42% higher than CVs of consumption surveys.[7] Once controlling for other features of survey design and country characteristics by estimating equation 2, the coefficient of acquisition in column (2) does not change much in its magnitude (0.35), but becomes statistically insignificant. Column (2) shows also that all other survey characteristics seem not to influence significantly the raw CVs.

In column (3) and (4) we estimate equation 2 for acquisition and consumption surveys differently. From results in column (3) we can see that controlling for seasonality does not influence CV. If a survey is carried out over an entire year, the raw CV increases. Regression results for the subsample of consumption surveys show no effect of survey design on the CV. This suggests that consumption surveys are more robust for measuring nutritional inequality.

Table 4: Regression result for the Coefficient of Variation of the DEC

	ln(Raw CV)			
	(1)	(2)	(3)	(4)
	All	All	Acqu.	Cons.
Acquisition only	0.35**	0.35*		
	(0.12)	(0.16)		
Mixed acquisition and consumtpion	-0.26'	-0.25		
	(0.14)	(0.15)		
Seasonality		0.16	0.24	0.13
		(0.13)	(0.33)	0.18
Recall		0.05	-0.18	0.14
		(0.15)	(0.29)	0.30
Ln(GDP)		-0.04	0.02	-0.03
		(0.05)	(0.20)	0.05
North Africa and Western Asia		-0.13	0.26	0.16
		(0.21)	(0.39)	0.37
SSA		0.16	0.06	0.55'
		(0.22)	(0.46)	0.27
Latin America		-0.04		0.48
		(0.25)		0.43
South-Eastern Asia		-0.28	-0.27	0.11
		(0.22)	(0.38)	0.30
Southern Asia		-0.48'	-0.48	-0.18
		(0.25)	(0.61)	0.29
Constant	3.83**	4.05**	3.90*	3.63**
	(0.08)	(0.50)	(1.85)	0.62
N	75	75	28	31
R^2	0.214	0.404	0.202	0.548

Note: Standard errors (in parenthesis) are clustered at country level, with ' p<0.10, * p<0.05 and ** p<0.01. In case that the coefficient has not been estimated due to multicollinearity, field is filled with a single dot.

[7] As we use the logarithm of CVs, the marginal effect of the dummy variable coefficient is its exponential value, i.e. exp(0.35)=1.42.

5. Case studies on the effect of survey design on food consumption measurement

The Survey of Household Welfare and Labour in Tanzania

The Survey of Household Welfare and Labour (SHWALITA) was conducted *ad hoc* for an experimental study carried out in Tanzania from September 2007 to August 2008 (Beegle et al., 2012). The survey respondents were split into eight representative subsamples, and food consumption data were collected from each subsample with a different type of enumeration method.[8] The designs differed according to: data collection method, ie diary versus recall; the level of the respondents, ie individual versus household; the reference period for which consumption is reported and the degree of detail of the items.

An individual diary, were enumerators supervised the respondents on a daily basis over a period of 14 days was found to be best way of capturing dietary energy consumption. Individual diaries are considered to be superior to household diaries, as a single reference person is not always aware of food consumption of other household members (especially of food away from home). Assessing individual food consumption is, however, more expensive than a household-level assessment. In the latter, avoiding multiple counting of food items of collective consumption can be difficult.

An interesting result of SHWALITA is that the seven day recall interviews on household food consumption turned out not to perform much worse than individual diaries.[9] Extending the recall period, in other words, resulted in a lower DEC, and the DEC of the 14 days recall was similar to household diaries. The worst method for assessing food consumption turned out to be the 12-month recall interviews; hence results advise against the use of interviews with long reference period in general.

As part of the experiment, in two types of recall interviews the degree of commodity detail was reduced from 58 to 11 items: this resulted in a considerable drop of the DEC. The reduction of the list of food items brought about only a relatively small reduction of the interview time. Therefore, Beegle *et al.* (2012) conclude that the potential saving on interview costs obtained with a shortened list of food items may not compensate for the loss of accuracy.

The Mongolia Socio-Economic Survey

A recent study has assessed the impact of survey design on the measurement of food consumption and acquisition using the 2007-08 Mongolia Socio-Economic Survey (MSES) (Troubat and Grunberger, forthcoming). The MSES was carried out from July 2007 to June 2008 and follows a diversified approach in measuring food consumption. The survey is organized with both diaries and supplemental recall interviews (both with a reference period of about 30 days). An additional feature of the MSES is that it collects data not only on food consumption, but also on acquisition and stocks. Troubat and Grunberger (forthcoming) use this variety of data collection methods to analyse the impact of the type of enumeration on food measurement. The main findings of this study as summarized below.

[8] SWALITA included also non-food household consumption, even though food was one of its key aspects.
[9] Nevertheless, as Beegle et al. (2012) mentioned, researchers need to be aware that recall interviews are likely subject to recall and telescoping error and are less adequate to capture personal food away from home than personal diaries.

*Impact of the diary length***.** Food consumption reported by household decreases as the period during which the dairy has to be filled increases. This is due to the respondent's fatigue in filling the diaries.[10] Shortening the diary length, not only reduces the bias of measurement, but results in a significant reduction of costs.

Cycles in food consumption and acquisition. Seasonality over the year is a commonly known issue in household surveys (Behrman and Deolalikar, 1989). In MSES, both food consumption and acquisition show remarkable variation not only between months, but also within the months.[11] Acquisition of food has been found to be higher at the beginning of each calendar month. For food consumption, also differences across weekdays have been attested. This suggests that enumeration needs to be spread across months and also across week days, in order to account for the cyclical behavior in food acquisition and consumption, and obtain measurements that are independent from it.

Food consumption diaries and recall interviews. Food consumption collected through diaries and recall interviews are not significantly different in the MSES. This result is surprising, as diaries are often considered to be more accurate than recall interviews. However, in the MSES recall interview may be influenced by the diaries, as they are performed just after three diaries have been filled. Hence, it is possible that the assessment of the recall method was biased.

Comparing food consumption and acquisition. In the MSES food acquisition is measured only through recall interviews, and food consumption is derived as a difference between beginning and ending stocks. This means that in principle, the MSES should report no difference between consumption and acquisition, as beginning and ending stocks may cancel out in the average. However, results show that food consumption is about 10% higher than acquisition, and the main reason for the discrepancy is the presence of monthly cycles in acquisitions.

6. Concluding remarks

Food consumption measurement is very sensitive to survey design. Collecting information on food acquisition rather than food consumption increases the estimated level of average Dietary Energy Consumed by more than 100 kcal/person/day. This range of difference may imply totally different results in certain countries, and can seriously bias the planning of food security and nutrition intervention programs. Not only the mean, but also the distribution of foods data among households is heavily affected by the survey design. There is a difference of more than 20 percent between the coefficients of variation of calories per person per day derived from surveys that focus on acquisition and those derived from surveys that focus on consumption. This difference may have strong impact on the estimates of Prevalence of Undernourishment, such as those that FAO undertakes in the framework of the global monitoring of food security. To overcome this problem, surveys employed by FAO in the global monitoring are treated for this extra variability (Wanner et al, 2014).

The analysis of the Mongolia 2007/08 Socio Economic Survey confirmed results of the panel, and showed how the length of period in which the food diary is administered to households

[10] This phenomenon of decreasing diary quality with respect to diary length has also been found in the previous literature (see review in Crossley and Winter (2013)).
[11] Systematical variations of expenditure within the month have been also found by Damon et al. (2013) and Hastings and Washington (2010).

impacts food consumption measurement. The analysis shows how to derive an optimal number of days for administering the diary. More analyses of surveys that collect information through different designs will facilitate the preparation of guidelines on best practices -- and bad practices to be avoided – when collecting food data in household surveys.

References

ADePT Food Security Module, available at http://www.fao.org/economic/ess/ess-fs/fs-methods/adept-fsn/it/ manual and software//

Cafiero C. (2014). "Advances in hunger measurement: traditional FAO methods and recent innovations". FAO Statistic Division. ESS working paper 14-04. FAO, Rome.

Cafiero C. and Gennari P. (2011). "The FAO indicator of the prevalence of undernourishment". Presented at the Workshop "Measuring Food Insecurity and Assessing the Sustainability of Global Food Systems", The National Academies of Science, Washington, DC, February 16-17, 2011.

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

FAO (2014)- "Food Balance Sheets and The Food Consumption Survey: A Comparison of Methodologies and Results"

FAO, IFAD and WFP (2013). *The State of Food Insecurity in the World. The Multiple Dimensions of Food Security*. Rome 2013.

FAO, IFAD and WFP (2014). *The State of Food Insecurity in the World. Strengthening the Enabling Environment for Food Security and Nutrition*. Rome 2014.

Fiedler J., Carletto C., and Dupriez O. (2010). "Still waiting for Godot? Improving Household Consumption and Expenditures Surveys (HCES) to enable more evidence-based nutrition policies".

Kaara J, Ramasawmy S. (2008), "Food data collected using acquisition and consumption approaches with a seven-day recall method in Kenya's KIHBS 2005/2006". In: Sibrian R, ed. Deriving food security information from national household budget surveys: experiences, achievements, challenges. Rome: Food and Agriculture Organization, 2008: 69–80.

Naiken LL. (2003)."FAO methodology for estimating the prevalence of undernourishment". In Measurement and Assessment of Food Deprivation and Undernutrition. Proceedings of the International Scientific Symposium convened by the Agriculture and Economic Development Analysis Division in Rome, 26-28 June 2002.

Naiken L. (2014) "Methodological issues in the estimation of the prevalence of undernourishment based on dietary energy consumption data : a review and clarification". FAO statistic Division. ESS working paper 14-03. FAO. Rome 2014.

Martirosova D. (2008), "Food data collected using acquisition and consumption approaches with daily diaries in Armenia's ILCS 2004". In: Sibrian R, ed. Deriving food security information from national household budget surveys: experiences, achievements, challenges. Rome: Food and Agriculture Organization, 2008:59–68.

Pereira C, Troubat N, Sibrian R. (2008), "Food data collected using acquisition and intake approaches in Cape Verde's IDRF 2001/2002". In: Sibrian R, ed. Deriving food security information from national household budget surveys: experiences, achievements, challenges. Rome: Food and Agriculture Organization, 2008:81–90.

Sibrian et al. (2008). “Deriving food security information from national household budget surveys. Experience, achievements, challenges”. FAO. Rome.

Smith L. C., and Subandoro A. (2005). “Measuring food security using household expenditure surveys. Food Security in Practice series. International Food Policy Research Institute, Washington, D.C.

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

Smith L.C., Harold Alderman and Dede Aduayom (2006) “Food Insecurity in Sub-Saharan Africa. New Estimates from Household Expenditure Surveys.” Research report 146. International Food Policy Research Institute, Washington, D.C.

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