Public Disclosure Authorized

Public Disclosure Authorized

WPS7090

Policy Research Working Paper 7090

The Impact of Household Food Consumption Data Collection Methods on Poverty and Inequality Measures in Niger

Prospere Backiny-Yetna
Diane Steele
Ismael Yacoubou Djima



WORLD BANK GROUP
Development Research Group
Poverty and Inequality Team
November 2014

Policy Research Working Paper 7090

Abstract

This paper assesses the impact of three methodologies of food data collection on the welfare distribution, and poverty and inequality measures in Niger. The first methodology is a 7-day recall period, the second one is a usual month, and the third one is a 7-day diary. The paper finds that there is a difference in the distribution of welfare between, on the one hand, the two first methodologies (7-day recall and a usual month, which give results close to each other) and, on the other hand, the 7-day diary method. When considering annual per capita consumption, the 7-day diary lags the 7-day recall by 28 percent. This gap is not only at the mean of the distribution, it has been found at any level. These differences lead to differences in poverty and inequality measures even when alternate poverty lines are used. This study underscores the problem that many developing countries face when it comes to monitoring poverty indicators over time where different methodologies have been used over the years.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at pbackinyyetna@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

The Impact of Household Food Consumption Data Collection Methods on Poverty and Inequality Measures in Niger

Prospere Backiny-Yetna
Diane Steele
Ismael Yacoubou Djima

Keywords: poverty measurement, consumption, expenditure, survey design
JEL: I32, C83, D12, C18

1. INTRODUCTION

As 2015 nears, many developing countries are assessing whether they have achieved the Millennium Development Goals (MDGs) first goal of halving poverty from its 1990 level. This process inevitably brings up the issue of monitoring poverty indicators, questions about the indicators themselves, and the methods used to measure these poverty indicators.

Monetary poverty indicators can be computed using either income or consumption. Each of the two variables has its own advantages and disadvantages as a measure of living standards. Income translates the real resources of the household at a particular point in time. In addition, it is possible to assign particular sources of income (agriculture, non-farm enterprise wages, pensions) to particular members of the household; giving the possibility to the analyst to work on intra-household inequality to some extent. Income is, however, very difficult to measure in developing economies where most of the active population derives its income from agriculture and other non-agricultural self-enterprises without using any accounting practices. Moreover, income can fluctuate from year to year because of different shocks, particularly in rural agricultural societies. Consumption is not closely tied to those short-term fluctuations and is smoother and less-variable than income. In those situations the ranking of households by income is less robust than the ranking by consumption (Deaton and Zaidi, 2002). As such, consumption (food and non-food) has been the standard variable to measure monetary poverty in developing contexts.

Food consumption is computed from information on how much food the household acquired from different sources (market purchases, home production on farms/gardens, and gifts or in-kind payments) in monetary terms and/or quantities. Methodologies of data collection can differ in many ways. They can differ between diary and recall; on the duration used to report household expenditures; on the list of items about which the data are collected (in the case of recall). Each of these elements affect the distribution of the expenditure and the methodology chosen has an effect on the distribution of expenditure, and therefore on the poverty levels (Lanjouw and Lanjouw, 2001; Tarozzi, 2007; Beegle *et al.*, 2010).

Niger is a case study of difficulties in poverty comparison due to differences in survey methodologies over time. Between 2005 and 2008, the *Institut National de la Statistique (INS)* of Niger implemented two national household surveys that have been used to measure poverty for, among other purposes: (i) monitoring poverty and reporting on the progress towards the MDGs;

(ii) assessing the impact of public policies to the poor; and (iii) targeting the poor (INS, 2006). The "*Questionnaire des Indicateurs de Base du Bien-être*" (QUIBB) was fielded in 2005. This survey collects food consumption information by the usual month approach for a comprehensive list of food items. The strategy of data collection consists of getting the usual monthly expenditure of each item and the number of months the item is consumed over the last 12 months. The second survey is the "*Enquête Nationale sur le Budget et la Consommation des Ménages*" (ENBC), which was fielded in 2007/08. This survey uses a 7-day diary to collect food consumption information from all sources.

In 2011, INS decided to institute a third survey which will be used as a baseline for future poverty monitoring. The third survey is the "*Enquête Nationale sur les Conditions de Vie des Ménages et l'Agriculture*" (ECVMA), an effort coordinated with the country's National Strategy for the Development of Statistics (*Stratégie Nationale de développement de la statistique or SNDS*) to support a new round of poverty estimates. The ECVMA methodology of data collection was conceived to take into account two constraints: the seasonality in household consumption and, the fact that agricultural data should be collected both during the planting and harvest period. Seasonality is likely to affect consumption. Niger is a rural country, and we expect that consumption will be high in the period after harvest and low as we move away from the harvest period.

The methodology used for collecting food consumption data for the ECVMA is a 7-day recall. The original idea was to use the same 7-day diary methodology as in 2007/08 in order to avoid issues of comparability of poverty indicators between the two surveys but this could not be done for at least two reasons. First, the enumerator would have to collect consumption data daily (this is true particularly in Niger where the literacy rate is low) and he would end up administering fewer questionnaires during the same time period than he would have for the 7-day recall. Second, the survey had to be conducted in at least two visits in order to collect agricultural data for the planting, and the harvest period. As a result, the idea of the 7-day diary was dropped, and the 7-day recall was chosen with the thinking that its methodology being close to the 7-day diary, it would produce more comparable results.

But we know from the extensive research cited in this paper that the specific design of the questionnaire can influence the data collected. When measuring poverty and comparing poverty measurements over time, it is difficult to know how to interpret differences between poverty

levels. Do these poverty measurements differ because of differences in the wellbeing of the population or because of differences in the methodologies? Comparing the three methodologies used in Niger through an experiment brings at least part of an answer to this question.

In preparation for the 2011 ECVMA survey and the changes in consumption measurement, an experiment was organized and conducted during the pilot survey to provide data using the three types of the methodologies implemented in Niger. The objective of the experiment was to assess how much of the differences in poverty indicators in Niger might be attributed to differences in data collection methodologies. The first methodology is the 7-day recall, the second one is the usual month and the last one uses a 7-day diary. We find that there is a large difference in consumption and poverty between the first two methodologies (which give close results to each other) and the 7-day diary method. The annual per capita consumption from 7-day recall method was, on average, 28 percent higher than that of the 7-day diary. Obviously, these differences lead to differences in poverty figures; and any analysis of poverty trends using different surveys with those methodologies, without some corrections, probably leads to errors.

The rest of the paper is organized as follows. In Section 2 a literature review is presented. Section 3 provides a description of the data, including the way the experiment has been implemented. Section 4 explores the impact of the methodology used in data collection on the distribution of economic welfare during the period reviewed and discusses the consequences on poverty and inequality. A conclusion follows.

2. SURVEY DESIGN AND CONSUMPTION DATA: A LITERATURE REVIEW

Measurement issues are at the heart of data collection. Regardless of the information being collected - employment, income, expenditures, mortality, etc. - the way that the data are collected matters for the use of those data. Measurement issues encompass all aspects of data collection. Among those issues, one can cite the level at which to collect the data (individual or household), the period of the year data are collected (employment and some other variables are affected by seasonality), the person who is providing the information (the individual him or herself or a proxy respondent).

To illustrate the influence of the methodology and the time of year in which a survey is administered on the statistics one can compute, we consider the proportion of the population not involved in the labor market and the percentage of people in each branch of activity. These are

simple statistics to produce from any household survey. We remain with the example of Niger. When using the 2005 and the 2011 households surveys in Niger, we find that in 2005, 46 percent of the active population (people older than 15) are not involved in the labor market and nearly 56 percent of the ones involved work in agriculture. The same statistics using the 2011 household survey are 20 percent not working and 77 percent of those in the labor market were involved in agriculture.

To be able to interpret the changes in those statistics one has to know that in 2005, the fieldwork was done from mid-April to mid-July, at the end of the off-season for agriculture when many people in the rural areas are still inactive; moreover, the reference period for employment was 7 days, which excluded all jobs done during the remainder of the year. In the 2011 household survey, the fieldwork was done from mid-July to mid-September, at the beginning of the agricultural season and the reference period was the year.

The debate over the best method to collect expenditure information has been considered a problem since the start of the Living Standards Measurement Study Program (Saunders and Grootaert, 1980).[1] Expenditure data (and particularly food expenditure) can be collected using a diary or recall. When using the diary approach, one has to determine how long the data have to be collected (a week, a month or longer). In the case of the recall, a list of items has to be determined, and the recall period has to be defined. For both methods, the time of year when the data collection occurs can be an issue, except if data are collected all year long. Because of those different issues, each method presents specific challenges. A body of research has been conducted that shows clearly that the method used to collect data has an impact on the results of analyses done using the data.

We first consider the diary method, which, if well implemented, one would expect to yield results closer to actual levels of households food consumption. First, there needs to be a respondent in the household who is literate and can record the entries in the diary. If no one in the household is literate, the interviewers must assist in completing the diary which blurs the line between a diary and a recall survey (Beegle *et al*., 2010). Second, diaries must be left with the household and picked up after the recording period is completed. This can pose logistical problems for the interview teams in making sure that someone collects the diaries and sends them for processing. Third, the use of a diary alters interviewing. It reduces the amount of time

[1] For more information on the LSMS program, go to the web site: www.worldbank.org/lsms.

that the interviewer spends interviewing the household, but may increase the amount of time that the interviewer spends traveling as s/he has to make an extra trip to the household to pick up the diaries. In addition, the interviewer may need to spend time helping illiterate households in filling in the diaries (Deaton and Grosh, 2000). For this last reason it turns out that in practice, in illiterate households, diaries become more like short recall surveys.

In testing the accuracy of data collected from diaries, several studies analyze changes in recording over time. McWhinney and Champion (1974) observe higher first-week expenditures in Canada; the first-week expenditures averaged 8.3 percent above second-week expenditures, and that has come to be accepted as a fact of life in record-keeping surveys. Wiseman, Conteh, and Matovu (2005) show that two-weeks diaries provide satisfactory estimates for food consumed at home, but are deficient in food consumed outside of the home. In addition, missing or unclear data may be difficult to resolve. If researchers have to go back and clarify entries with respondents, the data soon become retrospective and subject to recall biases.

The recall period also impacts the distribution of consumption. The choice of the ideal recall period is among the most important and difficult design issues for the consumption module. Longer recall periods are better than shorter ones for measuring the distribution of consumption because averaging consumption over many days eliminates the randomness of some of the household's day-to-day purchases that have nothing to do with its standard of living (Deaton and Grosh, 2000). However, people find it harder to remember more distant events – longer reporting periods lead to more forgetting, at least for common purchases like food (Deaton, 2001). It is a well-accepted assumption that the longer the recall period, the greater the likeliness of recall errors, but the longer the recall period, the more possible it is to cover a larger sample of transactions for a given number of interviews, and therefore for a given field cost (Scott and Amenuvegbe, 1990). Shorter recall periods may help the respondents to report more accurate information, but there is also the problem of "telescoping" in which respondents report events that lie outside the reference period. Telescoping appears to behave such that the more frequent the event, the greater the likelihood of confusion about dates (Bradburn, 2002). Therefore if a household has some food items that are purchased frequently, a purchase may be recorded in a recall interview even if that purchase is not made during the specific recall period.

The different survey methods affect estimates of the composition of expenditures as well as the estimates of average expenditures. The average food budget share may change depending

on the data collection method, and the food budget share plays an important role in the calculation of poverty lines. Therefore, the poverty line, and the measured poverty level, may depend on whether the expenditure data are gathered by the diary method or the recall method (Gibson, 1999). Tarozzi (2007) cites studies of poverty comparisons in India drawn from a number of cross-section surveys with changes in recall periods that show that when comparing over time, compensation methods need to be used in order to make up for the changes in methodology. His study shows that in urban areas, the poverty decline is consistent with the official estimates from the Office of Statistics, but that in rural areas, one-third of the measured decline in poverty is a statistical artifact. In El Salvador, Jolliffe (2001) compares poverty numbers using different questionnaire designs and finds that estimated levels of consumption are highly sensitive to questionnaire design. This leads to differences in the calculation of poverty based on the different questionnaire designs. The measures of absolute poverty are significantly affected by the differences in measured consumption levels. Measures of relative poverty do not show significant differences based on the questionnaire design, but the geographic distribution of relatively poor people is significantly different across the different questionnaire designs.

The three national household surveys conducted in Niger between 2005 and 2011 face some of the problems outlined above, making it difficult to track poverty over time. The experiment was conducted to inform the debate on poverty trends in the country.

3. THE DATA

3.1. Description of the experiment

The data used in this paper come from an experiment conducted as part of the ECVMA pilot survey. The purpose of the experiment is to understand the implications of the alternative survey designs on measuring food consumption and poverty in Niger.[2]

This survey consists of three questionnaires. The household questionnaire collects information on household demographics, education, health, anthropometrics, employment, non-agricultural enterprises, dwelling, durable goods, access to social services, non-salary income, transfers and remittances, and food security. In addition, food consumption and non-food expenditure information is collected in the household questionnaire. There is an agriculture and

[2] The pilot survey also included an experiment on the collection of milk production data. The details of that experiment are not reported here.

livestock questionnaire that collects information regarding the household's crop activities and livestock. Finally, a community questionnaire is used to collect information on the areas around the households selected for the survey on topics such as access to infrastructure and availability of services. The community questionnaire also collects price information that can be used to make regional adjustments to the welfare measure.

As noted above, the ECVMA pilot survey was conducted (in addition to the usual testing of questionnaires, and data entry programs to be used in field work) to evaluate the quality of food consumption data collection methods. For that purpose, the pilot needed to be conducted on a representative sample with a certain size. A nationally representative sample would be too disperse and would have increased the costs of data collection. Knowing that the further away from the capital Niamey the higher the transportation costs, the pilot survey is conducted in urban and rural areas near Niamey.

The final sample size, 648 households, is chosen both for financial and logistical reasons. The geographic scope of the operation is limited to Niamey, the capital city, and Tillabery (urban and rural areas), the nearest region. The sample is drawn randomly from the 2001 population census, excluding enumeration areas (EAs) already selected for the main survey. The sample is drawn as a two-stage random sample. In the first stage we randomly draw the enumeration areas (24 in Niamey, 12 in urban Tillabery and 12 in rural Tillabery). In the second stage, we select randomly households from updated household listings in the selected EAs. Nine households are selected from each EA in Niamey and 18 households in each EA outside of Niamey.

Within each EA, households are randomly assigned to one of three types of household questionnaires. Questionnaire 1 collects food consumption data using a 7-day recall method as in the ECVMA. Questionnaire 2 collects food consumption data using usual month methodology of the QUIBB. The last third of households receive Questionnaire 3 which collects food consumption using the 7-day diary method of the ENBC with 7 daily visits to the households. Questionnaire 1 (7-day recall) includes the complete household questionnaire as will be used in the full ECVMA survey. Questionnaire 2 (usual month) and Questionnaire 3 (7-day diary) have lighter household questionnaires excluding the modules on health, non-agriculture household enterprises and non-wage revenues. Households are randomly assigned to one of these three types of household questionnaires, with one-third of households in each EA assigned to each of

them. Thus in all three strata (capital city, other urban, rural), one third of the households (72 households) receive each type of questionnaire.

The data collection is organized into teams. A team consists of a team leader, three interviewers, a data entry operator and a driver. The workload of the interviewer consists of one third of each type of questionnaire, and the actual collection method depends on the nature of questionnaire. This is an important detail as it ensures that enumerators with a certain level of abilities are not assigned to a specific method of data collection, which could introduce some bias in the results. When administering the 7-day recall and the usual month questionnaires, the interviewer makes a single visit to the household to complete the entire questionnaire. For the 7-day diary, since the literacy rate is low in Niger (29.5 percent in 2011 according to the ECVMA main survey), the practice is a daily interview by the interviewer, rather than a personal notebook where the respondent can record his daily expenditures. The interviewer makes daily visits on 7 consecutive days.

As noted above, the methodology used for the data collection mirrors the three methodologies used in the QUIBB, the ENBC and the ECVMA. The experiment is conducted in February and March. During that time of year, the *contre-saison* is coming to an end and the planting season has not yet started.[3] So some food items might not be available at that time of the year and might not appear in data related to the 7-day recall and the 7-day diary, but be present in the data of the usual month methodology. For the purpose of the 2011 ECVMA, the list of food items used for the 2005 QUIBB survey and the classification used during the 2007/08 ENBC have been revised (see Appendix Table A3).[4] This new list is used during the experiment for the 7-day recall and the usual month, and is included as an annex in the interviewer manual to be used to assign codes for the 7-day diary (there is no pre-printed list of food items for the ENBC or this last methodology). For the usual month method which mirrors the QUIBB, the interviewer makes one single visit in the household and collects information on the usual monthly food consumption for each item during the past 12 months and the number of months the item is consumed. For the diary which mirrors the ENBC, the information is not collected in

[3] The contre-saison, or off-season agriculture, is the period after the harvest when there is still water available for small scale irrigation

[4] There are two food items, peanuts with shell (*arachides en coque*) and shelled peanuts (*arachides décortiquées*), in Questionnaire 1 (7-day recall) that do not appear in Questionnaire 2 (Usual month), they have been removed from the calculation of the consumption aggregate; and at any rate, they make up less than 0.2 percent of food consumption.

a diary in the strict sense that the households fill in daily journals of the food consumed by the household on their own. Instead, the interviewer visits the households on a daily basis and collects information on the food consumed on the previous day; so the methodology is not a real diary, but a one day recall. This is somewhat different from the methodology used in ENBC where, for the purpose of nutrition analysis, the interviewer had to weigh the quantity of food consumed in the morning and in the evening and for that reason, he visits the household twice a day (morning and evening).

The final sample for the pilot survey is well covered, with a small percentage of replacement households (2.6 percent). The planned sample size is 648 households, with 72 households per location and questionnaire type. In total 661 households were contacted for participation in the pilot survey. Of the 661, 5 households are eliminated because the households were vacant and 17 are eliminated because of refusals or because of incomplete questionnaires. We eliminate 12 more questionnaires which appear to be outliers regarding food consumption. Those questionnaires present an energy intake higher than the mean plus three times the standard deviation. In the end, 627 households are included in the data for analysis. Among those 627 questionnaires, there are 207 in Niamey, 207 in Tillabery urban and 213 Tillabery rural. By type of questionnaire, there are 205 for the 7-day recall, 211 for the usual month, and 211 for the 7-day diary.[5] (See Table A3 in the Appendix.)

3.2. Descriptive statistics

It is important to determine if any differences that may be found in the consumption aggregate can be attributed to any other factor than the type of questionnaire. Table 1 shows some basic characteristics of the households and the characteristics of the household heads, by type of questionnaires.

When looking at their basic characteristics, the households seem similar on a number of them, but there are also significant differences in some other key characteristics. The household size and the household composition (number of children, number of adults and number of elderly), two of the key welfare explanatory variables are consistent across questionnaire types. When considering the household size, the two extreme cases are 7-day recall and diary which

[5] The distribution of non-response suggests that there is no particular effect by location which can lead to selection bias. In Niamey where there was concern over the ability to find people in their homes, there is only a single household replaced due to absence and refusal.

have a difference in household size of 0.7 that is not statistically significant. But the head of household gender and age show some significant differences across types of questionnaire. Heads of household from 7-day diary data are more often women and are also older than those of the two other types of questionnaire. For example 19 percent of the heads of household in the sub-sample are women versus less than 10 percent for the two other types of questionnaire. This could be an issue for the validity of our results since many studies show the impact of the gender of the head of the household, and sometimes of the age of the head, on the distribution of welfare, and sometimes of the age of the head too. In Niger in 2005 and 2007/08, when controlling for other variables, it was estimated that a household whose head is female has a lower welfare level by 27 percent and 42 percent, respectively, than a household whose head is male (INS, 2008). The same study shows a very low negative correlation of the age of the head of the household and welfare, but still significant. In Guinea, the World Bank (2006) found a similar result. Being a female head of household has a negative correlation with welfare, but the age has a light positive correlation. In Senegal, neither of the two variables has a significant correlation with welfare (Banque mondiale, 2007). Those results lead to the fact that the gender and age of the head of the household being different among the different questionnaires type has to be taken into account when comparing the distributions of welfare in this experiment. In principle, the sample has been designed to avoid those types of disparities, but all errors cannot be avoided especially given the relatively small sample size.

Differences across the questionnaires types are less pronounced for other variables. The distribution of the education and labor market situations of household heads is very similar. The majority of household heads, regardless of the questionnaire type, have no education and more than three-fourths of them have achieved less than primary education. As for labor market, the participation is very high; in most cases, more than 80 percent of the heads have been involved in some type of activity during the last 12 months. Most household heads are employed first as self-employed in non-agricultural activities, second as self-employed in agricultural activities; and third as wage workers. Still from 8.4 percent of the heads for Questionnaire 1 (7-day recall) to 12.5 percent for Questionnaire 3 (7-day diary) are unpaid family workers. This distribution, with self-employed in non-agricultural activities being more important than being involved in self-agricultural activities is consistent with the fact that the sample was urban oriented. But even if

the structure of jobs is similar, there is a difference in heads involved in agriculture, which is relatively higher on the Questionnaire 1 (7-day recall) sample.

Table 1. Household and Head of Household Characteristics by Type of Questionnaire

	7-day recall		Usual month		7-day diary		Statistics	
	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	F	Prob
Household								
Household size	5.77	0.280	6.40	0.269	6.14	0.325	1.310	0.270
Number of children[a]	2.58	0.178	2.78	0.170	2.71	0.194	0.330	0.719
Number of adults[b]	2.87	0.136	3.21	0.162	3.08	0.181	1.377	0.253
Number of elderly[c]	0.22	0.046	0.26	0.043	0.24	0.048	0.195	0.823
Head of household demographic								
Female (%)	0.11	0.027	0.10	0.024	0.19	0.036	2.236	0.108
Average age (years)	45.33	1.248	47.71	1.376	48.93	1.373	1.985	0.138
Married (%)	0.87	0.028	0.86	0.029	0.85	0.032	0.119	0.888
Head of household education								
None	0.65	0.042	0.68	0.040	0.69	0.040	0.300	0.741
Primary	0.18	0.034	0.17	0.033	0.16	0.031	0.100	0.905
Secondary & Post-secondary	0.17	0.031	0.15	0.028	0.15	0.031	0.164	0.849
Head of household labor market								
Labor force participation-30 days	0.83	0.035	0.79	0.036	0.73	0.041	1.747	0.175
Labor force participation-12 months	0.89	0.029	0.84	0.033	0.82	0.036	1.177	0.309
Head of household occupation (12 months)								
Wage worker, except household pers.	0.22	0.034	0.16	0.029	0.17	0.031	1.074	0.342
Other dependent, incl. hh personal	0.08	0.025	0.13	0.031	0.10	0.028	0.730	0.482
Self, Agriculture	0.28	0.041	0.20	0.035	0.19	0.034	1.794	0.167
Self, non-Agriculture	0.31	0.041	0.36	0.042	0.37	0.043	0.641	0.527
Unemployed, not in labor force	0.11	0.029	0.16	0.033	0.18	0.036	1.177	0.309
Number of hhs	205		211		211		627	

Source: ECVMA Pilot survey

Note: [a] Children are defined as those 0-15 years old. [b] Adults are those 16-64 years old. [C] Elderly are those 65 years old and older.

Although the experiment focuses on food consumption, it is also worthwhile to assess if the differences which may be found in the consumption aggregate might come from non-food items. Across questionnaires types, we used an identical list of items, the same recall period, and the same methodology was used to calculate the non-food component of the consumption aggregates. It appears from Table 2 that there is no significant statistical difference in the non-food component of per capita expenditure between the three types of questionnaires. Actually, the overall methodology used for the collection of data could have influenced the results. For example the fact that data are collected on a single day for the first two methodologies gives the interviewer more time to connect with the household which can lead to a better interview. But this is not the case, or this factor is not a determinant. The small differences seen in Table 2 come from the usual sampling and data collection errors, but they cannot invalidate the fact that the three questionnaires are comparable on this issue.

Table 2. Household Non-food Per Capita Expenditure by Questionnaire Type (FCFA)

	Distribution						Test equality of means		
	Mean	P10	P25	P50	P75	P90	SE	F	Proba
7-day recall	143343	36933	50336	85053	144047	199428	255	0.308	0.734
Usual month	144259	41337	53891	79448	125570	273200	250		
7-day diary	131033	40819	56492	80650	154493	293858	124		

Source: ECVMA Pilot survey

4. WELFARE DISTRIBUTION AND POVERTY

4.1. Assessing welfare differences using simple tests

This section explores the impact of the methodology used in data collection on the distribution of economic welfare. For this exercise, we need an appropriate welfare indicator. We use per capita annual consumption, which has been used in the past in Niger. The consumption aggregate contains both food and non-food components. The food component comprises cash expenditures; and the value of auto-consumption and gifts received by the household. Data on food consumption have been collected using the three different methodologies described in the preceding section. The non-food module, which is the same in all three types of questionnaires, collects information on expenditures for non-food items over 7 days, 30 days, 3 months, 6 months and 12 months. Imputed rent has been estimated using a hedonic model for those households living in their own dwelling. A usage value for durable goods has also been

calculated. Health expenditures, which have been part of the Niger consumption aggregate in the past, are excluded, because they were collected only for the 7-day recall methodology. We also assume that because Tillabery is very close to Niamey, the cost of living is identical, so there is no spatial deflator.

When looking at the mean per capita expenditure, Questionnaire 1 (7-day recall) and Questionnaire 2 (usual month) are similar (see Table 3). On average, the difference in the mean per capita consumption is not statistically significant. It is interesting to note that this difference is not identical throughout the welfare distribution. For the households at the bottom (poorest) of the distribution up to the 25th percentile, the average per capita expenditure is higher for Questionnaire 2, and the opposite occurred after this percentile.

However, per capita consumption is significantly higher for the first two methodologies (7-day day recall and usual month) than for the 7-day diary. The average per capita expenditure for 7-day day recall is 32 percent higher than 7-day diary; and this average is 28 percent higher for usual month comparing with 7-day diary. An F-test confirmed that those differences are statistically significant.

The differences between Questionnaires 1 and 2 (7-day day recall and usual month) on the one hand, and Questionnaire 3 (7-day diary) on the other hand, are not only at the mean. There are differences along the distribution of the welfare. When comparing Questionnaires 1 and 3 which one can imagine will be very close because both are based on 7-day food data collection, we find that the differences are very important at the left tail of the distribution, less important in the middle of the distribution, and somewhat important at the right side of the distribution. For example the gap of consumption per capita for Questionnaire 3 compared to Questionnaire 2 represents 44 percent of average consumption of Questionnaire 3 at the 10th percentile, 22 percent at the 25th percentile and only 10 percent at the median. The same gap when comparing Questionnaires 1 and 2 are moderate and more identical at every level of the welfare distribution. Besides, the gap is higher at the right tail of the distribution and it reaches 22 percent at the 90th percentile, while it is not greater than 15 percent at any other level of the distribution.

The difference on the whole distribution of welfare among the three types of questionnaires can be assessed using a K-S test. When comparing two distributions, the K-S statistic quantifies the distance of the empirical distributions between them, with the null

hypothesis that the two distributions are drawn from the same sample. The only hypothesis when implementing the test is that the two distributions to compare are independent, which is supposed to be the case with our sample.

Table 3. Household Annual Per Capita Expenditure by Questionnaire Type (FCFA)

	Distribution						Test equality of means		
	Mean	P10	P25	P50	P75	P90	SE	F	Proba
All									
7-day recall	273612	77123	111745	187486	294482	422551	346	3.3	0.037
Usual month	272288	97876	123637	179689	286507	504801	335		
7-day diary	215375	68126	101001	157138	281686	414157	171		
Food									
7-day recall	130269	30423	48363	87339	169464	267786	127	12.28	0.000
Usual month	128029	38243	55070	93870	154225	233090	112		
7-day diary	84342	14926	28744	68023	121319	170594	67		

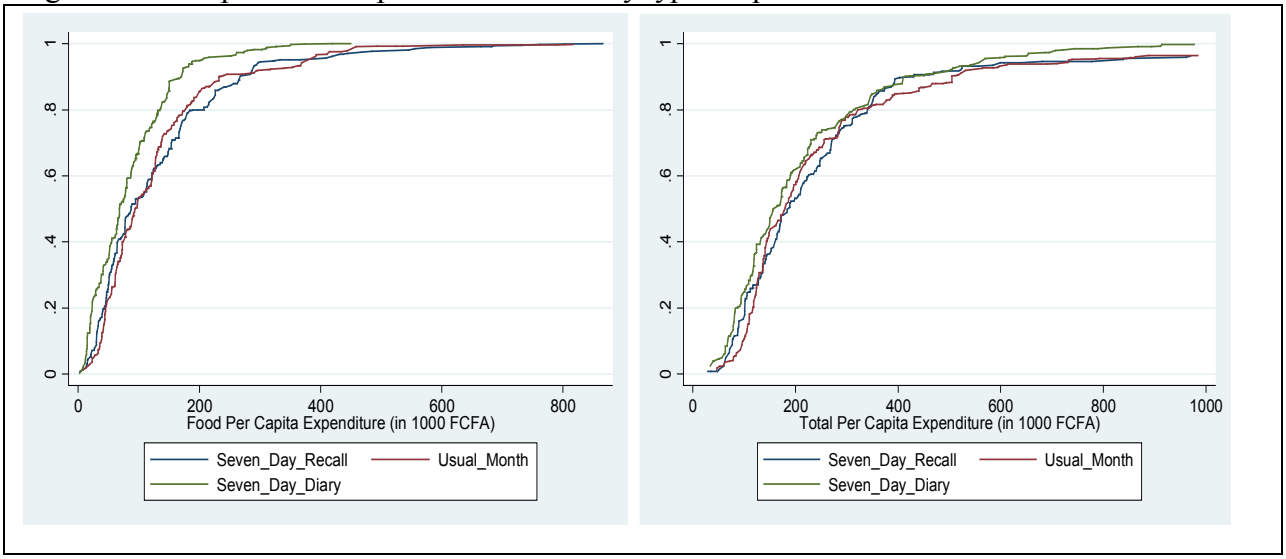
Source: ECVMA Pilot survey

When comparing distributions from the 7-day recall and the usual month methodologies, there is no significant difference. But when comparing distributions from the 7-day recall and the 7-day diary, there is no difference at the 1 percent level, but the difference is clearly significant at the 5 percent level, confirming that distribution of welfare from the 7-day recall is higher than from the 7-day diary. We get an identical result when comparing the usual month and 7-day diary. The differences assessed by a formal test can be visualized using Figure 1. On the plot of the food per capita expenditure on the left, the curve from 7-day diary is always above the two other curves. The two other curves cross at three places. As we see, the curves crossed near at the 25^{th} percentile; at the 80^{th} percentile and again at the 90^{th} percentile, showing that those two distributions are close to each other.

We have seen in the previous section that, except for head of household demographics (namely gender and age), the experiment has not been affected by some external factors. We deal with those household characteristics later, but at this point it is interesting to go further in the comparison of the welfare distribution using some exogenous variable. For this exercise, we use two exogenous welfare indicators. The first one is the per-capita non-food consumption. We have seen in the previous section that the non-food component of the consumption aggregate was not part of the experiment and has not been affected by it. Since the non-food consumption is also an indicator of the welfare, we use it as an exogenous variable while comparing the three

distributions. The second one is an asset welfare indicator. The asset welfare indicator is constructed using 6 housing variables and 16 durable goods variables. All the variables are dummy variables. The weight applied to the 22 variables is the inverse of the number of households having this good for 7-day recall. Interestingly, the distribution of the asset welfare indicator is close for the three types of questionnaires.

Figure 1. Per capita consumption distribution by type of questionnaire

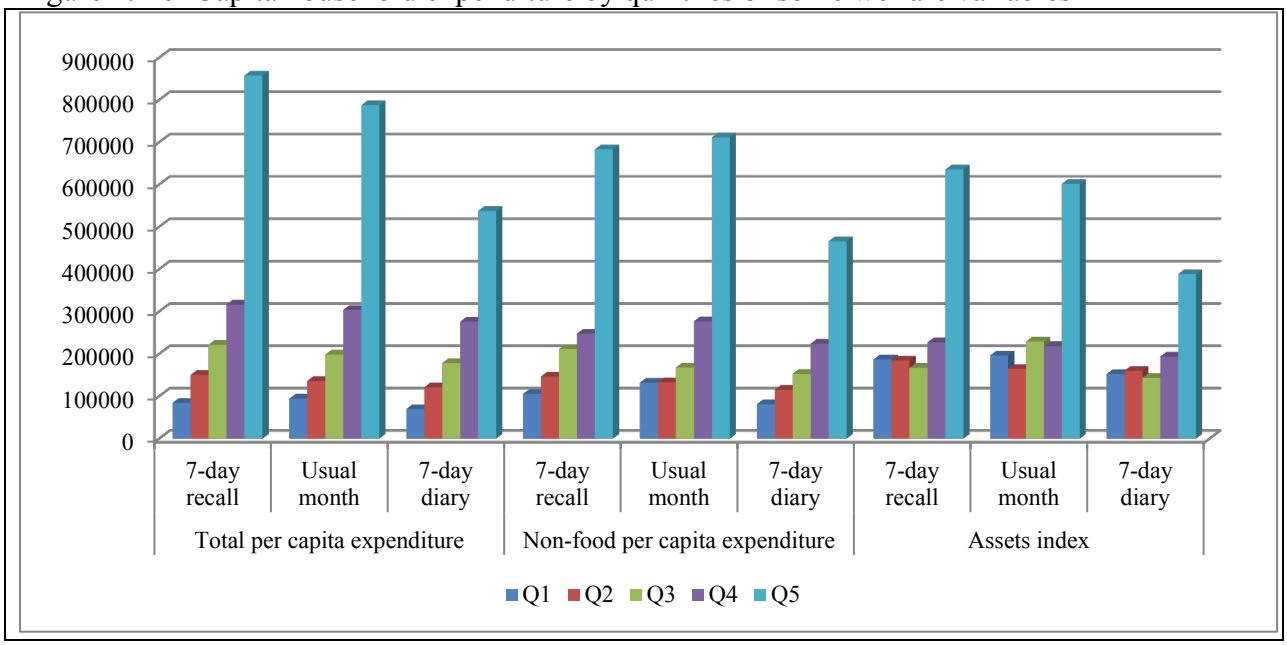


Source: ECVMA Pilot survey

We compare the per capita expenditure using the quintiles of the two welfare variables mentioned above, the per capita non-food consumption and the asset welfare indicator, and we also put in the same figure the quintiles of the per capita expenditure itself (Figure 2). This figure shows that the two welfare indicators are correlated with total consumption per capita; each of these indicators being an increasing function of household per capita consumption. More interesting, at each level of the distribution of each of the two exogenous indicators, the mean of the household per capita consumption for 7-day recall and usual month is greater than the one for the 7-day diary. Looking for example at the quintiles of the non-food consumption, one can see that the differences between the first two methodologies on one hand and the third one on the other is very high on the tails of the distribution, and somewhat less on the middle of the distribution (but with the assets index, we have no such clear pattern). If we accept that the best methodology to collect food consumption is the diary, then according the results above, poor

households and better-off households are the one which overestimate most their food consumption while using an alternative methodology.

Figure 2. Per Capita household expenditure by quintiles of some welfare variables



Source: ECVMA Pilot survey

Knowing that the distribution of non-food consumption (which was not part of the experiment) is statistically the same across three methodologies, we have an indication that the differences in welfare distribution come from the differences in food consumption. For households reporting food expenditures using 7-day recall and usual month, there are similar results for the total amounts spent at the mean. The comparison between the distributions of food per capita expenditure shows that for low food consumption up to the median, food per capita consumption is greater for usual month, and the reverse thereafter. All of the amounts reported by households using 7-day diary are significantly lower than using the other questionnaires. As we did for the total household per capita expenditure, the K-S tests confirm that those differences are statistically significant.

In theory, because households can immediately record an expense when it occurs, the 7-day diary (Questionnaire 3) is considered to be the gold standard for collecting data on food consumption. But the method has its own drawbacks, especially in the case of rural populations where illiteracy is high. To have a better understanding (even partial) in what might happen in

our case, we focus on the percentage of households who reported eating a particular food product, by type of questionnaire. Our interest is in the most consumed products (cereals, meat and fish, cooking oil, eggs, milk and dairy products) that amount to two-thirds of total food consumption expenditure. An interesting point comes from Table 4.

While the percentages of households consuming a specific product for 7-day recall and 7-day diary are similar, the usual month shows higher percentages of households that say that they consumed the items. This can be explained by the time of year at which the pilot survey was administered (as we mentioned earlier, the pilot survey was done in February and March during a time of year when some products are not available) and of the methodology. Households indicating that they did not consume a specific food product during the last 7 days, or during the collection of diary data may not have consumed that product because it was not available at the time of the survey. But during the last 12 months as collected in the usual month methodology, the product was available and households consumed it at some other period of the year.

Table 4. Household Annual Expenditure by Methodology of Data Collection and Nature of Products

	Percent hh consuming			Average expenditure			Average expenditure if >0			Share in total consumption		
	7-day recall	Usual month	7-day diary	7-day recall	Usual month	7-day diary	7-day recall	Usual month	7-day diary	7-day recall	Usual month	7-day diary
Maize	41.2	78.1	43.8	90,295	67,690	55,104	219,359	86,694	125,711	11.2	8.0	10.5
Millet	31.1	75.1	34.8	71,232	106,364	47,793	228,682	141,540	137,248	8.8	12.6	9.0
Rice	54.6	87.2	59.8	103,997	88,213	77,242	190,492	101,113	129,263	12.9	10.5	14.6
Other cereals	4.9	21.4	11.0	10,616	7,533	8,977	214,483	35,144	81,794	1.3	0.9	1.7
Bread/pasta	47.6	70.3	31.9	38,371	28,760	14,238	80,576	40,884	44,586	4.7	3.4	2.7
Meat & fish	66.1	91.5	54.6	108,691	121,150	58,003	164,404	132,355	106,206	13.4	14.4	11.0
Cooking oil	78.6	93.3	87.6	31,784	42,721	38,048	40,461	45,811	43,414	3.9	5.1	7.2
Milk product	43.7	69.6	49.5	22,177	25,506	15,275	50,792	36,655	30,880	2.7	3.0	2.9
Eggs & other	18.1	47.2	11.8	10,545	12,260	2,532	58,261	25,989	21,473	1.3	1.4	0.5
Other food	99.6	100	100	320,869	341,781	210,674	322,197	341,781	210,674	39.7	40.6	39.9
Total	100	100	100	808,578	841,978	527,885	808,578	841,978	527,885	100.0	100.0	100.0

Source: ECVMA Pilot survey

This is another drawback of the 7-day diary methodology: it is difficult to find all the products at any given time of the year. Even if the data were collected in November, after the harvests, we would have encountered a similar issue with some products. A shorter reference period leads to not taking into account the expenditure allocated to products, which are rare or even not available on the market at the time of data collection. One might counter that assuming that households substitute rare products for those available, the household's disposable income is allocated to other products, and finally consumption levels would be close. But that assumption is based on the idea that all consumption comes from purchases, which is not the case. An important part of consumption comes from own consumption and gifts, and this part is underestimated in the case of scarcity of the product. In addition there is the issue of consumer preference. The household can consume a product because it is available and may not be willing to use this money for a substitute product. Thus this factor may partly explain the fact that the level of household consumption of the usual month approach is higher than that of the 7-day diary approach.

One of the main results at this point is the fact that the average expenditure from the 7-day recall methodology is higher than the 7-day diary and that this difference is important. One would expect the opposite, the 7-day recall being lower and the two methodologies giving closer results on average expenditure. The logic underlying this expectation is that recording information using the diary is supposed to be more precise, while a recall, since the respondent has to use his memory, is supposed to lead to some forgotten information. We explore further the sources of the differences between the diary and the 7-day recall by looking at incidence of reporting expenditure.

We proceed to the comparison of the total number of observations recorded for each of the two methodologies (Table 5). There are nearly 14,000 observations for the 7-day diary, but as we know, a product can be purchased several times a week (even several times a day). When we aggregate these observations for a single product in each household, we obtain 4,121 observations, which is very close to the 4,152 observations obtained for the 7-day recall. Moreover, we generally find the same products in the two methodologies. Thus the difference might not come from some products having been purchased, but not recorded. But it turns out that the average expenditure per item (not weighted) with the 7-day recall approach is 52 percent higher than the 7-day diary. This result can be in parallel with the fact that the number of

observations of the 7-day diary decreases gradually from the first day to the fifth day. Although it increases somewhat during the last two days of data collection, the number of observations of the first day is no longer reached later.[6] Since there is no technical reason to explain this decrease, it can be attributed to the survey methodology, perhaps the fatigue of the respondent or the non-possession of part of information. It might happen for instance that for products which are purchased multiple times a day (items like doughnuts, bread, etc. can be purchased by different family members the same day) or meals which are taken outside the household, the respondent does not have all the information (Wiseman, Conteh, and Matovu, 2005).

Table 5. Comparison of the number of observations and expenditure per methodology

		Expenditure		
	Frequency	Total	Mean	SD
7-Day diary				
Day 1	2,128	404,575	190.1	711.6
Day 2	2,025	344,315	170.0	449.0
Day 3	2,050	331,521	161.7	384.4
Day 4	1,908	307,955	161.4	284.3
Day 5	1,843	271,520	147.3	245.2
Day 6	1,884	301,145	159.8	314.1
Day 7	1,980	363,045	183.4	600.9
Total	13,818	2,324,076	168.2	462.5
Collapse for a unique item per household	4,121	2,324,076	564.0	1195.8
7-Day recall				
Total	4,152	3,532,994	850.9	1985.1
Usual month				
Total	7,487	2.14E+07	2,862.7	5,190.3

Source: ECVMA Pilot survey

Even with this drawback on the side of the 7-day diary methodology, one cannot exclude the possibility of an overestimation of household expenditures when using the 7-day recall method. One can, for example, think that because of the memory effect, for some items which are frequently purchased during the week, the daily expenditure is multiplied by seven and not necessarily by the exact number of days in which the product has been purchased. To examine whether there might be an overestimation of expenses from the latter approach, we calculate the

[6] The same profile of the number of observations per day has been observed also during the 2007 ENBC in Niger. There has been a decrease of 9 percent of the number of observations between the first day of data collection and the 7th day of data collection.

number of calories per person per day from each of the two methodologies. The calculation is performed on the sample, without applying any extrapolation. The idea underlying our analysis is that a person should consume around 2,100 calories per day. If it deviates too much from the standard, the data are certainly unrealistic. The basket used (39 products) represents 80 percent of food consumption for the 7-day recall and 76 percent of the 7-day diary. We get an amount of 1,558 calories for 7-day diary and more than 2,462 for 7-day recall. If we assume that the part of the consumption not taken into account in the calculation (meals taken outside and some less consumed products for which prices are not available) provides the same amount of calories, we get 1,900 kilocalories for the 7-day diary and more than 2,700 kilocalories for the 7-day recall. Those figures tends to show that if in the specific case of Niger the 7-day diary underestimates the food consumption, the 7-day recall on the other hand tends to overestimate it.

Table 6. Comparison of the Calories intake per methodology of data collection

	All sample		1 percent trim		5 percent trim		Windsoring 5 percent	
	7-Day recall	7-Day diary	7-Day recall	7-Day diary	7-Day recall	7-Day diary	7-Day recall	7-Day diary
Mean	2442	1558	2212	1504	1887	1419	2134	1494
Median	1408	1363	1408	1363	1407	1363	1408	1363
# Obs.	213	212	207	208	192	192	213	212

Source: ECVMA Pilot survey

4.2. More on welfare impact using some regressions

The preceding analysis shows that the methodology used to collect data on food consumption makes significant differences in the welfare indicator (as measured by per capita annual consumption). In this sub-section we use linear regression techniques to explore further those differences. The first model is the regression of the logarithm of the total per capita food consumption (the welfare indicator) on the type of questionnaire (Model 1). The rationale of this model comes from the fact that since the experiment is a real randomization, the only variable that affects food consumption is the type of questionnaire. Other factors such as the interviewer should have no effect. Actually since the type of questionnaire is a dummy variable, the model is not different from a simple test of means among the three methodologies. In the previous section, we noted that for two characteristics of the head of the household, namely gender and age, there are statistically significant differences across households assigned to the three methodologies. To control for these differences, we introduce the interaction between the type of questionnaire and

each of those variables in a second set of regressions in the second model. So in a second model, we interact the type of questionnaire with gender (Model 2). In a third model, instead of the gender variable we have age (Model 3); and the last model includes both variables interacting with the type of questionnaire (Model 4).

Formally, y_i is the logarithm of the total per capita food consumption for household i, Q_{ki} is the questionnaire type (k=1, 2, 3; with type 1 being the 7-day recall, type 2 the usual month and type 3 the 7-day diary), $Female_i$ and Age_i are respectively the dummy for female head of household and Age of the head of the household and U_i is the error term. We estimate two models. The first model is (Model 1):

$$y_i = \alpha + \sum_{k=1}^{3} \beta_k Q_{ki} + U_i$$

And the Fourth model is (Model 4):

$$y_i = \alpha + \sum_{k=1}^{3} \beta_k Q_{ki} + \sum_{k=1}^{3} \gamma_k Q_{ki} * Female_i + \sum_{k=1}^{3} \delta_k Q_{ki} * Age_i + U_i$$

Models 2 and 3 are derived from Model 4 by removing respectively either the interaction between the type of questionnaire and the gender of the head of the household, or the interaction between the type of questionnaire and age of the head of the household. The comparison between the three methodologies is based on tests on the estimates of the parameters β.

When considering food consumption and the first regression (which has questionnaire dummies as the only explanatory variables) and taking the 7-day recall as the reference methodology, the 7-day diary has an estimated gap of -28 percent (exponential of -0.332), and the estimated parameter is statistically significant even at the 1 percent level. Always taking the same methodology as the reference, the difference with the usual month approach is not significant, confirming the results found using simple statistical tests.

Table 7. Regression of per capita expenditure on type of questionnaire*

	Dependent variable: Log of food per capita			
	Model 1	Model 2	Model 3	Model 4
Questionnaire				
7-Day recall	ref			
Usual month	0.071	0.102	-0.464	-0.515*
	(0.091)	(0.096)	(0.308)	(0.308)
7-Day diary	-0.332***	-0.282***	-0.804**	-0.876***
	(0.092)	(0.099)	(0.317)	(0.317)
Questionnaire#Female				
7-Day recall*Female		0.459**		0.598***
		(0.207)		(0.211)
Usual month*Female		0.197		0.203
		(0.214)		(0.219)
7-Day diary*Female		0.007		0.013
		(0.166)		(0.167)
Questionnaire#Age				
7-Day recall*Age			-0.011**	-0.014***
			(0.005)	(0.005)
Usual month*Age			0.000	-0.000
			(0.004)	(0.004)
7-Day diary*Age			-0.001	-0.001
			(0.004)	(0.004)
Constant	11.5	11.5	12.0	12.1
	(0.065)	(0.069)	(0.226)	(0.226)
# Observations	627	627	627	627

Source: ECVMA Pilot survey

Note: Standard errors in parentheses. *** indicates significant difference compared with 7-Day recall at 1 percent; ** at 5 percent; and * at 10 percent.

Introducing the possibility of the gender of the head of the household interacting with the type of questionnaire tells the same story, with on average a lag of 25 percent of the per capita food consumption from the 7-day diary methodology compared to the 7-day recall. When we remove the gender and introduce the second variable which is the age of the head of the household, the difference between the food per capita from the 7-day recall and the 7-day diary is even higher, suggesting that part of the gap was captured by the head of the household age effect. But the gap from this regression seems large (55 percent) compared to the previous ones.

At this point, the partial conclusion is that in the case of this Niger experiment of measuring food consumption using three different methodologies, the 7-day recall and the usual month method give higher values than the 7-day diary approach. This conclusion is robust, using

simple statistics tests and regressions. The next sub-section explores the impact on poverty and inequality.

4.3. Poverty and inequality

The trend in poverty, not its level, is what matters for the evaluation of public policies. With the MDG goal to halve poverty by 2015 from its 1990 level, the measurement of progress toward that goal is the poverty trend, i.e. the change in level over time (World Bank, 2011). It is thus important to have consistent poverty numbers over time. But as we have seen in the previous section, different methodologies of data collection result in different distributions of welfare, and potentially different levels of poverty. In this study, we use three different poverty lines to compute standard poverty indicators; the poverty headcount, the poverty gap and the squared poverty gap. The first poverty line (low) corresponds to $1 dollar per capita per day in PPP 1993; this gives 154,539 FCFA per capita per year in local currency. The second one (medium) is the national poverty line computed by the INS and the World Bank using the 2005 QUIBB survey, inflated by the ratio of the consumption price index between 2005 and 2011; this second poverty line is 166,087 FCFA per capita per annum (INS, 2006; INS, 2008). The third poverty line (high) corresponds to $1.25 dollar per capita per day in PPP 1993, which gives 193,174 FCFA per capita per year in local currency. Since the results using the three poverty lines converge, we present only the one with the national poverty line.

Table 8. Poverty indicators by type of questionnaire using the national poverty line

	7-Day recall		Usual month		7-Day diary		Test	
	Mean	SE	Mean	SE	Mean	SE	F	Level
Poverty headcount	0.425	0.052	0.465	0.050	0.510	0.051	1.470	0.232
Poverty Gap	0.150	0.024	0.136	0.018	0.199	0.028	4.940	0.007
Squared poverty gap	0.070	0.014	0.053	0.010	0.100	0.019	7.110	0.001

Source: ECVMA Pilot survey

The results in Table 8 show important differences among the three methodologies for all poverty indicators. When considering the poverty headcount, poverty is 8.5 percentage points higher with the 7-day diary compared to the 7-day recall; and 4 percentage points higher with the 7-day diary compared to the usual month. When implementing a formal test, the difference among the poverty headcounts using the three methodologies is not statistically significant; but when doing pairwise tests, the difference between the 7-day recall and the 7-day diary is

significant at the 10 percent level. As for the two other poverty indicators, the gap among the three methodologies go in the same direction and the differences are even significant at 1 percent level. One of the reasons why some of the differences are not statistically significant even if they appear large in absolute value is that the standard errors are high; the size of our sample is only 200 households for each methodology. Now because of the small sample, the standard errors are high and the magnitude of the differences is probably exaggerated. This should not undermine the credibility of our results especially considering that other studies with higher sample-sizes have obtained similar findings.

Table 9. Inequality indicators by type of questionnaire

	7-Day recall		Usual month		7-Day diary	
	Mean	SE	Mean	SE	Mean	SE
Gini	0.457	0.028	0.431	0.027	0.393	0.017
Atkinson(1)	0.297	0.030	0.264	0.029	0.231	0.019
Theil(1)	0.423	0.054	0.368	0.069	0.257	0.022

Source: ECVMA Pilot survey

Not only did we find that poverty measures are not consistent across methodologies, inequality is also inconsistent. We use three different inequality measures, the Gini index, the Atkinson index and the Theil index. The 7-day recall methodology which has the higher level of expenditure and the lowest poverty indicators comes up with the highest inequality, regardless of the inequality measure used. The differences between the 7-day recall and the usual monthly recall are not important, but they are when compared with the 7-day diary. Using a Gini index, up to 6 percentage points are just differences in methodology of data collection. These kinds of trends have been found in other countries when the methodology of data collection has changed between two surveys, including in Niger. In the case of Guinea for example, the methodology of food data collection in the 2002/03 national survey was built on 10 visits during a month, each visit collecting data for the last 3 days. The 2007 and 2012 surveys use the usual month approach. The Gini index is 0.403 in 2002/03; 0.312 in 2007 and 0.317 in 2012 at the national level. This trend in inequality measures in Guinea is difficult to understand; one was expecting inequality to fall or to remain constant between 2002/03 and 2007; because the rural population, the poorest, was less exposed to the crisis during the years 2004-2007 than the urban one. The correct justification of this trend is just the methodology problem (Banque mondiale, INS, 2012). In Niger the Gini index was supposed to decrease by more than 7 percentage points between

2005 (usual monthly) and 2007/08 (7-day diary); we think that part of this decrease, if not all, can be attributed to difference in methodologies (INS, 2008).

5. CONCLUSIONS

In this paper, we assess the impact of three methodologies of food data collection on the welfare distribution, and on poverty and inequality measures using a randomized experiment. We assigned one of the food data collection methods to each of the three randomly formed groups. The first methodology is the 7-day recall, the second one is the usual month and the last one used a 7-day diary. We find that there is a difference in the distribution of welfare, between on one hand, the 7-day diary and the usual month and, on the other hand, the 7-day diary approach. The gap on annual per capita consumption between the 7-day recall and the 7-day diary is estimated at the mean being one third of per capita annual consumption. The differences found between the 7-day recall methodology and the 7-day diary comes from both some overestimation of food consumption on the 7-day recall and an underestimation on the 7-day diary. These differences lead to differences in poverty and inequality measures even when using alternate poverty lines.

This last point is important since trends are important in evaluating the impact of public policies on poverty. In fact the three methodologies used in this study have been implemented in different household surveys in Niger blurring the real trend of poverty indicators, particularly when those trends are difficult to reconcile with economic growth. So to get a better understanding of poverty trends in Niger and assess how the country will do related to the first MDG which is to halve poverty from its 1990 level, it might be useful to revisit poverty trends in Niger. Revisiting the poverty trends used in Niger is not enough, one has to also prepare and ensure comparability of poverty measures in the future. The 2011 national household survey has used the 7-day recall approach. It is important to replicate the same method in the future, to be sure to get a more robust comparison on poverty.

Since the magnitude of the differences on poverty measures seems high, it might also be interesting to repeat the experiment in some other country facing the same problem (Guinea, Mali, Burkina Faso, etc.) with a larger sample. It can also be interesting to make these comparisons in the case where the 7-day recall and the 7-day diary methods are implemented with more than one round during the year.

BIBLIOGRAPHY

Backiny-Yetna, Prospere. 2011. "Rapport d'exécution de l'enquête pilote." World Bank. Processed.

Banque mondiale. 2006. Guinée : Diagnostic de la Pauvreté, The World Bank, Washington D.C.

Banque mondiale, Institut National de la Statistique (2012). Pauvreté et inégalité en Guinée 1994-2012 : Analyse à partir des enquêtes auprès des ménages (Version préliminaire). Ministère du Plan, Conakry.

Banque mondiale. 2007. Tendance et Profil de la Pauvreté au Sénégal 1994-2006 (version préliminaire), The World Bank, Washington D.C.

Beegle, Kathleen, Joachim De Weerdt, Jed Friedman and John Gibson. 2010. "Methods of Household Consumption Measurement through Surveys: Experimental Results from Tanzania." World Bank, Policy Research Working Paper, No. 5501. Washington, DC.

Bradburn, Norman M. 2010. "Recall Period in Consumer Expenditure Surveys Program." Paper presented at the Consumer Expenditure Survey Methods Workshop, December 8-9. Council of Professional Associations on Federal Statistics, Alexandria, Virginia.

Deaton, Angus. 2001. "Counting the World's Poor: Problems and Possible Solutions." *The World Bank Research Observer* 16(2): 125-147.

Deaton, Angus, Margaret Grosh. 2000. "Consumption." In Paul Glewwe and Margaret Grosh, eds., *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study*. Washington, DC: The World Bank.

Deaton, Angus, Salmon Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analyses." World Bank, LSMS Working Paper, No. 135. Washington, DC.

Gibson, John. 1999. "Are Poverty Comparisons Robust to Changes in Household Survey Methods?" Processed.

Institut National de la Statistique (2006). Questionnaire des indicateurs de base du bien-être, Profil de pauvreté – Rapport d'analyse. Ministère de l'économie et des finances, Niamey.

Institut National de la Statistique (2008). Tendances, profil et déterminants de la pauvreté au Niger 2005-2007/08. Ministère de l'économie et des finances, Niamey.

Jolliffe, Dean. 2001. "Measuring absolute and relative poverty: The sensitivity of estimated household consumption to survey design." *Journal of Economic and Social Measurement* 27 (2001): 1-23.

Lanjouw, Jean Olson, Peter Lanjouw (2001). "How to Compare Apples and Oranges: Poverty Measurement based on Different Definition of Consumption". Review of Income and Wealth, 47 (1), pp. 25-42.

McWhinney, Isabel, Harold E. Champion. 1974. "The Canadian Experience with Recall and Diary Methods in Consumer Expenditure Surveys." *Annals of Economic and Social Measurement* 3(4): 411-435.

Saunders, Christopher, Christiaan Grootaert. 1980. "Reflections on the LSMS Group Meeting." World Bank, Living Standards Measurement Study Working Paper No. 10, Washington, DC.

Scott, Chris, Ben Amenuvegbe. 1990. "Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana." World Bank, Social Dimensions of Adjustment in Sub-Saharan Africa, Working Paper No. 6, Washington, DC.

Tarozzi, Alessandro. 2007. "Calculating Comparable Statistics from Incomparable Surveys, with an Application to Poverty in India." *Journal of Business & Economic Statistics* 25(3):314-336.

Wiseman, V., L. Conteh, F. Matovu. 2005. “Using diaries to collect data in resource-poor settings: questions on design and implementation.” Health Policy and Planning 20(6): 394-404.

World Bank. 2012. Niger: A Poverty Assessment. Report No. 61393. The World Bank, Washington, D.C.

APPENDIX – Experiment Details

Table A1. Organization of the Interviewer Work Load

Questionnaire/ Household	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
	Listing the EA									Leave the EA
7-day recall HH1		Interview								
7-day recall HH2								Interview		
Usual month HH3			Interview							
Usual month HH4							Interview			
7-day diary HH5		Introduction	Daily food expenditures Household composition,	Daily food expenditures Education, Employment,	Daily food expenditures Dwelling, Revenues, Transfers	Daily food expenditures Shocks, Non-food expenditures	Daily food expenditures Agriculture	Daily food expenditures Revision	Daily food expenditures	
7-day diary HH6		Introduction	Daily food expenditures Household composition	Daily food expenditures Education, Employment,	Daily food expenditures Dwelling, Revenues, Transfers	Daily food expenditures Shocks, Non-food Expenditures	Daily food expenditures Agriculture	Daily food expenditures Revision	Daily food expenditures	

Note: Each interviewer was responsible for interviewing six households per EA. Two households for each type of questionnaire. For 7-day recall and usual month Questionnaires, the entire interview was administered in one day per household. For 7-day diary Questionnaire, seven daily visits were made to the household to collect daily food expenditure information and one or two other modules per day. The schedule was designed so that an interviewer could administer multiple households and multiple questionnaires type in a single day.

Table A2. Questionnaire Contents

7-day recall	Usual month	7-day diary
Socio-demographic characteristics of household members*	Socio-demographic characteristics of household members*	Socio-demographic characteristics of household members*
Education*	Education*	Education*
Health	Employment*	Employment*
Employment*	Dwelling characteristics*	Dwelling characteristics*
Non-agricultural enterprises	Durable goods*	Durable goods*
Dwelling characteristics*	Non-food expenditures*	Non-food expenditures*
Durable goods*	Transfers*	Transfers*
Non salary revenues	Shocks and survival strategies*	Shocks and survival strategies*
Non-food expenditures*	Food consumption – 12 month average	Food consumption – 7-day daily diary
Transfers*		
Shocks and survival strategies*		
Food security		
Food consumption – 7-day recall		
Complement to food consumption of the last 7 days		

Note: *Modules for Socio-demographic characteristics of household members, Education, Employment, Dwelling characteristics, Durable goods, Non-food expenditures, Transfers, and Shocks and survival strategies were identical in all three questionnaire types.
All households received an agricultural questionnaire that was identical.
A community questionnaire was applied in all enumeration areas.

Table A3. Distribution of surveyed households by type of questionnaire

	Total	7-day recall	Usual month	7-day diary
Interviewed				
Niamey	207	69	70	68
Tillabery urban	207	66	69	72
Tillabery rural	213	70	72	71
Total	627	205	211	211
Refusals/Incomplete				
Niamey	4	1	1	2
Tillabery urban	6	2	2	2
Tillabery rural	7	2	4	1
Total	17	5	7	5
Dropped as outliers(*)				
Niamey	5	2	2	1
Tillabery urban	4	3	1	0
Tillabery rural	3	3	0	0
Total	12	8	3	1

Note: Outliers are household having a daily per capita calorie intake greater than the mean plus 3 times the standard deviation.