Measuring macro- and micronutrient consumption in multi-purpose surveys: Evidence from a survey experiment in Tanzania

Hannah Ameye^{1*}
Joachim De Weerd^{1,2}
John Gibson^{1,3}

Abstract

The nutrition transition in developing countries has increased interest in moving the measurement and analysis of nutritional choice beyond calories to a more complete understanding of macro- and micronutrient consumption. To help move the literature on data collection forward we randomly assigned six different survey modules to measure food consumption across Tanzania, three using diaries and three using recall methods. These modules were chosen to reflect the variety of modules currently in use in multi-purpose household surveys collecting food consumption expenditures in some detail at a national scale. Using the percentage consumed relative to daily recommended intake of calories, protein, fats, sugars, fiber and 16 micronutrients, we find that the household diaries and 14-day recall systematically underestimate nutrient consumption compared to the benchmark personal diaries. The 7-day recall performs better, on average. We also calculate minimum cost diets and cost-of-basic needs food poverty lines in each region, the prevalence and depth of food poverty according to these lines, and the cost of targeted transfers designed to eliminate food poverty. The 7-day recall provides reasonable results whereas the 14-day recall grossly overestimates food poverty and therefore raises the cost of hypothetical monetary transfers needed to alleviate this poverty.

Keywords: Consumption, Household surveys, Nutrition, Poverty

¹ LICOS Centre for Institutions and Economic Performance, KU Leuven, Belgium

² IOB, Institute of Development Policy, University of Antwerp, Belgium

³ Department of Economics, University of Waikato, New Zealand.

*Corresponding author: hannah.ameye@kuleuven.be

Joachim De Weerd: Joachim.DeWeerd@uantwerpen.be

John Gibson: jkgibson@waikato.ac.nz

1. Introduction

Nutritious diets are important for human health. Health is not only a prime development outcome, but also feeds into many, if not most, other dimensions of well-being. The development community used to be primarily concerned with hunger, focusing on how many people had insufficient dietary energy and considering what the best ways were to ensure adequate calorie consumption. With economic development came a ‘nutrition transition’ adding, alongside hunger, excess calories and nutrient deficiencies to the list of public health concerns in developing countries.¹ The World Health Organisation estimates that one out of three people in the world suffer from either wasting, stunting, vitamin and mineral deficiency, being overweight or obese, or from a diet-related non-communicable disease (WHO, 2017).

At the same time that increased attention has been directed to the nutrition transition, researchers and policy-makers have also gained a deeper appreciation of the physiological connections between nutritional consumption and in-utero development, children’s cognitive and physical development and adult health. Consequently, we now have a reasonable idea of what a healthy and environmentally sustainable diet may look like, although it remains unclear how to make it universally available and affordable (Willett et al., 2019; Hirvonen et al., 2019; FAO et al., 2020). Policies that aim to improve nutrition require an understanding of how incomes, prices and preferences come together to determine the nutritional choices that consumers make. Understanding such choices lies at the heart of what economics is about and it is therefore no surprise that there is

¹ See, for example, Ramakrishnan (2002), Black et al. (2013), Forouzanfar et al. (2013), Pingali and Sunder (2017), Popkin (2017), Perez-Escamilla et al. (2018) and Imamura et al. (2015).

increased interest to move beyond calories towards a ‘new economics of nutrition’ (Finaret and Masters, 2019).

Studying the economics of nutrition is complex as it requires us to understand the consumption of a large number of different macro- and micronutrients that together with individual and environmental circumstances will determine nutritional status. This means moving beyond measuring the effects of nutrient deficiencies (which is what we do when we collect anthropometric data or take haemoglobin test kits to the field) to getting a more complete picture of nutritional consumption, broken down into specific food items and their respective macro- and micronutrient contents. Moreover, these detailed nutrition data should, preferably, be part of a survey that collects additional relevant socio-economic information as this information is needed to model the determinants of consumption.

Recognising that data collection budgets are finite, one promising avenue would be to assess nutrient consumption through on-going initiatives. A good starting place would be the food consumption modules included in household consumption and expenditure surveys (Zezza et al., 2017), which are conducted at national level across the developing world. They are repeated over time and often include a wide range of other socio-economic information on households and individuals, such as health, education, agriculture and so forth. These modules were not designed to measure nutritional content, being primarily included to measure the monetary value of consumption, from which poverty and inequality indicators can be derived. However, given that food constitutes such a large portion of the consumption expenditures of poor people, such surveys will tend to collect a wealth of food consumption data. For example, food consumption is typically recorded from various sources (whether purchased, home produced or given)

and information on quantities consumed of individual food items is often available, which can be linked to its nutritional content. In short, household consumption and expenditure surveys (HCES) are ubiquitous and hold the potential to reveal something about the nutritional content of what people are eating (De Haen et al., 2011; Fongar et al., 2019).

One concern is that the design of these surveys varies in several dimensions and there is little guidance on the best data collection method (Fiedler et al., 2008; Carletto et al., 2013). This design variation raises concerns about the analytical consequences of using different survey methods and about the potential fragility of results that depend upon comparing across surveys that use different methods. Survey methods differ across countries and some countries also change their survey method, which introduces differences over time. For example, recent FAO and World Bank (2018) guidelines for food data collection in household surveys recommend using a 7-day recall, necessitating a switch in survey method for many countries. Another example is the adoption of mixed recall periods in India's National Sample Survey, with some items having a 7-day recall and others get a longer recall period, whereas previously it was 30-days for all items, with this change motivated by results from a survey experiment (NSSO, 2002).

To provide evidence on these questions, we randomly assigned six different modules to measure the quantity of food consumption across Tanzania, three using diaries and three using recall methods. The modules differ by survey observation period, by length of the food recall list, by type of survey reporter (individual reporting or a single reporter per household) and by frequency of interviewer visits. Previous analyses of this same experiment have shown that poverty and inequality measures are sensitive to survey design (Beegle et al., 2012) and that errors tend to be mean-reverting (Gibson et al., 2015). This has important implications for poverty comparisons over space and time and for

analyses using surveyed consumption data. Gazeaud (2020) uses the experiment to show how the predictive performance of proxy means tests is affected by survey design. Two papers have focussed on food consumption specifically. First, Friedman et al. (2017) show how errors in the extensive margin (consumed yes or no) and intensive margin (the monetary value) of food consumption reporting can off-set each other. De Weerdt et al. (2016) looked at how survey design impacted the accuracy of calorie consumption. The current paper is the first to look at micro- and macronutrient consumption.²

In contrast to previous studies we provide three innovations. First, we have access to detailed descriptions of the actual foods reported in diaries by the survey respondents. This allows us to disaggregate more finely, using 99 individual food items that we link to Tanzanian food composition tables (Lukmanji et al., 2008). Second, with this link we are able to assign macro- and micronutrient information to all consumed items. From these data we calculate the percentage consumed relative to dietary reference intakes of calories, protein, fats, sugars, fiber and 16 micronutrients, taking into account the age and gender composition of each household. Our third innovation is to use linear programming to calculate the regional cost of least cost diets that meet all of the macro- and micro-nutrient recommendations. These go beyond the more typically used cost-of-basic-needs food poverty lines. Our motivation for these cost calculations is that survey design not only affects measurement of how apparent consumption compares to recommendations, but also affects calculated costs of potential public transfers that are designed such that recipients could afford minimally adequate diets.

² A final paper using data from this same survey is Dillon et al. (2020), that exploits the transaction-level granularity of the consumption data to understand why households tend to buy very small quantities frequently, even while slightly larger purchases made less frequently could lead to substantial savings.

Our analysis uses the data from the highly supervised individual diaries (based on daily checking on respondents) as our ‘true’ measure of the inflows and outflows of food into the household. When comparing the data from the other five survey modules (household diaries and recall modules) to our ‘true’ food consumption module, the apparent consumption of total and animal-sourced protein, cholesterol and fiber are more sensitive to the use of different consumption modules than is the case for calorie consumption. Thus, previous findings that focus solely on calories, such as De Weerd et al (2016), may overstate the robustness of nutritional indicators to the variation in survey methods. We also find that using household diaries rather than the highly monitored individual diaries results in the understatement of vitamin and mineral consumption.

Whilst household diary modules underestimate macronutrient consumption slightly, 14-day recall modules do so more substantially, underestimating calories by 17%, protein by 40%, fat by 27%, and carbohydrate consumption by 11%. Conversely, the 7-day long-list recall module performs similarly to the household diaries, although overestimating the consumption of carbohydrates, sugar and fiber consumption by about 20%. Vitamin and mineral consumption for 7- and 14-day recall modules is also generally underestimated, except for nutrients derived from higher reported cereal consumption. Important mechanisms that may contribute to the degree of mismeasured nutrient consumption, such as respondent’s education, household size, wealth (based on an asset index) and whether they live in urban or rural locations, are also discussed.

When exact macro- and micronutrient consumption cannot be measured, food expenditures can be compared to food poverty lines instead. We construct two such lines. First, the minimum cost (MC) diet, derived through linear programming, is calculated per region to account for price differences across Tanzania. Dar es Salaam is the most

expensive region in our sample, with a minimum cost macro- and micronutrient-fulfilling diet of 412 TZS per person, per day (equivalent to US\$ 0.36 per day). The resulting food-poverty prevalence rate varies by survey module. Consistent with the results for understated nutrient consumption, the 14-day recall overestimates the prevalence of extreme food poverty, suggesting that 11% of people in our sample live in households where total food expenditure is below what is needed to buy the minimum cost diet. This is more than twice the rate of extreme food poverty shown with the benchmark individual diary. In contrast, the frequently checked household diary and the 7-day long-list recall perform most closely to the individual diaries at the national level, although with several discrepancies at the regional level.

We then determine how much it would cost to transfer the amount of the shortfall to each household under the food poverty line. Using our benchmarks, the best estimate is USD 51 million per year, with the 7-day recall coming closest with an estimate of USD 69m. The 14-day recall module, by contrast, gives a much more exaggerated annual figure of USD 159m for the same hypothetical programme, working through both an overestimation of the number of people under the MC line and their average shortfall.

Along the same lines, we also calculate a food poverty line based on the typical diet of the poor (specifically, the poorest two quintiles of households) scaled to provide exactly the daily calorie requirement. There is much greater regional variation in this line than in the one calculated through linear programming. That is because this food poverty line reflects actual consumer choice, while the linear programming exercise of finding the cheapest combination of foods that satisfy dietary reference intakes does not take any taste considerations into account. We see that the food poverty line is two to three times higher than what comes out of the linear programming calculations in all regions. But,

the discrepancy is especially large in Dar es Salaam and Pwani, indicating a stronger preference for more expensive micro- and macronutrients in these places. Together, these two sets of calculations suggest a practical implication of the lack of robustness across the different survey designs; setting possible transfer programs to eliminate food poverty using some survey designs may produce excessive budgetary requests, which could affect the political feasibility of attempting to eliminate hunger with targeted transfers.

Overall, we find that the 7-day recall with a long list of food items is relatively good at predicting average nutrient consumption, as well as providing decent predictions for the prevalence and depth of food poverty. However, the performance of the long list 7-day recall does seem to depend on household characteristics, so it needs to be used cautiously when studying the correlates of nutrient consumption.

2. Data

The data used in this paper come from a survey experiment where eight types of consumption survey modules were randomly assigned across 4000 households in Tanzania, about 500 households per module. Data were collected from September 2007 to August 2008. We analyze six modules as the two remaining modules lack data on food quantities, which are needed to calculate nutrient content. The experiment was fielded in several regions including Dodoma, Pwani, Dar es Salaam, Shinyanga and Kagera, and covers both urban and rural areas. Randomization was conducted within each of the 168 sampling units ('villages'), with each survey module assigned to three households per village. The randomization was successful according to the balance tests reported in Beegle et al (2012).

The survey modules differ by observation period, by length of the food consumption list, by type of survey reporter (individual or a single reporter per household) and by intensity

of interviewer visits. Three modules use the recall method, where food consumption is recalled by a single respondent on behalf of the household, in a single verbal interview. The recall modules differ in that they cover either the last 7 days or the last 14 days of consumption, as well as in the number of food and drink groups used to prompt the recall, based on either 58 groups or a subset of 17 important groups. The first recall module uses a long 58-item list with a 14 day recall. The second recall module is identical to the first one, except for the recall period which is reduced to 7 days. A third recall module is identical to the second one, except for the list of items, which is reduced to 17. All three recall modules included an identical module asking about 7 different categories of food and drink consumed outside of the house in the past 7 days. Note that the typical food consumption methods used by nutritionists i.e. the individual 24h recall (Bouis et al., 1992; Bouis, 1994), are not included in this study. HCES data already include food consumption and expenditures, which is why the modules used in these surveys are interesting to dissect.

The other three modules use the diary method, where respondents record daily transactions in open form diaries for a 14 day period. The diary modules are of an acquisition type, adding all products coming into the household through harvests, purchases, gifts and stock reductions as well as subtracting all items not consumed, such as through sales, stock increases and gifts. Food waste and food given to animals are taken into account and guest meals and meals consumed outside of the household are also recorded. The diary modules differ in two ways. First, the reporting is done by either each individual adult (who also reports on behalf of children and others for whom they are responsible and who are incapable of completing the diaries) or at the household level where there is a single diary-keeper for the whole household. Individual level recording entails that each adult member records products they or their charges have brought into

the house. These are then stored (so show up as a stock increase) or consumed by the household as a whole. Second, the frequency of interviewer visits differs. Frequent visits entails daily visits by the local assistant and visits every other day by the survey enumerator for the duration of the two-week diary. The infrequent visit treatment entails three visits: to deliver the diary (day one), to pick up the first week diary and drop off the week two diary (day eight), and to pick up the second week diary (day 15). There are three types of diaries: a frequent visit individual diary, a frequent visit household diary and an infrequent visit household diary.

During this survey experiment, investments were specifically made to ensure the personal diaries could provide a benchmark for analytic comparisons. Diary holders received daily visits from resident enumerators to help them fill in the diaries and to avoid double counting across multiple diaries. Such daily visits can be particularly important when collecting diary information from illiterate households. Every second day there would be an additional visit from the survey team.

There are at least two diagnostics that support our interpretation of the results from the highly supervised individual diaries as a benchmark. First, a common pattern with diary surveys is a decline in the number of items recorded over time, as respondents become fatigued. This decline, which shows up as lower reported consumption in week 2 than in week 1, is sometimes suggested as a reason to not use diary surveys as a benchmark (Brzozowski et al., 2017). However, for the experiment that we use in Tanzania, Beegle et al. (2012) compare the number of diary entries and the total consumption expenditures between the first and second week and find no differences, overall or by key household characteristics. This absence of diary fatigue attests to the care with which the diaries were administered.

We also supplement this analysis by using Benford's Law as a diagnostic. Specifically, we look at the distribution of the first digit of the consumption value reported in the diaries across days. If people were starting to fatigue as the diary went on, or were otherwise reporting fabricated numbers we would expect to see bigger deviations from Benford's law towards the end of the diary period. Figure A1 in the appendix shows that the general pattern predicted by Benford's Law holds and deviations from it (such as the higher frequency of 5 and lower frequencies of 7, 8 and 9) do not differ by day. In other words, there is no evidence of either diary-keeping respondents or the interviewers taking shortcuts and fabricating data as the survey progresses, which again attests to the care with which the diaries were administered. Based on these analyses and the use of the highly supervised individual diaries as a 'gold standard' benchmark from previous studies using this particular survey experiment, we also treat the evidence from this arm of the experiment as our benchmark for actual food consumption.

As the diary modules recorded detailed Swahili descriptions of all foods consumed, we exploited this information to disaggregate the reported transactions into 99 food groups. This lets us split up important food groups containing multiple types of vegetables or fruit, as well as specific snacks and full meals.

For both diaries and recall modules we are able to link the individually consumed products or meals to their macro- and micronutrient content using the Tanzanian Food Composition tables (Lukmanji et al., 2008). Using the quantities of food consumed, we calculate calorie consumption, as well as macronutrient consumption of protein, animal protein, fat, saturated fat, cholesterol, carbohydrates, fiber and sugars. Micronutrient consumption was also calculated for vitamins (vitamin A, B1, B2, B3, B6, B9, B12, C

and E) and minerals (calcium, phosphorus, iron, sodium, potassium, magnesium and zinc).³

In order to facilitate data interpretation, we present calorie⁴ and nutrient consumption relative to the Dietary Reference Intakes (DRI).⁵ The calorie calculations follow Smith et al. (2006) in determining the daily-required calorie consumption per household member taking into account age, sex and breastfeeding status. DRI values for macronutrient, vitamin and mineral consumption are derived from Otten et al. (2006) and are scaled by age, gender and breastfeeding status.

When determining the cost of calorie and nutrient sufficient diets in Section 3.3. we add the additional consideration that the DRI values of potentially unhealthy nutrients such as saturated fat, cholesterol, sugars and sodium should not exceed maximum values at which they start negatively impacting health (Otten et al., 2006). We also verify that consumption of vitamins and minerals does not exceed toxicity levels. In those calculations we include a minimum requirement for animal-based protein. While in dietary terms this is not strictly necessary, we use it as a shortcut way to ensure 9 key amino-acids in protein are covered in the calorie and nutrient sufficient diet.⁶

³ The recall modules followed standard practice in HCES, in only soliciting reports on expenditures within each of the 7 categories of food taken outside of the household. Dividing these expenditures by local prices surveyed during the experiment lets us derive approximate quantity information here.

⁴ In the remainder of the text, for ease of writing, we will refer to all dietary substances as nutrients.

⁵ Dietary reference intakes (DRI) include 4 subsets of recommendations. We use the recommended dietary allowance (RDA), determining the average daily dietary nutrient intake level that is sufficient to meet the nutrient requirements of nearly all (97–98 percent) healthy individuals in a particular life stage and gender group, and the tolerable upper intake level, the highest average daily nutrient intake level that likely poses no adverse health risks to almost all individuals in the general population (Otten et al., 2006).

⁶ Animal products such as meat, fish, poultry, eggs and dairy products contain nine essential amino acids necessary for the body's growth, maintenance and repair. These protein sources are called "complete proteins", as opposed to plant-sourced proteins found in nuts, seeds, legumes, grains and vegetables, which tend to be deficient in one or more of the essential amino acids (Otten et al., 2016). The alternative would be to include these amino-acids separately in the list of requirements, but we chose not to do this to make the presentation of the results more readable and tables more concise.

3. Results

3.1. Diary vs recall modules

We start by reporting average percentage deviations from the DRI, calculated over the ca. 500 households given each of the six survey modules. In the results in Table 1 and 2, positive values represent the percentage by which the requirement is exceeded and negative values are the percentage shortfall from the DRI.⁷ For example, for the sample using individual diaries, calorie consumption appears to be 20% above DRI, while animal protein consumption is 51% below DRI. The results for the sample using individual diaries in column (1) serve as a benchmark to compare with the results from the other survey approaches, given that the individual diary is the most resource intensive survey approach that should come closest to the truth. If the household diaries are used instead of the individual diaries, calories appear to be only 7% (for frequent visits) or 12% (for infrequent visits) above requirements, while animal protein appears to be 61% (for frequent visits) or 56% (for infrequent visits) below the DRI.

In column (4) of Table 1, we display the mean consumption levels across all three types of diaries. It is evident that the main dietary components below requirements are animal protein, fat, Vitamin B9 (folate), Vitamin E, Calcium and Zinc. We also observe that, on average, sugar and saturated fat consumption lie above DRI, but consumption of cholesterol and sodium remain below potentially harmful levels. A rise in the

⁷ We use the percentage relative to Dietary Reference Intakes (DRI) as this provides a comparable and understandable measure, as opposed to reporting our findings in grams, for example. This measure also provides more information and variation than a simple binary distinction of consuming sufficiently or not. We refer readers interested in further distributional data to Table A1, where we use the intensively supervised individual diaries to show, per nutrient, (i) prevalence of consumption under DRI, (ii) for those households with a shortfall, the average percentage of that shortfall, (iii) for those households with a shortfall, the squared average percentage of that shortfall.

consumption of animal protein and healthy fats would improve the diets of these households.

Table 1 | Nutrient consumption relative to DRI (% over or under DRI) – Diary modules

	(1)	(2)	(3)	(4)
	Individual diary, frequent visits, 14 days	Household diary, frequent visits, 14 days	Household diary, infrequent visits, 14 days	Average of all 3 diaries
Calories	20.28	6.69	11.89	12.96
Macronutrients				
Protein	44.74	25.43	34.83	35.01
Animal protein	-50.55	-61.14	-56.19	-55.96
Fat	-11.03	-24.70	-20.72	-18.81
Saturated fat	63.69	34.42	44.35	47.50
Cholesterol	-74.57	-82.46	-78.18	-78.41
Carbohydrates	50.72	37.39	44.62	44.25
Sugars	33.06	10.13	19.51	20.91
Fiber	66.28	49.33	56.33	57.32
Vitamins				
Vitamin A	34.45	13.13	23.85	23.81
Vitamin B1	25.18	13.56	20.14	19.63
Vitamin B2	164.41	135.83	165.33	155.19
Vitamin B3	40.13	23.54	29.59	31.09
Vitamin B6	75.39	52.99	61.48	63.29
Vitamin B9	2.27	-9.90	-1.87	-3.16
Vitamin B12	72.09	32.03	38.96	47.72
Vitamin C	38.44	17.65	27.71	27.94
Vitamin E	-8.99	1.65	-4.59	-3.98
Minerals				
Calcium	-44.48	-51.74	-47.77	-47.99
Phosphorus	103.76	80.89	91.68	92.12
Iron	34.82	20.58	27.50	27.64
Sodium	-52.83	-58.71	-57.95	-56.49
Potassium	19.36	6.62	12.35	12.78
Magnesium	48.00	35.59	44.20	42.60
Zinc	-1.27	-12.86	-5.39	-6.50

Note: DRI is Dietary Reference Intakes. Positive values denote % overconsumption relative to DRI, negative values % under-consumption relative to DRI.

Table 2 presents the results for the recall modules. According to the 14-day recall module, calories seem almost exactly at the recommended values, despite the more accurate individual diaries suggesting that calorie consumption was 20% greater than the requirement. The 7-day recall comes closest to the individual diaries with an overconsumption of 22%. The 7-day recall using a subset of food groups overestimates this value. Further, we find similar patterns across nutrients as in Table 1, but with

noticeably different magnitudes. For example, consumption of animal protein seems to be 62 - 69% below the requirement compared to 51% below in the benchmark results. It is the long food list 7-day recall that most closely mirrors the personal diary results, although it still overstates the amount by which dietary needs are not being met. The greatest discrepancy from the benchmark results comes from using the 14-day recall module.

Table 2 | Nutrient consumption relative to DRI (% over or under DRI) – Recall modules

	(1)	(2)	(3)	(4)
	14 day recall	7 day recall	7 day recall (subset of food groups)	Average of all 3 recall modules
Calories	3.36	22.12	31.98	19.15
Macronutrients				
Protein	6.17	25.26	23.73	18.38
Animal protein	-69.23	-62.86	-67.59	-66.56
Fat	-37.55	-24.92	-29.14	-30.54
Saturated fat	14.47	36.80	51.50	34.26
Cholesterol	-80.70	-77.65	-85.88	-81.41
Carbohydrates	39.70	65.11	84.78	63.20
Sugars	22.98	53.87	34.86	37.24
Fiber	56.94	87.32	91.51	78.59
Vitamins				
Vitamin A	14.36	23.92	81.98	40.08
Vitamin B1	11.85	30.01	47.21	29.69
Vitamin B2	90.61	124.88	191.00	135.49
Vitamin B3	12.05	32.06	26.81	23.64
Vitamin B6	67.10	99.90	113.63	93.54
Vitamin B9	15.65	37.57	44.64	32.62
Vitamin B12	-3.29	4.88	38.79	13.46
Vitamin C	45.13	71.73	89.36	68.74
Vitamin E	14.89	3.27	7.19	8.45
Minerals				
Calcium	-59.08	-50.60	-45.13	-51.60
Phosphorus	63.00	89.98	106.10	86.36
Iron	17.67	37.00	50.07	34.91
Sodium	-67.40	-60.90	-76.51	-68.27
Potassium	29.36	58.62	67.04	51.67
Magnesium	64.48	95.11	104.20	87.93
Zinc	-8.69	7.19	7.40	1.97

Note: DRI is Dietary Reference Intakes. Positive values denote % overconsumption relative to DRI, negative values % under-consumption relative to DRI.

In order to empirically test differences in apparent nutrient consumption by module, we use a regression analysis. Our ordinary least squares model has the following structure:

$$N_{ik} = \beta_k M_k + e_{ik} \quad (1)$$

where N_{ik} represents the percentage of nutrient consumed relative to DRI, by average household member i assessed using questionnaire k . M_k is a vector of dummy variables for each module type where the personal diaries become our reference category, and e_{ik} is the error term.

Table 3 shows how apparent macronutrient consumption relative to DRI varies by module, where the sample given the personal diary module are the reference group. If diaries are conducted at the household level, with one single diary recorder, they show lower levels of macronutrient consumption. This is apparent with both frequent and infrequent enumerator visits. Thus, it appears that relying on a single recorder rather than on individual recording is the design dimension that seems to affect the data.

The understatement of macronutrient consumption is even more apparent when a 14-day recall (again with a single reporter per household) is used. If a 7-day recall is used, with either type of recall list (58 food and drink groups or 17 major items) there is consumption overstatement of carbohydrates, sugar, fiber and (for the subset list) calories, with protein and cholesterol understated. These patterns may reflect a household respondent remembering a higher consumption of starchy foods like maize, rice, bread and cakes than is truly the case. Another reason for overstatement may be that because the recall modules were unbounded, with no initial visit to mark the start of the recall period, telescoping errors could cause respondents to report consumption occurring before the recall period as if it happened during the recall period.

Table 4 shows how apparent consumption of vitamins vary by survey module. For frequently visited household diary and 14-day recall, the discrepancies from the benchmark are often around 20% or larger. These are lower and less prominent for the

infrequently visited household diaries. Apparent consumption using 7-day modules are more diverse, with vitamin B2, B3 and B12 understated, but vitamin B6, B9 (folate), C and E overstated. If the subset food list is used, all vitamin consumption, bar vitamin B3 and B12, appears to be overstated.

Table 5 presents similar findings for mineral consumption. For the sample given the benchmark diaries, calcium consumption averaged 44% below requirement, as seen from the constant term in Table 5 (and from Table 1). Yet, consumption would seem even lower, by a further factor of up to 14% points (for 14-day recall) if the other survey modules were used. The pattern of understated consumption holds for most minerals, except potassium and magnesium, and is especially apparent if household diaries or the 14-day recall are used. In contrast, the 7-day recall using the full food list provides slight understatements in the consumption of calcium, phosphorus and sodium, but an overstatement in potassium, magnesium and zinc consumption. The 7-day recall using the subset list of foods is more likely to overstate mineral consumption. It must be noted that minerals coming from foods that are difficult to record i.e. sodium from salt, are more likely to be understated across the board.⁸

⁸ The results presented in Table 3, 4 and 5 determine the effect that module types have on nutrient consumption on a one-by-one basis. However, in total we assess these effects on 25 outcomes, therefore raising the possibility of multiple hypothesis testing errors. Using the Holm (1979) sequentially-rejective Bonferroni method, we reconsider the significance of individual coefficients when viewed as a family of n hypothesis, ranking the n outcomes in increasing order of their p-values i.e. $p_1 \leq p_2 \leq \dots \leq p_n$, we test whether $p_1 \leq \alpha/n$ to reject the null effect of the outcome with the smallest p-value. If this is the case, the null effect of all following outcomes can also be rejected. Therefore, the smallest p-value in our analysis must be significant at for example the 0.04% level instead of that at the nominal 1% level, or at the 0.2% level instead of at the nominal 5% level. We confirm that this is the case for our results as we have multiple coefficients with a p-value of 0.000. Thus, our findings remain robust to potential family-wise errors.

Table 3 | Effect of module type on apparent macronutrient consumption as a percentage of dietary reference intakes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Calories	Protein	Animal protein	Fat	Saturated fat	Cholesterol	Carbohydrates	Sugars	Fiber
Diary: HH, frequent	-13.60*** (5.125)	-19.31*** (4.462)	-10.59*** (2.604)	-13.67 (14.06)	-29.27 (40.51)	-7.890*** (1.652)	-13.33*** (4.675)	-22.93*** (6.427)	-16.95*** (5.621)
Diary: HH, infrequent	-8.389 (5.127)	-9.912** (4.464)	-5.633** (2.605)	-9.689 (14.06)	-19.35 (40.53)	-3.610** (1.653)	-6.102 (4.677)	-13.55** (6.430)	-9.958* (5.624)
Recall: Long, 14 day	-16.93*** (5.120)	-38.58*** (4.457)	-18.68*** (2.601)	-26.52* (14.04)	-49.22 (40.47)	-6.129*** (1.650)	-11.02** (4.670)	-10.07 (6.420)	-9.342* (5.615)
Recall: Long, 7 day	1.835 (5.120)	-19.48*** (4.457)	-12.31*** (2.601)	-13.89 (14.04)	-26.89 (40.47)	-3.072* (1.650)	14.39*** (4.670)	20.81*** (6.420)	21.04*** (5.615)
Recall: Subset, 7 day	11.70** (5.120)	-21.02*** (4.457)	-17.04*** (2.601)	-18.11 (14.04)	-12.19 (40.47)	-11.31*** (1.650)	34.06*** (4.670)	1.801 (6.420)	25.22*** (5.615)
Constant	20.28*** (3.622)	44.74*** (3.153)	-50.55*** (1.840)	-11.03 (9.934)	63.69** (28.63)	-74.57*** (1.167)	50.72*** (3.304)	33.06*** (4.542)	66.28*** (3.973)
Observations	3,018	3,018	3,018	3,018	3,018	3,018	3,018	3,018	3,018
R-squared	0.014	0.027	0.024	0.001	0.001	0.019	0.048	0.018	0.031

Note: Each column presents estimates of equation (1), where the left hand side variable is the deviation from the dietary reference intake of the nutrient indicated in the column heading. Standard errors are in parentheses under the coefficient. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4 | Effect of module type on apparent vitamin consumption as a percentage of dietary reference intakes

VARIABLES	(1) Vitamin A	(2) Vitamin B1	(3) Vitamin B2	(4) Vitamin B3	(5) Vitamin B6	(6) Vitamin B9	(7) Vitamin B12	(8) Vitamin C	(9) Vitamin E
Diary: HH, frequent	-21.32* (11.24)	-11.63*** (3.940)	-28.58** (12.04)	-16.59*** (4.416)	-22.40*** (6.581)	-12.17** (4.922)	-40.06*** (14.14)	-20.79** (8.490)	10.65 (6.984)
Diary: HH, infrequent	-10.60 (11.25)	-5.041 (3.942)	0.922 (12.05)	-10.54** (4.418)	-13.90** (6.585)	-4.136 (4.925)	-33.13** (14.15)	-10.73 (8.494)	4.408 (6.988)
Recall: Long, 14 day	-20.09* (11.23)	-13.33*** (3.937)	-73.80*** (12.03)	-28.07*** (4.411)	-8.287 (6.575)	13.39*** (4.918)	-75.38*** (14.13)	6.689 (8.482)	23.89*** (6.977)
Recall: Long, 7 day	-10.53 (11.23)	4.827 (3.937)	-39.54*** (12.03)	-8.062* (4.411)	24.52*** (6.575)	35.30*** (4.918)	-67.21*** (14.13)	33.29*** (8.482)	12.26* (6.977)
Recall: Subset, 7 day	47.53*** (11.23)	22.03*** (3.937)	26.57** (12.03)	-13.32*** (4.411)	38.25*** (6.575)	42.37*** (4.918)	-33.30** (14.13)	50.92*** (8.482)	16.18** (6.977)
Constant	34.45*** (7.944)	25.18*** (2.785)	164.4*** (8.513)	40.13*** (3.121)	75.39*** (4.651)	2.269 (3.479)	72.09*** (9.995)	38.44*** (6.000)	-8.996* (4.936)
Observations	3,018	3,018	3,018	3,018	3,018	3,018	3,018	3,018	3,018
R-squared	0.017	0.035	0.028	0.015	0.041	0.063	0.012	0.033	0.005

Note: Each column presents estimates of equation (1), where the left hand side variable is the deviation from the dietary reference intake of the nutrient indicated in the column heading. Standard errors are in parentheses under the coefficient. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5 Effect of module type on apparent mineral consumption as a percentage of dietary reference intakes

VARIABLES	(1) Calcium	(2) Phosphorus	(3) Iron	(4) Sodium	(5) Potassium	(6) Magnesium	(7) Zinc
Diary: HH, frequent	-7.259*** (2.317)	-22.87*** (6.044)	-14.24*** (4.269)	-5.874** (2.389)	-12.74** (6.032)	-12.41** (5.379)	-11.59*** (3.258)
Diary: HH, infrequent	-3.294 (2.318)	-12.08** (6.047)	-7.322* (4.271)	-5.118** (2.390)	-7.007 (6.035)	-3.808 (5.382)	-4.122 (3.260)
Recall: Long, 14 day	-14.60*** (2.315)	-40.76*** (6.038)	-17.15*** (4.265)	-14.56*** (2.387)	10.00* (6.026)	16.48*** (5.374)	-7.424** (3.255)
Recall: Long, 7 day	-6.125*** (2.315)	-13.78** (6.038)	2.183 (4.265)	-8.067*** (2.387)	39.26*** (6.026)	47.11*** (5.374)	8.460*** (3.255)
Recall: Subset, 7 day	-0.654 (2.315)	2.342 (6.038)	15.24*** (4.265)	-23.68*** (2.387)	47.68*** (6.026)	56.20*** (5.374)	8.667*** (3.255)
Constant	-44.48*** (1.638)	103.8*** (4.272)	34.82*** (3.017)	-52.83*** (1.688)	19.36*** (4.263)	48.00*** (3.802)	-1.267 (2.303)
Observations	3,018	3,018	3,018	3,018	3,018	3,018	3,018
R-squared	0.016	0.022	0.024	0.039	0.055	0.086	0.021

Note: Each column presents estimates of equation (1), where the left hand side variable is the deviation from the dietary reference intake of the nutrient indicated in the column heading. Standard errors are in parentheses under the coefficient. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.2. Heterogeneity analysis

Various household characteristics may affect module performance. For example, a more educated respondent may provide more accurate reports of food consumed over the past period and this may matter more as the period lengthens compared to a daily diary entry (especially as interviewers and survey assistants checked daily). On the other hand, households in urban areas or with more wealth likely consume a greater variety of food, adding to the complexity of the reporting task, especially for recall modules. Lastly, household size and how many young children are present in the household could affect the focus of the respondent and the time spent on diary recording. In other words, the potential accuracy of some of the designs depends more heavily on respondent characteristics than is the case for other designs. Therefore, we run similar regressions to those in Section 3.1., but now including several selected household characteristics and an interaction term between the household characteristic and module type, as follows:

$$N_{ik} = \beta_k M_k + \beta_x X_{ik} + \delta_k M_k X_{ik} + e_{ik} \quad (2)$$

where X_{ik} is a single household characteristic and δ_k represents the coefficient for the interaction term. This equation is estimated individually, per household characteristic. The household characteristics we use are household size, education of the household head, household wealth, the share of children under 6 and whether the household lives in an urban area.

Table 6 presents a summary of the results of the heterogeneity analysis.⁹ Panel A of this table shows which modules have positive (non-bold) or negative (bold) interaction effects

⁹ Full regression tables can be found in Table A2-A6 in the appendix. Table A7 in the appendix repeats Panel B of Table 6 considering all characteristics jointly in the same regression.

for the household characteristic indicated in the column when assessing the consumption of the nutrient indicated in the row. Panel B of Table 6 summarises these results by simply counting per characteristic and per module how many statistically significant (at the 5% level) interaction effects there are. When interpreting Panel B it is useful to keep in mind that Panel A assesses the modules for 25 nutrients across 5 household characteristics and so is assessed 125 times. From Panel B we can see that the characteristics that appear to matter most to the relative performance of survey modules are household size and urban location, with 32 and 36 significant interaction effects. Household wealth, education of the household head and the share of children under 6 have far fewer significant interaction effects (16, 13 and 11, respectively, out of a possible total of 125).

For household size, it appears that respondents in larger households report relatively less when using recall modules (of either 7 or 14 day duration). The effect holds for all minerals, for all vitamins except Vitamin B2, B12 and E, and for calories and protein. In other words, the gap between the consumption measured when using benchmark individual diaries and what is measured using recall methods is bigger for larger households. This pattern is consistent with prior evidence of recall surveys increasingly understating food consumption as household size grows, because larger households have more people to report on and so in a given time interval will have more food transactions to recall (Gibson & Kim, 2007). In contrast, almost no nutrient consumption is differentially affected by household size when various types of diaries are used.

Whether the household resides in a rural or urban area also seems to matter. Recall modules often seem to mistakenly understate the macro- and micronutrient consumption of urban households compared to what the benchmark shows. Recall likely misses more complex urban consumption patterns that are less reliant on common-pot eating; thus, a

recall respondent in an urban area may be less informed about the food consumption of the other household members than is the case in a rural area. As can be seen from column (3) this is not a pure wealth effect, as wealth has half the number of significant interaction effects.

Finally, the share of children under 6 years of age and the years of formal education of the household head have the lowest count of significant interaction effects and the effects are overwhelmingly concentrated in the recall modules.

Table 6 | Significant interactions between module type and selected household characteristics

Panel A: Significant effects

	Household characteristics				
	(1)	(2)	(3)	(4)	(5)
	Household size	Education household head	Wealth	Share children under 6	Urban household
Calories	4	-	5	-	3, 4
Protein	4	1	3, 4	4	-
Animal protein	-	-	1, 3, 4, 5	4	1, 3, 4, 5
Fat	5	-	5	-	-
Saturated fat	5	-	5	-	-
Cholesterol	-	1, 5	1, 2	-	1, 2
Carbohydrates	3, 4, 5	4, 5	5	-	3, 4, 5
Sugars	1, 4, 5	-	5	-	-
Fiber	3, 4, 5	4, 5	5	-	3, 5
Vitamin A	5	-	-	-	5
Vitamin B1	3, 4, 5	4, 5	5	-	3, 4, 5
Vitamin B2	-	-	2	-	-
Vitamin B3	4	-	-	-	3, 5
Vitamin B6	3, 4, 5	-	-	4, 5	3, 4, 5
Vitamin B9	3, 4, 5	-	-	4, 5	5
Vitamin B12	-	-	1	-	1, 2
Vitamin C	4	-	5	4, 5	5
Vitamin E	4	-	-	-	5
Calcium	3, 4, 5	-	3, 4, 5	3, 4, 5	3, 4
Phosphorus	3, 4, 5	-	-	4	5
Iron	3, 4, 5	-	-	-	3, 4, 5
Sodium	4	3, 4	3, 4	3, 4	3, 4
Potassium	3, 4, 5	-	5	5	3, 4, 5
Magnesium	4, 5	3, 4, 5	3, 4, 5	-	3, 4, 5
Zinc	4	3	-	4	-

Panel B: Count of significant effects in Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
	Household size	Education household head	Wealth	Share children under 6	Urban household	TOTAL
1. Diary: HH, frequent	1	1	2	0	3	7
2. Diary: HH, infrequent	0	0	1	0	2	3
3. Recall: Long, 14 day	9	2	5	2	12	30
4. Recall: Long, 7 day	11	5	5	5	10	36
5. Recall: Subset, 7 day	11	5	3	4	9	32
TOTAL	32	13	16	11	36	108

Note: Panel A shows the results of estimating equation (2) separately for each combination of nutrients indicated in the row and the household characteristic indicated in the column. The cells indicate for which module the interaction effects are significant at the 5% level, with a bold number representing a negative coefficient and a non-bold number a positive one. The modules are represented with the following numbers 1 (Diary HH, frequent), 2 (Diary HH, infrequent), 3 (Recall 14 days), 4 (Recall 7 days), 5 (Recall 7 days subset). A dash (-) means there are no positive effect for any module for that combination of nutrient and household characteristic. Panel B then counts the number of significant effects in Panel A (positive or negative) per module and per household characteristic. Table A7 repeats Panel B, but for regressions including all characteristics jointly.

Looking at the last column in Panel B, which summarises the total number of significant interaction effects across all 5 modules, we see that the household diaries have much lower counts than the recall modules. All recall modules have many significant interaction effects. Especially the 7-day recall stands out, with 36 significant effects out of a total of 125 possible effects. These patterns confirm that the potential accuracy of some module designs, particularly recall modules, are more subject to variation according to respondent characteristics than is the case for other designs.¹⁰

3.3. Food poverty

Where analysts lack data on food quantities consumed or surveys do not have the required granularity in food items to accurately determine macro- and micronutrient consumption,

¹⁰ Note that the 5% level is chosen to assess the significance for individual outcomes, however, if we were to assess significance in a multiple hypothesis testing context, we would have to determine which module types interact significantly with household characteristics at the 0.2% level. In doing so, we find that most interactions become insignificant, except those for the 7-day recall modules. It remains the case, therefore, that the 7-day recall is more sensitive to household characteristics compared to other survey modules when taking possible family-wise errors into account.

there may be interest in comparing the value of total food consumption expenditures against a minimum threshold. These comparisons may also be motivated by the insight that interventions to improve nutrition ultimately have to be budgeted for, and so monetary-based calculations can be useful because they are in the same metric (dollars or TZ shillings in this case) as those used when nutritional intervention programs are funded. We conduct such an exercise with two thresholds.

In the first we use linear programming to determine, per region, the minimal cost of a diet that meets all caloric and nutrient requirements per adult equivalent capita (Allen, 2017).¹¹ We call the diet in question the MC diet and the corresponding cost the MC diet poverty line. We then compute the share of households whose food expenditures lie below the MC diet poverty line. This provides a minimum bound on the calculation of nutrient deprived households: we can be sure that these households are facing nutrient deficiencies, although we cannot be sure that those whose value of food consumed is above the threshold are eating healthy diets. For those with a food consumption shortfall, we also calculate the average gap between the value of the household's food consumption and the price of the MC diet in their region.

The second threshold calculates the cost of the typical diet of the 40% poorest, in terms of total income, to ensure it provides sufficient calories for the household. If the typical diet of the 40% poorest in a region provides, say, 1800 calories, then the cost of this diet is scaled up by for example 2200/1800 to meet the target daily recommended calorie

¹¹ We are calculating the minimum cost of nutrient adequacy. Nutrient adequacy is more strict than calorie adequacy, but less strict than, for example, the 'healthy diet' used in FAO et al. (2020), which is based on national food-based dietary guidelines, or the diet proposed by the EAT-Lancet commission which includes further considerations related to, for example, sustainability. Hirvonen et al. (2020) show how the EAT-lancet diet is, on average, 60% more expensive than the minimum cost of nutrient adequacy. The FAO et al. (2020) healthy diet is also 60% more expensive than the minimum cost nutrient adequate diet. For similar exercises see also Drewnowski and Darmon (2005), Chastre et al. (2007), Masters et al. (2018), Headey and Alderman (2019).

intake.¹² We call this the food poverty line (as the scaled basket is one component of cost-of-basic-needs poverty lines). We also calculate the share of households whose food expenditures lie below it and, conditional on that, the average gap.

Both poverty lines are calculated using the highly supervised individual diaries, which we believe contain the most accurate information on prices and expenditures. For the MC diet, prices are median region-level unit prices for each of the 99 food items recorded in the diaries. Ideally, we would have used a price survey for this exercise, but the one conducted as part of this survey experiment was not done with sufficient granularity for all food items to confidently calculate the least cost diet.

Table 7 | Minimum cost diet and cost of basic needs food poverty lines per region

	Minimum cost diet food poverty line (TZS)	Cost of basic needs food poverty line (TZS)
Dodoma	320.17	710.40
Pwani	227.84	746.08
Dar es Salaam	411.76	1444.25
Shinyanga	310.42	617.94
Kagera	257.75	695.46

Table 7 shows the two poverty lines, broken down by region.¹³ A number of interesting facts emerge. First, the food poverty line is much higher than the MC diet poverty line in all regions. The diets of the poorest 40% are 2-3 times more expensive than what would be strictly needed to meet all nutrient requirements. Second, there is much more regional

¹² The average individual expenditure (per adult equivalent) on food by the 40% poorest is scaled by the average calorie deficit or overconsumption of the average 40% poorest, in order to ascertain the food expenditure needed to reach the daily calorie requirements.

¹³ Masters et al. (2018) also calculates the minimum cost diet in Tanzania and includes the necessary detail for a comparison with our own results. In order to do so, we run our LP algorithm on the 8 products included in Masters et al. (2018) and account for inflation. We see that once we restrict diets to these 8 products instead of the 99 items included in our analysis, the price of our MC diet increases substantially, in line with the results found by Masters and co-authors. We believe this is due to the fact that we are optimizing over a much larger set of items, allowing us to find much cheaper sources of nutrients. Using the 8-item optimization, the gap between the CBN and MC lines closes everywhere, but it remains, like before, relatively larger in Pwani and Dar es Salaam. We hypothesise that this is due to the specific food preferences in these places.

variation in the food poverty line than in the MC diet poverty line. Especially Dar es Salaam has a higher poverty line than other regions, which is less evident from Dar es Salaam's MC diet poverty line. Therefore, the higher food poverty line is not just due to higher prices in Dar es Salaam, but also reflects the choice of those consumers for more expensive sources of macro- and micronutrients.

Another interesting example is Pwani, which has the cheapest MC diet at TZS 228 per person per day, yet the food poverty line in Pwani is the highest after Dar es Salaam in our sample, at TZS 746 per person per day. Thus, it appears that consumers in Pwani are choosing relatively higher priced micro- and macronutrients. In this specific case, this could perhaps be due to a local culture of eating outside the household.

Table 8 | Percentage of the poorest 40% households consuming below the DRI

	Percentage of poorest 40% households consuming below DRI (%)
Calories	43.70
Macronutrients	
Protein	34.00
Animal protein	87.30
Fat	79.70
Saturated fat	65.00
Cholesterol	98.50
Carbohydrates	17.30
Sugars	50.30
Fiber	15.20
Vitamins	
Vitamin A	62.40
Vitamin B1	37.60
Vitamin B2	17.80
Vitamin B3	35.50
Vitamin B6	7.61
Vitamin B9	57.90
Vitamin B12	71.10
Vitamin C	29.40
Vitamin E	55.30
Minerals	
Calcium	84.80
Phosphorus	16.80
Iron	40.10
Sodium	94.40
Potassium	23.90

Magnesium	20.30
Zinc	64.50

Note: Figures are for the poorest 40% of households in the benchmark highly supervised individual diary group. 40% poorest is determined according to the asset index, a proxy for overall income, not simply food poverty. $N=1,206$.

Finding that the food poverty line is much higher than that of the MC diet poverty line, may insinuate that the food consumption of Tanzanian households suffices. However, if we calculate the percentage of households within the 40% poorest income group that consume under the dietary reference intakes, we find that this is not the case (Table 8). Although spending more than what is necessary to fulfil all nutritional needs, about 44% of these households consume too little in terms of calories, 88% too little in terms of animal protein and 80% too little in terms of fat. Vitamin and mineral deficiencies are also prevalent for more than half of this subsample. These findings echo other data from Tanzania on child stunting levels, at about 42% at the time of this survey experiment, and still remaining high at about 32% in 2018 (World Bank, 2018).

Let us now look at how the different survey methods would assess the prevalence and depth of food poverty compared to these lines. We do so by comparing total household food expenditures, as measured by each of the modules, to the region-specific poverty lines listed in Table 7. Table 9 does this for the set of MC diet poverty lines, with Panel A showing the percentage of households whose food expenditures fall below the threshold and Panel B the percentage of the shortfall per person per day, i.e. the average proportionate gap for the food-poor.¹⁴

Table 9 portrays a situation that is consistent with the results for understated nutrient consumption. The 14-day recall overestimates the prevalence of extreme food poverty, suggesting that 11% of people in Tanzania live in households where total food

¹⁴ Table A8 contains similar results for the CBN food poverty line. The findings are similar to those in Table 9, but with higher rates of households with expenditures below the threshold and with larger shortfalls.

expenditure is below what is needed to buy the MC diet. This is more than twice the rate of extreme food poverty shown with the benchmark individual diary. Likewise, with the 14-day recall, the poverty gap of the food poor is overstated by about one third. In contrast, the infrequent household diary and 7-day recall modules perform most closely to the individual diaries at the national level, although with several discrepancies at the regional level.

Another instructive exercise is to sum the value of all food consumption shortfalls in our sample. A hypothetical perfectly targeted transfer of this size could give all households the freedom to eat a healthy diet without compromising on current non-food expenditures. Of course there is no guarantee that households would spend this amount on nutritious food, nor that the transfer would not spillover into increasing non-food expenditures, but it is a useful benchmark to keep in mind how much households are short when it comes to purchasing a healthy diet. According to the personal diaries, our hypothetical transfer programme would need to provide each MC diet food poor household 60 TZS per person per day to lift it up to the MC diet line. Extrapolating the 4.73% of the total sample under this line to the 57 million people in the whole of Tanzania implies a cost of just under TZS 162 million in total per day or about USD 51 million a year.¹⁵ However, if we were to budget based on the guidance provided by a 14-day recall survey, the estimated cost would be over 100 million USD higher, at USD 159m per year, which is a significant bias in the forecast costs. Carrying on with these calculations, the 7-day recall gives a much more reasonable estimate, putting the cost of perfectly targeted transfers at USD 69m.¹⁶

¹⁵ The average exchange rate during the study period was TZS 1,150 for one USD.

¹⁶ These values are much higher if calculated based on the CBN food poverty line and food poverty shortfall at USD 2.119 billion for personal diaries, USD 3.418 billion for the 14-day recall and USD 2.293 billion for the full food list 7-day recall.

Table 9 | Extreme food poverty

Panel A: Households spending below minimum cost diet threshold (%)						
	Dodoma	Pwani	Dar es Salaam	Shinyanga	Kagera	Total
Diary: Personal	5.65	9.64	0.00	6.77	1.25	4.73
Diary: HH, frequent	7.37	1.61	1.69	14.47	4.55	7.61
Diary: HH, infrequent	6.85	2.31	3.83	11.28	2.09	6.19
Recall: Long, 14 day	4.79	3.10	0.00	22.70	6.90	11.35
Recall: Long, 7 day	1.96	3.35	0.00	16.39	4.89	7.77
Recall: Subset, 7 day	8.65	0.00	0.93	22.09	2.26	9.54
Panel B: Average poverty gap of the food-poor (%)						
	Dodoma	Pwani	Dar es Salaam	Shinyanga	Kagera	Total
Diary: Personal	17.43	14.20	0.00	24.70	5.38	19.63
Diary: HH, frequent	29.17	41.59	17.65	19.69	17.06	20.91
Diary: HH, infrequent	24.41	4.51	5.74	17.32	25.67	17.50
Recall: Long, 14 day	17.76	11.60	0.00	27.29	22.23	25.39
Recall: Long, 7 day	2.76	2.30	0.00	19.11	9.96	16.04
Recall: Subset, 7 day	8.20	0.00	14.25	13.86	31.36	14.32

4. Discussion and Conclusions

Using a unique survey experiment, we have been able to assess survey performance in recording macro- and micronutrient consumption, as well as food poverty rates, across 3000 households in Tanzania. Our analyses are motivated by two factors: the growing attention to the nutrition transition means that the focus of many interventions is moving beyond calories; and, some of the most available sources of information are multi-topic household surveys. Yet, even while the confluence of these two factors may see more nutritional analyses based on data from multi-topic household surveys, a threat to the validity of these analyses is that the food consumption modules of these surveys differ in several key dimensions. While these issues have been studied before, especially in the collection of articles introduced by Zezza et al (2017), there remains much that is unknown, especially when the focus moves beyond calories.

One of the key results to emerge from the current experiment is that the 14-day recall modules led to the largest understatements of both macro- and micronutrient consumption. The long food list 7-day recall module is found to perform closest to the personalized

diaries, although particular attention to cereal and other carbohydrate consumption must be considered when using this module design. Household diaries also prove to be less accurate than highly supervised individual diaries, with it being the reliance on a single reporter rather than the frequency of enumerator visits that seems to matter. Comparing survey modules in terms of their performance at the nutrient level as well as at the calorie level has also proven important, with greater sensitivity of apparent protein, cholesterol and fibre consumption than is shown for calorie consumption. Apparent vitamin and mineral consumption also varies substantially with choice of survey module, which has important implications for surveys that aim to take dietary quality measures into account. Previous findings that focus solely on calories may therefore understate the lack of robustness of nutritional indicators to the variation in survey methods.

Furthermore, we found that recall modules are more sensitive to household characteristics than diary modules are. For example, larger household size and urban location are cause for recall modules to further underestimate actual consumption levels. It seems that having more members to recall for, a more diverse availability of foods, and less reliance on common-pot eating in urban areas can make recall modules relatively more difficult for respondents than they are in rural settings. Education levels of the household head and wealth had slightly less effect on the results. Some of the factors that underlie these patterns will occur more universally, but there is one important pattern that we find may not translate so easily to other settings.

In our experiment the diary modules generally perform better than the recall modules overall. First, the intensively monitored individual diary is judged as sufficiently close to the truth that we can use it as a benchmark. Second, the other diary modules have results

that are less affected by respondent characteristics than is the case for the recall modules, and give results that are fairly close to the benchmark.

We caution that this good performance of the diary modules may not translate into other settings. We had great cooperation from respondents, and interviewers were skilled and experienced, and more intensively trained and supervised than typical for developing country surveys. Thus, as mentioned in Section 2, common diagnostics of problems with diary surveys, such as incomplete diaries or fatigue as the diary-keeping period progresses, are not apparent in this experiment. In contrast, in other experiments, and under more usual field conditions with statistics offices, diary surveys often have many partially completed or incomplete records and rates of compliance are likely to vary with household characteristics so the missing information is not missing-at-random (Sharp et al, 2019). For example, Beegle et al. (2012) found that leaving an unsupervised diary with a household lacking a literate member introduces mismeasurement. This can be overcome with daily visits, which improve the quality of the data, but come at a high cost. Such costs are obviously context-specific, but to give an order of magnitude, implementing an unsupervised household diary is 2 or 3 times more expensive than a recall module, while a daily supervised individual diary is 9 times more expensive. Clearly intensively supervised individual diaries cannot be rolled out at any large scale, but they can be useful as a benchmark in studies addressing measurement.

Policy Implications

Accurate measurement of macro- and micronutrient consumption is needed for policymakers to be informed about the scale of nutritional problems and to be equipped with tools that can help to monitor the impact of interventions that aim to overcome these problems. A key finding from our study is that the disparities in apparent nutrient

consumption between survey modules are also observed when we compare the value of food consumption with food poverty lines. In particular, all module designs except for the 7-day recall would overestimate the prevalence and depth of food poverty in Tanzania. Consequently, any calculation of possible cash or in-kind transfer programs to eliminate food poverty would likely lead agencies into making excessive budgetary requests if the evidence base was built with surveys such as these. These excessive budgetary requests could affect the political feasibility of attempting to eliminate hunger with targeted transfers.

Although each survey design comes with potential benefits and shortcomings, the results here could be taken as support for the FAO/WB recommendations for food consumption surveys to use a 7-day recall, rather than longer period recall or diaries (Zezza et al., 2017). In particular, we find that 7-day recalls are relatively good at predicting average nutrient consumption, and also provide good predictions for the prevalence and depth of food poverty. Thus, such surveys could be a reasonable basis for estimating the needed scale of transfer programmes designed to alleviate food poverty.

Suggestions for Future Research

Notwithstanding the point above that these results could be taken as support for the FAO/WB recommendation to use 7-day recall surveys for food, there is still much that we do not know. In particular, we caution that even though the nutrient consumption values recorded through the 7-day recall are more likely than other modules to be correct on average, this design is rather sensitive to household characteristics. One reason for this, explored in some detail in Friedman et al. (2017), is that the 7-day recall has offsetting errors that seem to roughly balance: there are incidence errors within several important food groups (that is, entirely forgetting to report any consumption) that are offset by the

overstatement of the value of what was consumed conditional on reporting any consumption. One explanation for this pattern is telescoping errors, where consumption that occurred before the recall period is reported as if it was in the recall period; such errors will have a larger impact on a 7-day recall than on a 14-day recall (as they are amortized over fewer days). Hence, a recommendation for any future experiments on this topic is to include treatments that can inform about telescoping, such as the use of an initial visit to bound the recall period.

References

- Allen, R. (2017). Absolute Poverty: When Necessity Displaces Desire. *American Economic Review*, 107(12), 3690-3721.
- Beegle, K., De Weerd, J., Friedman, J., & Gibson, J. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, 98(1), 3-18.
- Black, R., Victora, C., Walker, S., Bhutta, Z., et al., (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 382(9890):427–51.
- Bouis, H., Haddad, L., & Kennedy, E. (1992). Does it matter how we survey demand for food?: Evidence from Kenya and the Philippines. *Food Policy*, 17(5), 349-360.
- Bouis, H. E. (1994). The effect of income on demand for food in poor countries: Are our food consumption databases giving us reliable estimates? *Journal of Development Economics*, 44(1), 199-226.
- Brzozowski, M., Crossley, T. F., & Winter, J. K. (2017). A comparison of recall and diary food expenditure data. *Food Policy*, 72(1), 53-61.
- Carletto, C., Zezza, A., and Banerjee, R. (2013). Towards better measurement of household food security: harmonizing indicators and the role of household surveys. *Global Food Security*, 2, 30-40.
- Chastre, C., Duffield, A., Kindness, H., LeJeune, S., Taylor, A. 2007. The minimum cost of a healthy diet: Findings from piloting a new methodology in four study locations. Save the Children. London: Save the Children, UK.

- De Haen, H., Klasen, S., & Qaim, M. (2011). What do we really know? Metrics for food insecurity and undernutrition. *Food Policy*, 36(6), 760-769.
- De Weerd, J., Beegle, K., Friedman, J., & Gibson, J. (2016). The challenge of measuring hunger through survey. *Economic Development and Cultural Change*, 64(4), 727-758.
- Dillon, B., De Weerd, J., & O'Donoghue, T. Paying More for Less: Why Don't Households in Tanzania Take Advantage of Bulk Discounts? *The World Bank Economic Review*. <https://doi.org/10.1093/wber/lhz020>
- Drewnowski, A. and Darmon, N., 2005. The economics of obesity: dietary energy density and energy cost. *American Journal of Clinical Nutrition*, 82(1), pp.265S-273S.
- FAO, & the World Bank. (2018). Food data collection in Household Consumption and Expenditure Surveys. Guidelines for low- and middle-income countries. Rome. 104 pp. Licence: CC BY-NC-SA 3.0 IGO
- FAO, IFAD, UNICEF, WFP and WHO. 2020. The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets. Rome, FAO.
- Finaret, A. & Masters, W. (2019). Beyond calories: the new economics of nutrition. *Annual Review of Resource Economics*, 11, 237–259.
- Fiedler, J., Smits, M., Dupriez, O., & Friedman, F. (2008). Household income and expenditure surveys: a tool for accelerating the development of evidence-based fortification programs. *Food and Nutrition Bulletin*, 29(4), 306-319.

- Fongar, A., Gödecke, T., Aseta, A., & Qaim, M. (2019). How well do different dietary and nutrition assessment tools match? Insights from rural Kenya. *Public Health Nutrition*, 22(3), 391-403.
- Forouzanfar H, Alexander L, Anderson H, et al., (2013). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*, 386, 2287–323.
- Friedman, J., Beegle, K., De Weerd, J., & Gibson, J. (2017). Decomposing response error in food consumption measurement: Implications for survey design from a randomized survey experiment in Tanzania. *Food Policy*, 72(1), 94-111.
- Gazeaud, J. (2020). Proxy Means Testing vulnerability to measurement errors? *The Journal of Development Studies*, 56(11) 2113-2133.
- Gibson, J., Beegle, K., De Weerd, J., & Friedman, J. (2015) What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxford Bulletin of Economics and Statistics*, 77(3), 466-474.
- Gibson, J., & Kim, B. (2007). Measurement error in recall surveys and the relationship between household size and food demand. *American Journal of Agricultural Economics*, 89(2), 473-489.
- Headey, D. and Alderman, H. 2019. The Relative Caloric Prices of Healthy and Unhealthy Foods Differ Systematically across Income Levels and Continents. *Journal of Nutrition*, 149(11), pp.2020-2033.
- Hirvonen, K., Bai, Y., Headey, D., Masters, W. (2019). Affordability of the EAT-Lancet reference diet: a global analysis. *Lancet Global Health* 8: e59–66.

- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65-70.
- Imamura, F., Micha, R., Khatibzadeh, S., Fahimi, S., Shi, P., Powles, J., et al. & Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (NutriCoDE). (2015). Dietary quality among men and women in 187 countries in 1990 and 2010: a systematic assessment. *Lancet Global Health*, 3(3), e132-e142.
- Jann, B. (2007). DIGDIS: Stata module to analyze the distribution of digits, *Statistical Software Components* S456853, Boston College Department of Economics.
- Lukmanji, Z., Hertzmark, E., Mlingi, N., Assey, V., Ndossi, G., & Fawzi, W. (2008). Tanzania food composition tables. *MUHAS-TFNC, HSPH, Dar es Salaam Tanzania*.
- Masters, W.A., Y. Bai, A. Herforth, D. Sarpong, F. Mishili, J. Kinabo, and J.C. Coates. (2018). Measuring the affordability of nutritious diets in Africa: Price indexes for diet diversity and the cost of nutrient adequacy. *American Journal of Agricultural Economics*, 100(5), 1285-1301.
- National Sample Survey Organisation [NSSO], Ministry of Statistics and Programme Implementation, Government of India (2002). *Results of a Pilot Survey on Suitability of Different Reference Periods for Measuring Household Consumption*, January-June 2000. Report number 475.
- Otten, J., Meyers, L., & Hellwig, J. (Eds.). (2006). *Dietary Reference Consumption: The Essential Guide to Nutrient Requirements*. National Academies Press.
- Perez-Escamilla, R., Bermudez, O., Buccini, G., Lutter, C. & Victora, C. (2018). Nutrition disparities and the global burden of malnutrition. *BMJ* 361:k2252.

- Pingali P, & Sunder N. (2017). Transitioning toward nutrition-sensitive food systems in developing countries. *Annual Review of Resource Economics* 9:439–59
- Popkin, B. (2017). Relationship between shifts in food system dynamics and acceleration of the global nutrition transition. *Nutrition Reviews*, 75(2), 73-82.
- Ramakrishnan, U. (2002). Prevalence of micronutrient malnutrition worldwide. *Nutrition Reviews* 60 (suppl 5): S46-52.
- Sharp, M., Buffière, B., Himelein, K., & Gibson, J. (2019). Effects of Data Collection Methods on Estimated Household Consumption and Poverty, and on Survey Costs: Evidence from an Experiment in the Marshall Islands. IARIW-World Bank conference.
- Smith, L. C., Alderman, H., & Aduayom, D. (2006). *Food insecurity in sub-Saharan Africa: New estimates from household expenditure surveys* (Vol. 146). Intl Food Policy Res Inst.
- Willett, W., Rockström, J., Loken, B., et al., (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *Lancet*, 2019, 393: 447–92.
- WHO. (2017). The double burden of malnutrition. Policy brief. Geneva: World Health Organization.
- World Bank (2018). *World Development Indicators*. 2018
- Zeza, A., Carletto, C., Fiedler, J. L., Gennari, P., & Jolliffe, D. (2017). Food counts. Measuring food consumption and expenditures in household consumption and

expenditure surveys (HCES). Introduction to the special issue. *Food Policy*, 72, 1-6.

APPENDIX

Table A1 | Nutrient deficiencies measured in highly supervised individual diaries

	(1)	(2)	(3)
	Percentage consuming under DRI	Average percent shortfall for those with a shortfall	Average squared percent shortfall for those with a shortfall
Calories	37	23	756
Macronutrients			
Protein	30	30	1,198
Animal protein	86	66	4,886
Fat	71	47	2,763
Saturated fat	48	46	2,700
Cholesterol	98	77	6,469
Carbohydrates	15	18	561
Sugars	41	43	2,560
Fiber	17	20	639
Vitamins			
Vitamin A	53	52	3,439
Vitamin B1	39	24	875
Vitamin B2	16	34	1,685
Vitamin B3	32	25	929
Vitamin B6	10	22	798
Vitamin B9	60	27	1,072
Vitamin B12	59	68	5,262
Vitamin C	38	41	2,320
Vitamin E	44	65	7,558
Minerals			
Calcium	89	55	3,515
Phosphorus	14	21	635
Iron	37	27	1,040
Sodium	93	60	4,024
Potassium	31	26	1,030
Magnesium	19	18	552
Zinc	60	32	1,451

Note: N=503

Table A2 | Interaction terms between module type and household size

VARIABLES	(1) Calories	(2) Protein	(3) Animal protein	(4) Fat	(5) Saturated fat	(6) Cholesterol	(7) Carbohydrates	(8) Sugars	(9) Fiber
Diary: HH frequent * household size	0.792 (1.734)	0.335 (1.500)	0.612 (0.882)	2.948 (4.797)	8.506 (13.84)	0.589 (0.560)	0.754 (1.567)	4.209** (2.117)	0.176 (1.892)
Diary: HH infrequent * household size	0.943 (1.715)	2.119 (1.483)	1.518* (0.872)	2.806 (4.744)	6.382 (13.68)	0.617 (0.553)	0.436 (1.549)	1.699 (2.093)	0.550 (1.870)
Recall: Long, 14 day * household size	-3.084* (1.742)	-2.706* (1.507)	-0.236 (0.886)	1.064 (4.819)	3.231 (13.90)	-0.653 (0.562)	-5.216*** (1.574)	-3.732* (2.127)	-6.676*** (1.900)
Recall: Long, 7 day * household size	-4.864*** (1.814)	-3.718** (1.569)	-0.545 (0.923)	1.082 (5.018)	2.651 (14.47)	-0.166 (0.585)	-8.279*** (1.639)	-8.053*** (2.214)	-8.904*** (1.978)
Recall: Subset, 7 day * household size	-0.295 (1.811)	-1.990 (1.566)	-0.276 (0.921)	12.48** (5.008)	34.66** (14.45)	0.322 (0.584)	-6.686*** (1.636)	-5.847*** (2.210)	-6.414*** (1.975)
VARIABLES	(10) Vitamin A	(11) Vitamin B1	(12) Vitamin B2	(13) Vitamin B3	(14) Vitamin B6	(15) Vitamin B9	(16) Vitamin B12	(17) Vitamin C	(18) Vitamin E
Diary: HH frequent * household size	3.076 (3.802)	0.125 (1.331)	-3.880 (4.094)	0.659 (1.482)	1.669 (2.198)	0.301 (1.630)	2.242 (4.808)	2.193 (2.883)	-1.007 (2.374)
Diary: HH infrequent * household size	1.552 (3.760)	0.266 (1.317)	-3.570 (4.048)	2.232 (1.465)	1.499 (2.173)	1.306 (1.612)	6.264 (4.754)	0.0980 (2.851)	-1.063 (2.347)
Recall: Long, 14 day * household size	-0.457 (3.820)	-3.996*** (1.337)	-5.309 (4.112)	-2.517* (1.489)	-4.919** (2.208)	-5.541*** (1.637)	1.692 (4.830)	-4.709 (2.896)	0.131 (2.384)
Recall: Long, 7 day * household size	-7.560* (3.977)	-5.290*** (1.393)	-6.599 (4.282)	-3.735** (1.550)	-10.11*** (2.299)	-8.482*** (1.705)	0.745 (5.029)	-9.629*** (3.016)	4.956** (2.483)
Recall: Subset, 7 day * household size	-8.847** (3.970)	-4.716*** (1.390)	-7.895* (4.273)	-2.614* (1.547)	-6.610*** (2.294)	-4.772*** (1.702)	0.210 (5.019)	-5.709* (3.010)	1.481 (2.478)
VARIABLES	(19) Calcium	(20) Phosphorous	(21) Iron	(22) Sodium	(23) Potassium	(24) Magnesium	(25) Zinc		
Diary: HH frequent * household size	-0.754 (0.787)	-0.311 (2.044)	-0.302 (1.440)	0.329 (0.795)	1.287 (2.014)	0.788 (1.794)	0.796 (1.085)		
Diary: HH infrequent * household size	0.105 (0.778)	2.015 (2.021)	0.877 (1.423)	1.061 (0.786)	1.019 (1.991)	2.229 (1.774)	1.995* (1.072)		
Recall: Long, 14 day * household size	-2.345*** (0.790)	-5.553*** (2.053)	-4.558*** (1.446)	-1.383* (0.799)	-4.814** (2.023)	-3.555** (1.802)	-1.967* (1.089)		
Recall: Long, 7 day * household size	-3.035*** (0.823)	-6.871*** (2.138)	-5.258*** (1.506)	-3.039*** (0.832)	-9.591*** (2.106)	-5.769*** (1.877)	-2.963*** (1.134)		
Recall: Subset, 7 day * household size	-2.483*** (0.821)	-5.059** (2.134)	-3.957*** (1.503)	0.805 (0.830)	-7.574*** (2.103)	-4.683** (1.873)	-1.252 (1.132)		

Table A3 | Interaction terms between module type and years of education (hhh)

VARIABLES	(1) Calories	(2) Protein	(3) Animal protein	(4) Fat	(5) Saturated fat	(6) Cholesterol	(7) Carbohydrates	(8) Sugars	(9) Fiber
Diary: HH frequent * education (hhh)	-1.315 (1.359)	-2.355** (1.180)	-1.208* (0.678)	-1.666 (3.719)	-3.899 (10.72)	-1.140*** (0.420)	-1.027 (1.235)	2.156 (1.686)	-1.651 (1.470)
Diary: HH infrequent * education (hhh)	-0.173 (1.348)	-1.396 (1.171)	-0.648 (0.672)	-0.220 (3.690)	0.996 (10.64)	-0.642 (0.417)	-0.294 (1.225)	0.780 (1.672)	-1.716 (1.458)
Recall: Long, 14 day * education (hhh)	-1.518 (1.342)	-0.486 (1.166)	0.570 (0.669)	-2.396 (3.673)	-5.371 (10.59)	-0.0633 (0.415)	-1.635 (1.220)	-0.348 (1.665)	-2.021 (1.452)
Recall: Long, 7 day * education (hhh)	-1.810 (1.354)	0.333 (1.176)	1.088 (0.675)	-0.933 (3.705)	-2.277 (10.68)	-0.0296 (0.419)	-3.150** (1.231)	-0.171 (1.679)	-3.363** (1.464)
Recall: Subset, 7 day * education (hhh)	-0.518 (1.343)	-0.288 (1.167)	0.539 (0.670)	1.815 (3.677)	6.076 (10.60)	-0.987** (0.415)	-3.061** (1.221)	2.709 (1.667)	-3.191** (1.453)
VARIABLES	(10) Vitamin A	(11) Vitamin B1	(12) Vitamin B2	(13) Vitamin B3	(14) Vitamin B6	(15) Vitamin B9	(16) Vitamin B12	(17) Vitamin C	(18) Vitamin E
Diary: HH frequent * education (hhh)	-1.965 (2.956)	-1.418 (1.033)	-0.517 (3.173)	-1.418 (1.172)	-0.00764 (1.743)	-0.0849 (1.306)	-5.131 (3.635)	0.581 (2.245)	0.0661 (1.854)
Diary: HH infrequent * education (hhh)	-1.931 (2.933)	-1.092 (1.025)	-5.072 (3.148)	-0.806 (1.162)	-1.075 (1.729)	-0.623 (1.296)	-2.309 (3.606)	-1.282 (2.227)	0.623 (1.839)
Recall: Long, 14 day * education (hhh)	-0.588 (2.919)	-1.408 (1.021)	2.066 (3.133)	-0.838 (1.157)	0.404 (1.722)	-0.0223 (1.290)	-5.791 (3.590)	2.426 (2.217)	-0.912 (1.831)
Recall: Long, 7 day * education (hhh)	-0.889 (2.945)	-2.212** (1.030)	-1.773 (3.161)	-0.739 (1.167)	-2.162 (1.737)	-1.085 (1.301)	-4.159 (3.622)	-2.869 (2.237)	-0.513 (1.847)
Recall: Subset, 7 day * education (hhh)	1.782 (2.923)	-2.501** (1.022)	-2.282 (3.137)	-1.277 (1.158)	-1.828 (1.723)	-0.850 (1.291)	-1.843 (3.594)	-0.795 (2.220)	-0.272 (1.833)
VARIABLES	(19) Calcium	(20) Phosphorous	(21) Iron	(22) Sodium	(23) Potassium	(24) Magnesium	(25) Zinc		
Diary: HH frequent * education (hhh)	-0.284 (0.611)	-2.692* (1.600)	-1.650 (1.126)	0.202 (0.625)	0.317 (1.596)	-0.765 (1.417)	-1.002 (0.861)		
Diary: HH infrequent * education (hhh)	-0.688 (0.606)	-1.923 (1.588)	-1.337 (1.117)	0.249 (0.620)	-0.391 (1.584)	-0.257 (1.406)	-0.535 (0.854)		
Recall: Long, 14 day * education (hhh)	0.758 (0.603)	-1.932 (1.581)	-1.189 (1.112)	1.404** (0.617)	0.387 (1.577)	-3.832*** (1.400)	-2.345*** (0.850)		
Recall: Long, 7 day * education (hhh)	0.776 (0.609)	-1.166 (1.595)	-1.348 (1.122)	1.371** (0.623)	-2.524 (1.591)	-4.382*** (1.412)	-1.554* (0.858)		
Recall: Subset, 7 day * education (hhh)	0.370 (0.604)	-1.687 (1.582)	-1.747 (1.113)	-0.468 (0.618)	-2.124 (1.578)	-4.334*** (1.401)	-1.093 (0.851)		

Table A4 | Interaction terms between module type and wealth

VARIABLES	(1) Calories	(2) Protein	(3) Animal protein	(4) Fat	(5) Saturated fat	(6) Cholesterol	(7) Carbohydrates	(8) Sugars	(9) Fiber
Diary: HH frequent * asset index	-3.351 (5.122)	-7.234 (4.398)	-5.579** (2.540)	-9.343 (14.02)	-24.42 (40.44)	-4.864*** (1.564)	-1.279 (4.666)	7.442 (6.340)	-3.422 (5.581)
Diary: HH infrequent * asset index	-2.723 (5.129)	-4.485 (4.403)	-2.273 (2.543)	-9.039 (14.04)	-18.33 (40.50)	-3.409** (1.566)	-1.861 (4.673)	3.792 (6.348)	-3.996 (5.588)
Recall: Long, 14 day * asset index	-5.198 (5.038)	9.970** (4.325)	7.395*** (2.498)	-8.665 (13.79)	-27.02 (39.78)	1.522 (1.538)	-5.822 (4.589)	2.448 (6.235)	1.488 (5.489)
Recall: Long, 7 day * asset index	-3.307 (5.173)	14.48*** (4.441)	10.21*** (2.565)	-0.630 (14.16)	-6.936 (40.84)	2.127 (1.579)	-8.004* (4.712)	7.533 (6.402)	-0.779 (5.635)
Recall: Subset, 7 day * asset index	-14.41*** (5.151)	7.739* (4.422)	8.814*** (2.554)	-32.91** (14.10)	-96.45** (40.67)	-0.506 (1.573)	-17.16*** (4.693)	14.76** (6.376)	-13.47** (5.612)
VARIABLES	(10) Vitamin A	(11) Vitamin B1	(12) Vitamin B2	(13) Vitamin B3	(14) Vitamin B6	(15) Vitamin B9	(16) Vitamin B12	(17) Vitamin C	(18) Vitamin E
Diary: HH frequent * asset index	-17.12 (11.13)	-2.393 (3.901)	-21.30* (12.05)	-4.454 (4.423)	-4.132 (6.553)	-2.820 (4.931)	-29.74** (13.13)	-2.371 (8.405)	5.280 (6.992)
Diary: HH infrequent * asset index	-15.07 (11.14)	-2.194 (3.906)	-24.99** (12.06)	-3.867 (4.429)	-3.452 (6.561)	-3.301 (4.938)	-22.35* (13.14)	-3.866 (8.416)	5.013 (7.001)
Recall: Long, 14 day * asset index	-5.204 (10.94)	-1.125 (3.837)	9.693 (11.85)	3.527 (4.351)	3.742 (6.445)	4.751 (4.850)	-23.62* (12.91)	7.945 (8.266)	-4.988 (6.876)
Recall: Long, 7 day * asset index	5.971 (11.24)	-4.253 (3.940)	-0.393 (12.17)	4.817 (4.467)	-1.605 (6.617)	2.612 (4.980)	-6.621 (13.25)	-1.326 (8.487)	-8.014 (7.060)
Recall: Subset, 7 day * asset index	2.903 (11.19)	-8.694** (3.923)	-1.886 (12.12)	-1.772 (4.449)	-8.762 (6.590)	0.595 (4.959)	10.09 (13.20)	-18.09** (8.452)	8.660 (7.031)
VARIABLES	(19) Calcium	(20) Phosphorous	(21) Iron	(22) Sodium	(23) Potassium	(24) Magnesium	(25) Zinc		
Diary: HH frequent * asset index	-1.751 (2.292)	-6.687 (6.053)	-3.138 (4.277)	0.221 (2.243)	-2.372 (5.973)	-2.794 (5.368)	-2.368 (3.193)		
Diary: HH infrequent * asset index	1.637 (2.295)	-3.218 (6.061)	-3.737 (4.283)	0.527 (2.246)	-1.912 (5.981)	-3.706 (5.375)	-2.971 (3.197)		
Recall: Long, 14 day * asset index	10.70*** (2.255)	6.881 (5.953)	3.133 (4.206)	14.23*** (2.206)	-1.712 (5.875)	-11.99** (5.280)	-2.271 (3.140)		
Recall: Long, 7 day * asset index	12.49*** (2.315)	9.689 (6.113)	2.438 (4.319)	19.08*** (2.265)	-6.127 (6.032)	-13.47** (5.421)	0.805 (3.224)		
Recall: Subset, 7 day * asset index	6.391*** (2.305)	0.261 (6.088)	-2.678 (4.301)	-1.974 (2.255)	-13.88** (6.007)	-17.90*** (5.399)	-0.0263 (3.211)		

Table A5 | Interaction terms between module type and share of children below 6 years old

VARIABLES	(1) Calories	(2) Protein	(3) Animal protein	(4) Fat	(5) Saturated fat	(6) Cholesterol	(7) Carbohydrates	(8) Sugars	(9) Fiber
Diary: HH frequent * share children below 6	10.37 (30.58)	19.30 (26.52)	5.717 (15.53)	12.83 (84.07)	23.65 (242.4)	1.348 (9.832)	17.77 (27.86)	-11.78 (38.17)	29.66 (33.60)
Diary: HH infrequent * share children below 6	15.94 (30.52)	22.70 (26.46)	7.968 (15.49)	41.66 (83.88)	66.53 (241.8)	5.095 (9.811)	10.26 (27.80)	0.172 (38.09)	19.11 (33.53)
Recall: Long, 14 day * share children below 6	-12.05 (30.20)	-35.37 (26.19)	-18.29 (15.33)	-4.320 (83.01)	2.132 (239.3)	-11.10 (9.708)	-16.66 (27.51)	5.108 (37.69)	-18.45 (33.17)
Recall: Long, 7 day * share children below 6	-30.93 (30.66)	-64.69** (26.59)	-31.29** (15.57)	-2.360 (84.28)	12.26 (243.0)	-19.29* (9.857)	-50.08* (27.93)	-18.75 (38.27)	-54.83 (33.68)
Recall: Subset, 7 day * share children below 6	6.824 (29.73)	-21.79 (25.78)	-22.49 (15.09)	96.54 (81.72)	246.8 (235.6)	1.003 (9.557)	-31.79 (27.09)	-34.12 (37.11)	-16.22 (32.66)
VARIABLES	(10) Vitamin A	(11) Vitamin B1	(12) Vitamin B2	(13) Vitamin B3	(14) Vitamin B6	(15) Vitamin B9	(16) Vitamin B12	(17) Vitamin C	(18) Vitamin E
Diary: HH frequent * share children below 6	29.27 (67.01)	28.81 (23.51)	64.79 (72.04)	29.59 (26.29)	19.22 (39.15)	10.71 (29.28)	-35.75 (84.35)	35.12 (50.17)	-19.91 (41.68)
Diary: HH infrequent * share children below 6	55.25 (66.86)	15.52 (23.46)	7.293 (71.88)	41.05 (26.23)	13.20 (39.07)	4.807 (29.21)	28.56 (84.16)	-3.807 (50.06)	-27.97 (41.59)
Recall: Long, 14 day * share children below 6	15.61 (66.16)	-19.28 (23.21)	-3.571 (71.13)	-11.01 (25.96)	-21.72 (38.66)	-40.52 (28.91)	-16.54 (83.28)	-29.97 (49.54)	3.548 (41.15)
Recall: Long, 7 day * share children below 6	-73.52 (67.17)	-44.21* (23.57)	-33.71 (72.22)	-39.98 (26.36)	-81.35** (39.25)	-69.36** (29.35)	-30.72 (84.56)	-107.7** (50.30)	65.00 (41.78)
Recall: Subset, 7 day * share children below 6	-110.7* (65.13)	-26.98 (22.85)	-43.94 (70.02)	-4.036 (25.56)	-84.37** (38.06)	-57.67** (28.46)	23.02 (81.99)	-102.4** (48.77)	54.60 (40.52)
VARIABLES	(19) Calcium	(20) Phosphorous	(21) Iron	(22) Sodium	(23) Potassium	(24) Magnesium	(25) Zinc		
Diary: HH frequent * share children below 6	-0.842 (13.82)	27.44 (36.03)	20.52 (25.45)	-7.505 (14.15)	17.75 (36.03)	22.25 (32.16)	4.991 (19.35)		
Diary: HH infrequent * share children below 6	-8.966 (13.79)	25.48 (35.95)	15.71 (25.40)	-6.095 (14.12)	12.32 (35.95)	25.76 (32.09)	11.22 (19.30)		
Recall: Long, 14 day * share children below 6	-33.43** (13.65)	-50.62 (35.57)	-30.24 (25.13)	-48.17*** (13.97)	-1.332 (35.57)	3.433 (31.76)	-23.12 (19.10)		
Recall: Long, 7 day * share children below 6	-44.00*** (13.86)	-86.74** (36.12)	-46.27* (25.52)	-50.06*** (14.19)	-38.62 (36.12)	-28.54 (32.24)	-40.52** (19.39)		
Recall: Subset, 7 day * share children below 6	-37.97*** (13.44)	-43.25 (35.02)	-15.50 (24.74)	-0.296 (13.76)	-72.07** (35.02)	-0.154 (31.26)	1.583 (18.81)		

Table A6 | Interaction terms between module type and households living in urban regions

VARIABLES	(1) Calories	(2) Protein	(3) Animal protein	(4) Fat	(5) Saturated fat	(6) Cholesterol	(7) Carbohydrates	(8) Sugars	(9) Fiber
Diary: HH frequent * urban hh	-11.17 (10.75)	-17.92* (9.239)	-11.13** (5.425)	-20.69 (29.37)	-48.49 (84.84)	-8.195** (3.410)	-7.309 (9.712)	6.834 (13.54)	-10.24 (11.76)
Diary: HH infrequent * urban hh	-10.33 (10.76)	-14.33 (9.248)	-6.723 (5.430)	-25.23 (29.39)	-56.82 (84.92)	-8.062** (3.414)	-7.451 (9.721)	-1.195 (13.55)	-9.335 (11.78)
Recall: Long, 14 day * urban hh	-32.99*** (10.74)	-0.385 (9.228)	13.59** (5.418)	-26.98 (29.33)	-56.42 (84.74)	3.065 (3.406)	-44.57*** (9.700)	-1.306 (13.52)	-24.42** (11.75)
Recall: Long, 7 day * urban hh	-25.35** (10.74)	12.60 (9.228)	14.10*** (5.418)	-6.399 (29.33)	-17.34 (84.74)	2.964 (3.406)	-42.05*** (9.700)	-2.167 (13.52)	-16.06 (11.75)
Recall: Subset, 7 day * urban hh	-17.40 (10.74)	-6.289 (9.228)	14.21*** (5.418)	31.99 (29.33)	122.5 (84.74)	0.904 (3.406)	-60.31*** (9.700)	6.096 (13.52)	-58.45*** (11.75)
VARIABLES	(10) Vitamin A	(11) Vitamin B1	(12) Vitamin B2	(13) Vitamin B3	(14) Vitamin B6	(15) Vitamin B9	(16) Vitamin B12	(17) Vitamin C	(18) Vitamin E
Diary: HH frequent * urban hh	-33.11 (23.65)	-9.422 (8.212)	-20.20 (25.30)	-17.95* (9.239)	-10.57 (13.56)	-7.286 (10.30)	-71.59** (28.77)	-8.506 (17.37)	13.29 (14.65)
Diary: HH infrequent * urban hh	-42.61* (23.68)	-7.987 (8.220)	-49.19* (25.32)	-13.45 (9.248)	-9.951 (13.57)	-15.68 (10.31)	-64.34** (28.80)	-15.86 (17.38)	18.17 (14.67)
Recall: Long, 14 day * urban hh	-40.11* (23.63)	-33.51*** (8.202)	3.368 (25.27)	-25.35*** (9.228)	-33.77** (13.54)	-14.75 (10.28)	-43.71 (28.73)	-28.88* (17.35)	13.76 (14.64)
Recall: Long, 7 day * urban hh	-27.36 (23.63)	-31.12*** (8.202)	-0.0773 (25.27)	-16.26* (9.228)	-44.08*** (13.54)	-17.60* (10.28)	-42.74 (28.73)	-30.34* (17.35)	-0.702 (14.64)
Recall: Subset, 7 day * urban hh	-52.76** (23.63)	-44.22*** (8.202)	-13.68 (25.27)	-39.54*** (9.228)	-50.61*** (13.54)	-34.43*** (10.28)	3.258 (28.73)	-68.95*** (17.35)	40.99*** (14.64)
VARIABLES	(19) Calcium	(20) Phosphorous	(21) Iron	(22) Sodium	(23) Potassium	(24) Magnesium	(25) Zinc		
Diary: HH frequent * urban hh	-3.511 (4.846)	-19.28 (12.67)	-10.95 (8.948)	-2.359 (4.829)	-8.101 (12.14)	-7.234 (11.23)	-7.673 (6.676)		
Diary: HH infrequent * urban hh	1.330 (4.850)	-11.71 (12.68)	-11.28 (8.956)	-3.076 (4.834)	-9.472 (12.15)	-12.47 (11.24)	-10.25 (6.682)		
Recall: Long, 14 day * urban hh	14.75*** (4.840)	-22.07* (12.66)	-26.08*** (8.937)	23.71*** (4.824)	-38.84*** (12.13)	-40.37*** (11.22)	-8.863 (6.668)		
Recall: Long, 7 day * urban hh	19.94*** (4.840)	-4.510 (12.66)	-18.96** (8.937)	37.74*** (4.824)	-53.04*** (12.13)	-36.57*** (11.22)	1.814 (6.668)		
Recall: Subset, 7 day * urban hh	5.667 (4.840)	-32.25** (12.66)	-33.61*** (8.937)	-4.437 (4.824)	-71.73*** (12.13)	-61.21*** (11.22)	-8.919 (6.668)		

Table A7 | Count of significant interaction effects (Repeating Table 6, with all characteristics included jointly in the regression)

	(1)	(2)	(3)	(4)	(5)	(6)
	Household size	Education household head	Wealth	Share children under 6	Urban household	TOTAL
1. Diary: HH, frequent	1	1	0	0	0	2
2. Diary: HH, infrequent	1	0	0	0	0	1
3. Recall: Long, 14 day	12	1	14	1	16	44
4. Recall: Long, 7 day	13	0	9	2	14	38
5. Recall: Subset, 7 day	14	2	11	4	17	48
TOTAL	41	4	34	7	47	133

Note: This table repeats Panel B of Table 6 in the main text, but now estimating equation (2) with all 5 household characteristics included jointly in a single regression. The numbers are a count of the number of times an interaction between the module assignment indicated in the row and the household characteristic indicated in the column is significant at 5% level.

Table A8 | Cost of basic needs food poverty

Panel A: Percentage of households spending below threshold						
	Dodoma	Pwani	Dar es Salaam	Shinyanga	Kagera	Total
Diary: Personal	23.22	48.69	53.30	43.06	47.78	43.75
Diary: HH, frequent	43.95	70.30	76.81	60.04	53.74	60.05
Diary: HH, infrequent	31.11	68.21	65.51	51.13	57.52	54.43
Recall: Long, 14 day	39.80	63.02	45.56	76.58	56.47	61.64
Recall: Long, 7 day	25.09	53.08	30.21	57.18	50.53	47.38
Recall: Subset, 7 day	36.41	48.80	44.84	68.74	46.69	53.25
Panel B: Average poverty gap of the food-poor						
	Dodoma	Pwani	Dar es Salaam	Shinyanga	Kagera	Total
Diary: Personal	30.80	42.15	31.52	31.38	27.56	31.76
Diary: HH, frequent	35.01	33.56	38.29	34.07	32.73	34.43
Diary: HH, infrequent	28.46	37.26	42.70	30.36	29.37	32.92
Recall: Long, 14 day	23.22	31.86	32.48	38.28	40.68	36.37
Recall: Long, 7 day	25.44	30.02	28.30	35.18	29.99	31.74
Recall: Subset, 7 day	36.04	30.70	27.25	37.56	28.81	33.47

Figure A1| Benford’s Law and first digits of transaction values reported in the highly supervised individual diaries

Note: The blue bars indicate the frequency of the first digit of reported value of consumption in the individual diary, including 95% confidence intervals. The green line indicates what this distribution is expected to look like under Benford’s Law. We conducted this analysis using the DIGDIS command in Stata (Jann, 2007).