INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

RESEARCH PROGRAM ON Agriculture for Nutrition and Health

Led by IFPRI

IFPRI Discussion Paper 01793

December 2018

Income Variability, Evolving Diets, and Demand for Processed Foods in Nigeria

Alan de Brauw

Sylvan Herskowitz

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Alan de Brauw (a.debrauw@cgiar.org) is a Senior Research Fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute, Washington, DC.

Sylvan Herskowitz (s.herskowitz@cgiar.org) is an Associate Research Fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute, Washington, DC.

Income Variability, Evolving Diets, and Demand for Processed Foods in Nigeria

Alan de Brauw and Sylvan Herskowitz*

International Food Policy Research Institute

December 2018
Link to most recent version

Abstract

We present evidence on evolving dietary patterns in Nigeria using three waves and six total rounds of household consumption data from the Nigerian Living Standards and Measurement Surveys between the years of 2011 and 2016. First, following conventional definitions in the literature, we show that Nigeria has not shown any aggregate increase in consumption of highly processed foods over this time period, contrary to studies elsewhere in the region. In fact, consumption of highly processed foods at home has decreased, while food away consumed away from home has risen substantially. We then show that estimates of food expenditure elasticities of different food types are highly sensitive to different estimation approaches, raising concerns regarding the existing evidence base on food consumption patterns reliant on estimation of food expenditure elasticities. Different specifications can lead to broadly differing conclusions about whether highly processed food is either the most or least elastic food category. In our preferred specifications, we find that elasticity of demand for food away from home is highest for the relatively wealthy and in the urban South. Within households, elasticities are highest in times of scarcity, suggesting that households cut food away from home when resources are relatively scarce.

Keywords: Nutrition, Household Shocks, Vulnerability
JEL Codes: I15, E21, Q18

*Markets, Trade and Institutions Division, International Food Policy Research Institute, 1201 Eye St NW, Washington, DC 20005 USA. We thank Rachel Huang and Katie Sproule for their research support. This work was undertaken as part of, and funded by, the CGIAR Research Program on Agriculture for Nutrition and Health (A4NH). Any opinions expressed in this paper belong to the authors, and do not necessarily reflect those of A4NH or CGIAR. Email: s.herskowitz@cgiar.org.

Contents

1 Introduction

While the twenty-first century has seen substantial poverty reduction and improved food security from a caloric perspective, several types of malnutrition persist. Micronutrient malnutrition remains stubbornly high in much of the developing world (Butta and Salam, 2012), while the incidence of obesity has been rising throughout the world (International Food Policy Research Institute, 2016). More people are able to meet their caloric needs, but food systems may not always make nutritious food options accessible to consumers. In many countries, even relatively wealthy consumers frequently lack access to the types of food needed for a healthy, balanced diet (Alston et al., 2016).

As the economies of most countries in the developing world have been growing in recent decades, their food systems have been undergoing rapid change. Political biases favor policies that promote improved grain yields and grain self-sufficiency, leading to higher relative prices for more nutrient dense foods (Pingali, 2015). At the same time, Rising rates of urbanization (Food and Agriculture Organization, 2018) are bringing people to cities, shifting diets towards foods with less micronutrient density (e.g. Cockx et al., 2018). And the continued development of supermarkets is lengthening value chains and affected the composition of food supply (Reardon et al., 2013). All of these factors are likely contributing to a broad increase in demand for processed foods which, along with a shift towards more sedentary lifestyles, are harming global health and nutrition as evidenced by rising rates of overweight and obesity (Popkin et al., 2012).

While these shifts have been documented globally and in more developed countries, the evidence on consumption trends in developing countries is thinner. One exception is Tschirley et al. (2015), who argue that the consumption of highly processed foods is rising rapidly across Africa, even among the relatively poor. They present descriptive evidence showing demographic shifts and trends in food consumption towards highly processed foods. Additionally, they calculate food demand elasticities to project future demand for processed foods across different regions and for different economic groups. Arguing that food demand elasticities are greater for foods with higher levels of processing, they suggest that rising incomes will result in greater demand for highly processed foods.

This paper uses the Nigeria Living Standards Measurement Survey (LSMS) data to contribute to the growing literature on shifting diets, demand for processed foods, and food expenditure elasticities. The paper makes three main points. First, we show that data limitations affect the way we understand trends related to both the incidence and the increase in consumption of processed foods. Second, the estimation of elasticities is strongly dependent on choices made about functional form. Different analysis choices substantially affect estimates, which has important implications for conclusions about how future demand will change for different classes of goods. Third, we show that within household income fluctuations also substantially affect demand for processed foods and food away from home, in ways that are obscured by broad descriptive analyses or estimation of food expenditure elasticities.

To make these contributions, the paper uses consumption expenditure modules enumerated in the Nigerian Living Standards Measurement Survey (LSMS) panel, collected between 2011 and 2016. We begin by noting that other authors have categorized all food eaten away from home as highly processed (Monteiro et al., 2010; Tschirley et al., 2015). Following this categorization, the data in Nigeria suggest that there has not been a dramatic increase in consumption of highly processed foods over this period. Average weekly value of food consumption of highly processed foods held constant at just over 5 USD per capita and 27% of total food consumption value.

However, most surveys on food consumption do not include details about how meals and snacks consumed away from home are prepared, and the Nigerian LSMS is no exception. Assuming that all food consumed away from home is highly processed almost certainly overstates the amount of highly processed foods being consumed. Therefore, we split highly processed food into that consumed at home and away from home, and show consumption of food away from home has increased substantially over this time period in Nigeria, from 2.39 USD per capita to 3.04 USD per capita, while consumption of highly processed foods eaten at home has, in fact, declined. Increasing food consumption away from home has important potential policy relevance, as the implications of increased food demand away from home are different than those for an increase in demand for highly processed foods consumed at home.

Economists and other researchers have long been interested in the estimation of food demand elasticities as a way to understand how diets adjust as prices and incomes change. Bouis and Haddad (1992) was one of the first papers to consider measurement issues in relating variables associated with nutrition outcomes to income. A more recent literature has attempted to measure the responsiveness of micronutrients in the diet to income (e.g. Skoufias et al., 2012). However, consumers do not typically demand micronutrients, but the foods that contain them. Following others such as Monteiro (2009), an alternative is to measure elasticities of food demand based on the level of processing, as in Tschirley et al. (2015). Yet to the best of our knowledge, the literature lacks an analysis of food demand by level of processing that follows the same group of consumers over time.

We contribute to this literature by estimating food expenditure elasticities for different categories of food by their level of processing. Well known issues arise when applying a logarithmic transformation to data containing zeros conventionally used in the estimation of elasticities: these observations get dropped from the analysis. In response, adjusted logarithmic functions such as $\log(1 + x)$ and, more recently, the inverse hyperbolic sine transformation are used as a way to approximate the properties of a log transformation without dropping zeros. However, these alternatives result in different estimated elasticities while introducing a new set of problems, highlighting that a clear "right" method does not exist. Our analysis shows that the choice of specification used to estimate these elasticities has first order implications for their resulting magnitudes as well as for their relative ranking across categories.

Acknowledging these concerns about estimating food demand elasticities, we next observe meaningful heterogeneity in demand response within Nigeria. Focusing on the elasticity of food consumption away from home, we find that the North, rural areas, and the relatively poor, all have substantially lower estimated elasticities for food away from home than their counterparts. Finally, we observe considerable variation in overall food expenditures across the six waves of the survey. We use this variation to separately estimate food demand elasticities when households experience times of relative scarcity or abundance, and find that they vary substantially. We confirm this finding with semi-parametric estimates of food expenditures away from home on overall food expenditures.

The paper proceeds as follows. The next section describes the data used in the paper and presents descriptive statistics related to food demand. Section three presents estimates of income elasticities, and the section four concludes with implications for further research.

2 Data Description and Summary Statistics

2.1 Data and Data Treatment

This paper uses all six rounds of interviews conducted as part of the Nigerian Living Standards Measurement Survey-Integrated Surveys for Agriculture (LSMS). There have been three waves of the LSMS in 2011-2012, 2013-2014, and 2015-2016. Each wave targeted the same set of 5,000 households and was designed to be representative of the population at the zonal level, which includes three Northern and three Southern regions as well as by rural and urban areas within these zones. Each wave of the LSMS contained two separate visits timed to learn about different points in the agricultural season: post-planting and post-harvest. The survey therefore attempted to interview each household six times. Response rates declined over time but we use the full set of data wherever possible.[1]

[1]Wave 1, post-planting included 4,206 respondents, from the initial targeted list of 5,000. Wave 3, post-harvest (the final visit) included 3,937 respondents. The possibility of selective attrition provide additional motivation for our preferred analysis specification which looks at variation in consumption within house-

A comprehensive household expenditures module was collected in each survey round. The section on food consumption asked respondents about more than 100 different food items.[2] For each food type, the respondent, typically the household head or spouse responding on behalf of the household, reported how much of this item had been consumed by the household in the previous week. They then reported how much was purchased for consumption and how much was produced by the household. Using reported expenditures for purchased foods and reported quantities for consumption from household production, we calculate the value of consumption for each type of food.[3]

In addition to modules on food prepared and consumed within the home, the LSMS also asks about meals outside the home. Respondents were asked about nine different categories of food away from home: breakfast, lunch, dinner, side dishes, snacks, dairy based beverages, vegetables, non-alcoholic drinks, and alcoholic drinks.[4] However, instead of having the respondent report the ingredients or quantities of these meals, these broad categories were not further defined. Respondents were asked whether anyone in the household had consumed any of each type of meal "prepared and consumed outside the home." Following an affirmative response for a category, they were asked to report how much was spent on that category in the last seven days. The type and quantity of food consumed is unknown.[5] Unfortunately, given the broad nature of these groups, it

holds while controlling for time by survey cluster fixed effects.

[2]In total, the consumption module included 120 different food types included in all three waves. Later waves expanded the list of foods or split apart foods that demanded more precision, such as maize being split into shelled and non-shelled groups.

[3]To assign values to reported quantities, we followed the Deaton and Zaidi (2002) method, attributing the median price given in the primary sampling unit for purchases of that item. If a price is not available at the primary sampling unit level, we used the median at the smallest possible geographic level above that.

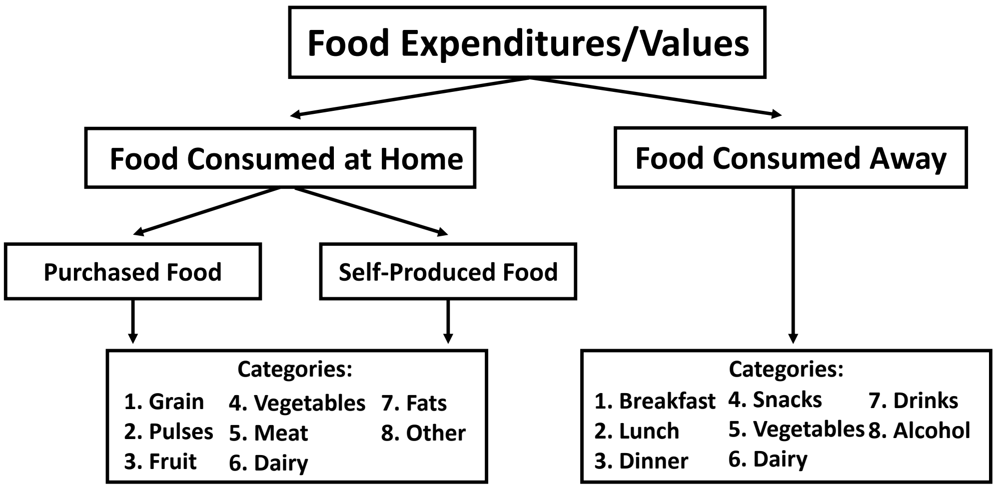
[4]The survey only provides limited additional description of these meals prepared and consumed outside the home. Breakfast, lunch, and dinner: "Full meals (e.g. rice and stew, pounded yam and egusi, etc)". Side dishes: "Side dishes like pepper soup, nkwobi, suya etc." Snacks: "Snacks such as sandwiches, biscuits, meatpies, donuts, pofpof, etc." Dairy: "Dairy based beverages such as milk, yogurt etc." Vegetable: "Vegetables and roasted such as (carrot, pears, roasted corn and plantain, sugar."

[5]This issue is common with many of the LSMS surveys and with multi-topic surveys in general. The Ethiopia LSMS uses the same module as the Nigerian LSMS for food away from home. The Tanzania LSMS asks the same set of questions, but has respondents report the values for each household member separately and thus has the same limitations. The Uganda LSMS asks about food away from home even more broadly, using a single question on total expenditures at restaurants over the past seven days.

is not possible to compute the quantities, processing level, or nutritional content of these meals.

Figure 1 summarizes the categories of food expenditures captured in the data.

Figure 1: Categories of Food Expenditures in LSMS



Notes: Figure created by authors, based off of LSMS survey questions.

For analysis, we follow Monteiro et al. (2010) and Tschirley et al. (2015) as closely as possible, categorizing each food item by its degree of processing. Tschirley et al. describe their criteria for high, low, and unprocessed foods as follows:

> Foods are 'unprocessed' if they undergo no transformation from their original state beyond removal from the plant and (for non-perishables) drying; examples include pulses, whole grains and fresh fruit and vegetables. Processed foods are assigned to the 'low value added' category if they satisfy only one of the following three conditions: have multiple ingredients; underwent physical change induced by heating, freezing, extrusion or chemical processes (i.e. more than simple physical transformation); and have packaging more complex than simple paper or plastic. Examples in this 'low processed' class include maize meal and milled rice. Foods meeting two of the three categories are classified as high value-added processed; examples are breads and other bakery products, industrially packaged vegetable oils and food away from home.

Appendix Table A.2 documents our assignment of processing for each of the 120 types of food recorded in the LSMS surveys.

2.2 Food Consumption Trends

We first note that overall household per capita weekly expenditures on food are relatively low. Table 1 shows consumption levels and shares of a set of different food categories. Households were ranked by percentiles of food per capita value. Values reported are the mean levels or shares for households within five percentile points of the percentile listed at the top of each column. The "Mean" column is the mean from all respondents.[6]

The table shows that mean (median) total food value is 19 USD (17 USD) per capita per week during the study period. The table also splits the value of food consumption by location, at home or away from home, and for food eaten at home, whether this food was produced or purchased by the household.

The average respondent household purchases approximately 68% of the food they consume by value. The mean (mean at the median of food expenditures) share of total food expenditures on food consumed away from home is 11% (10%). The table also shows the division of consumption by degree of processing. The largest share of consumption is non-processed food, at 47.5% on average, while highly processed foods constitute 28% of total consumption value. However, a large share (39%) of food categorized as "highly processed" is food consumed away from home. As noted above, the preparation of that food is, in fact, unknown and may or may not be highly processed.

We note that households with higher levels of overall food expenditures have higher shares of food value consumed away from home, from 7.8% at the tenth percentile to 13.9% at the 90th percentile. Consumption value from own production falls from 25% to 13% while the share on highly processed foods rises modestly from 24.6% to 31.8%.

The last two rows of Table 1 describe the overall variability or "spread" of both total

[6]Note that, because values listed are averages of people in a given bin, group totals, such as share of food value eaten at home, may not perfectly match the sum of the sub-groups: purchased and produced.

Table 1: Food Consumption Shares by Percentile of Food Expenditures

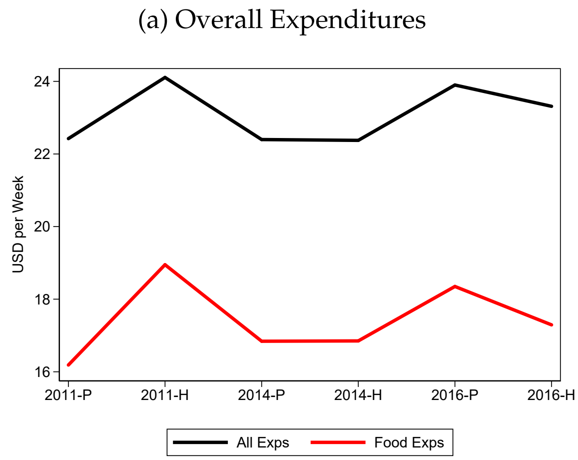
	Mean	p10	p25	p50	p75	p90
Per Capita Weekly Expenditure Totals:						
Total Expenditures (USD)	25.883	12.138	15.946	22.431	32.259	44.403
Food Value (USD)	19.460	9.599	12.573	17.029	23.727	32.330
Food Value / Total Exps	0.777	0.812	0.808	0.771	0.753	0.747
Share of Food Value by Source and Location:						
Away	0.111	0.078	0.097	0.098	0.133	0.139
Home	0.889	0.922	0.903	0.902	0.867	0.861
Purchased	0.683	0.647	0.665	0.684	0.704	0.703
Produced	0.186	0.250	0.223	0.200	0.139	0.130
Processing:						
Non-Processed	0.475	0.541	0.500	0.474	0.455	0.438
Low Processed	0.244	0.214	0.229	0.259	0.249	0.245
High Processed	0.282	0.246	0.272	0.268	0.296	0.318
Home	0.171	0.164	0.174	0.169	0.172	0.178
Away	0.111	0.082	0.097	0.099	0.125	0.140
Expenditure Spread: (Max - Min)/Mean						
All Expenditures	1.220	1.148	1.096	1.133	1.246	1.357
Food Expenditures	1.126	1.089	1.020	1.042	1.145	1.235

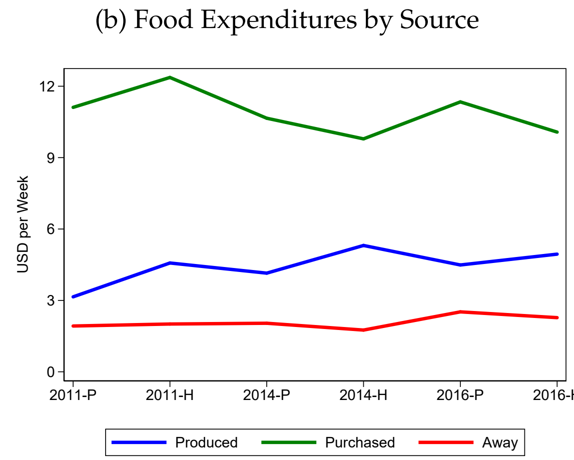
Notes: Reported expenditures are means of per capita household expenditures in US dollars. Expenditures include the value of food produced and consumed by the household. Households were ranked by levels of overall value of food consumption. Other expenditures, shares, and spreads are reported as the mean *for people within 5 percentile points of the percentile listed in the column header*. Expenditure spreads are calculated for each household by taking reported expenditures for each of the six survey rounds, subtracting the smallest total from the largest, and scaling by the household's mean. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

and food expenditures, defined for each household as the difference between the maximum computed expenditures and the minimum, divided by average expenditures for each household. This measure shows considerable variability in expenditures across the six survey rounds. The average of 1.22 for total expenditures suggests that, on average, the difference between the maximum and minimum reported expenditures is greater than that household's mean for the six survey rounds by 15 percent. Food expenditures are slightly less variable, with a spread of 1.126, consistent with food being a less elastic than non-food expenditures. Some of this variability may reflect measurement error, but it implies that there is a great deal of within household variation to exploit in measuring elasticities.

While these descriptive statistics on food shares provide a useful characterization of the state of Nigerian food consumption, they do not tell us how they are changing over time. Panel (a) of Figure 2 shows that average total and average food expenditures may have grown slightly on average, but were mostly level over this period. Panel (b) shows that the value of food produced at home increased more dramatically over the study period, while the value of food purchased for consumption at home declined over time.

Figure 2: Food Consumption Trends by Survey Round





Notes: P = survey conducted post-planting. H = survey conducted post-harvest. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

Recently considerable concern has emerged about the way diets are evolving in most countries and, in particular, about increased demand for highly processed foods that may be detrimental to health and nutritional status (International Food Policy Research Institute, 2016). The black lines in Figure 3 show the values and shares of expenditures on highly processed foods. The levels, in Panel (a) are mostly flat across the six years of the study while the shares of consumption value, in Panel (b), show some modest decline. While informative, these aggregate trends obscure meaningful heterogeneity between processed food consumed at home, where processing levels at the time of purchase are well-known, and the value of meals eaten away from home, for which the processing level at time of consumption is unknown. Whether we examine the expenditure levels or shares, there is a decline in processed foods consumed at home and an increase in food eaten away.

Table 2 tests these differences in regression form, taking the first and last two survey rounds of the study. The main expenditure groups are listed at the top of each column and were regressed on an indicator for being in the final wave (the fifth and sixth visits) and a dummy for whether it was the post-harvest visit (the second and sixth visits) to account for seasonality. Similar to the figures we observe a modest (and only marginally statistically significant) increase in reported total and food expenditures. We also see a large increase in produced food and decrease in purchased food. We see no significant difference in consumption of highly processed food, however we observe a statistically significant increase in food eaten away of approximately 19%, while highly processed foods eaten at home fall by 14%, which is also statistically significant.

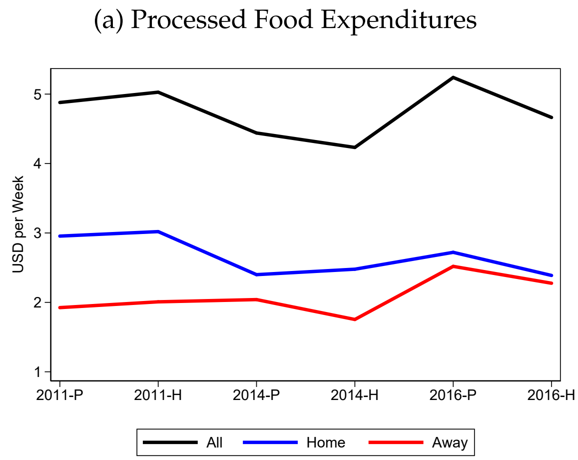
Table 2: Test of Per Capita Expenditures Changes between First and Last Survey Wave

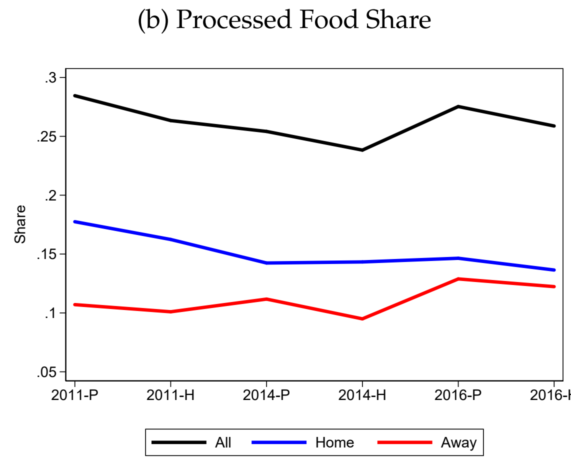
	(1) Total Exps	(2) Food Exps	(3) Produced Food	(4) Purchased Food	(5) High Process (HP)	(6) HP Home	(7) HP Away
Final Wave (0/1)	0.565*	0.384*	0.869***	-0.895***	0.011	-0.399***	0.410***
	(0.248)	(0.191)	(0.127)	(0.146)	(0.077)	(0.048)	(0.053)
Post Harvest (0/1)	0.696***	0.933***	0.905***	0.078	-0.167**	-0.117**	-0.050
	(0.160)	(0.138)	(0.103)	(0.106)	(0.059)	(0.041)	(0.039)
Mean	23.428	17.688	4.275	11.238	4.953	2.777	2.176
Scaled Difference	0.024	0.022	0.203	-0.080	0.002	-0.144	0.188
N	14467	14538	14538	14538	14538	14538	14538
R2	0.585	0.472	0.380	0.535	0.501	0.429	0.477

Notes: Regressions include household fixed effects. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

Given their large share of foods typically categorized as "highly processed," an understanding of these trends is unlikely to be possible without better information on the quantity, treatment and processing of meals eaten away from home. For example, neither a bowl of rice and beans or a piece of fruit would be considered highly processed if eaten at home, but that same bowl of rice and beans or fruit would be labeled highly processed if eaten away from home, thus leading to a likely overstatement of the trends towards increased consumption of highly processed food over time. An increase in a categorization of highly processed foods which includes both categories does not tell us which, *a priori* is changing more quickly, and in this case the disaggregation tells a different story than the aggregate.

Figure 3: Processed Food Value Share by Survey Round



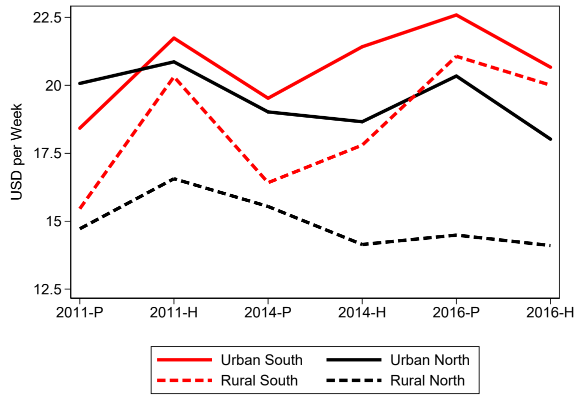


Notes: P = survey conducted post-planting. H = survey conducted post-harvest. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

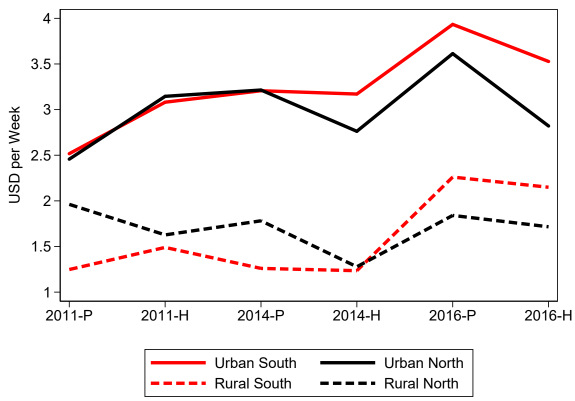
The previous figures and table reveal important national trends, but may mask meaningful heterogeneity by region or location. Food availability and wealth vary considerably across different parts of Nigeria; therefore, we split the data into four groups: the Urban South, Rural South, Urban North, and the Rural North (Figure 4). Panel (a) shows substantial heterogeneity in overall food expenditure trends between the North and the South. In both urban and rural regions of the South, food expenditures have been climbing over time, whereas expenditure levels are falling slightly in both North locations. Panels (b) and (c) show that expenditures on food away from home have grown in urban areas of both the North and the South, as well as in the rural South. However, the rural North has shown a slight decline.

Figure 4: Overall and Food Away from Home Expenditures over Time

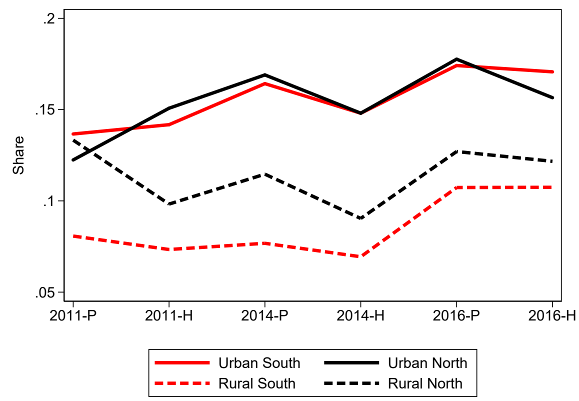
(a) Food Expenditures Per Capita



(b) Exps Per Capita, Food Away from Home



(c) Share of Exps, Food Away from Home



Notes: P = survey conducted post-planting. H = survey conducted post-harvest. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

We note that these trends are difficult to disentangle from other national or regional level shocks. For example, shocks to prices of key commodities or components of household consumption could substantially affect the assessed value of food consumption.[7] Furthermore, differences in data collection protocols or practices across survey rounds could also lead to trends in the data that do not accurately reflect realities on the ground. One way to address these large correlated changes and gain further insight on food de-

[7]Though we control for general inflation in putting the data in real terms, if specific prices rose relative to the mean it could affect the assessed value of food consumption without changing the consumption bundle.

mand patterns is to switch to a regression based approach and estimate food demand elasticities where we can control for survey round fixed effects and adjust for potentially problematic cross-sectional differences in levels. We explore both the challenges and insights of this type of approach in the following two sections.

3 Income Elasticities of Food Demand

3.1 Estimation Challenges of Food Demand Elasticities

In this section, we aim to estimate food expenditure elasticities for foods with different levels of food processing. Elasticities can be modeled with a modified version of the equation that results from utility maximization by a household facing a Cobb-Douglas utility function:

$$\log(Y^k) = \beta_0 + \beta_1 \log(X) + \sum_{i=1}^{k} \beta_k \log(p_k) + \mu \log(H) \quad (1)$$

where Y is the value of food consumed within a specific type k, X is total food expenditures, the vector $\mathbf{p}$ are prices for all types of food, and H represents household size. To develop a version of equation (1) that can be estimated, we first note that prices are nearly impossible to define for classes of food as broad as the types of processing being used, so we replace them with a set of geographic area by time fixed effects, $\mu_r \times \eta_t$, where r indexes geographic areas and t indexes time. These survey enumeration areas are very small, a town or cluster of villages, and we therefore expect these time by geographic area fixed effects to capture local prices better than the broad geographic areas used in many other studies where geography or local infrastructure are likely to have considerably greater variation. We further specify a set of household size fixed effects, θ_{st}, where

s indexes each specific household size.[8] Our initial estimating equation is therefore:

$$\log(Y)^k_{isrt} = \beta_0 + \beta_1 \log(X)_{it} + \theta_{st} + \mu_r \times \eta_t + \varepsilon_{isrt} \tag{2}$$

where i indexes households. This approach is similar to a cross-sectional approach in that it does not control for household fixed effects, and it may be specified with large geographic areas over which prices for single commodities may vary. With six rounds of data in the Nigerian LSMS, we can instead use panel methods. In this preferred approach, we control for household level fixed effects, δ, and additionally switch to a lower geographic unit replacing the geographic area index with a cluster level index, g:

$$\log(Y)^k_{isgt} = \beta_0 + \beta_1 \log(X)_{isgt} + \delta_{ig} + \theta_{st} + \mu_g \times \eta_t + \epsilon_{isgt} \tag{3}$$

The household level fixed effects capture time-invariant household-level factors that influence food demand such as taste or dietary restrictions. Further, the interaction term $\mu_g \times \eta_t$ now controls for cluster level effects over time, which should better capture relative prices for classes of goods. Estimates for β_1 can be interpreted as the percent change of Y associated with a percent change in X, holding other observed factors constant.[9] We estimate equation (2) to better match the cross-sectional elasticities that have been estimated elsewhere in the literature, and equation (3) as a preferred specification that better controls for both fixed differences among households and dynamic factors such as local prices.

Setting aside important issues of endogeneity, several challenging issues affect estimation of food demand elasticities. Notably, when households do not report any consump-

[8]Household size fixed effects are included to account for systematic differences in levels of reported food consumption that research attributes to economies of scale in reported food consumption (Deaton and Paxson, 1998).

[9]Without either a valid instrument for X or randomized variation in X, we cannot put a causal interpretation on the coefficient. Unmeasured factors correlated with both the dependent and independent variables can bias these estimates in either direction.

tion within a specific consumption category, a logarithmic transformation of the outcome variable, as in equations (2) and (3), is undefined and therefore omitted in the estimation. This omission distorts the estimated relationship between independent and dependent variables by ignoring variation on the extensive margin as households switch in and out of positive consumption: these estimations suffer from a selection problem. Additionally, selection will be different across food categories for any household with positive food consumption when zeros are present, which exacerbates selection issues when comparing elasticities across food groups. One could trim the data to only include households with positive consumption of all food categories of interest but, depending on the number of categories desired for analysis, doing so will make selection even more severe. As each outcome variable imposes an additional sample restriction, the analysis sample will become less and less representative of the study population.

To address these selection issues, economists have used an assortment of ad hoc adjustments such as adding a constant to the consumption level before taking its logarithm in order to avoid the selection issues discussed above (e.g. $\log(1+Z)$ or $\log(0.01+Z)$). More recently, use of the inverse hyperbolic sine transformation (IHST), $IHST(Z) = \log(Z + (Z^2+1)^{\frac{1}{2}})$, has become more frequent in the literature.[10] However, issues resulting from this transformation have begun to emerge. Bellemare and Wichman (2018) provide a detailed discussion of how to convert regression estimates with IHST into conventional elasticities, adjusting for a distortion that results from low average values of the underlying variables. They suggest that if the mean of the underlying variables is small (below ten), that they be rescaled by multiplying the original value by a factor large enough to reduce the influence of the arbitrary addition of one in the (Z^2+1) term.

Despite enthusiasm for the IHST transformation in the literature, the (Z^2+1) term also represents an arbitrary distortion of the underlying variable similar to that in logarithmic adjustments. This results in two further issues. First, whereas we can derive

[10] Examples of recently published papers using the IHST include McKenzie (2017), Clemens and Tiongson (2017), Jayachandran et al. (2017), and Bahar and Rapoport (2018).

equation (2) and (3) from a well-behaved preference relation, the same is not true when using the IHST in place of a log transformation. As a result, similar to adding an arbitrary constant to the variable, the use of the IHST reduces the theoretical validity of elasticity estimates. Second, we show in Appendix B that as the incidence of zeroes in the dependent variable increases, the elasticities resulting from a log transformation with a constant added and an IHST transformation deviate sharply, depending on the scaling or shifting factor of the transformation. As a result, the IHST transformation may be introducing a different type of bias than the selection issues that result from use of unadjusted logarithms.

Since there is no ideal way to estimate these elasticities, we proceed as follows. First, we provide the incidence of consumption of each category of foods by processing level: unprocessed, low processing, and high processing. Due in part to this concern about its definition, we further break up highly processed food into two subcategories: that consumed at home and that consumed away from home. Next, we provide an elasticity for "any" consumption– meaning that we estimate the percentage increase in incidence of positive consumption within that category if total food expenditures were to increase by a specific percentage. Then we estimate elasticities based on the value of consumption using five specifications– levels, unadjusted logarithms, logarithms adding one, and then two versions of the IHST, which vary by the multiplier on the monetary value in dollars.[11]

3.2 Elasticities and Functional Form

Table 3 shows elasticities and regression results that use the value of consumption from each level of processing as dependent variables in equations (2) and (3). Panel (a) provides elasticity estimates using the LSMS as a set of repeated cross-sections, using equation (2).

[11] By levels, we mean that we are not adjusting the underlying food values in the regressions with either a logarithmic or IHST transformation. In all cases except the binary outcome variable for positive consumption, we treat the dependent and key independent variables in the same way. For example, if we use $\log(Y+1)$ as the dependent variable, we use $\log(X+1)$ as the independent variable.

This specification resembles most recent work in the literature. The implicit assumption in Panel (a) is that location-region fixed effects proxy for prices, and that there are are no unobserved variables (say, at the village level) that would be correlated with both food shares of different food types and total food expenditures. Panel (b) uses equation (3), which includes household level fixed effects to control for time invariant household tastes and unobservables, and switches to a finer measure of location by time fixed effects that more closely capture dynamic factors affecting household consumption decisions. Panel (b) is motivated by the assumption that within household variation can better identify the relationship between food budget allocations in these elasticities.

Table 3: Elasticities from Different Functional Forms

Panel (a): Cross-Sectional Estimation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Transformation	Incidence	Any (0/1)	Level	Log(Z)	Log(1+Z)	IHST(100xZ)	IHST(10000xZ)
Unprocessed	0.997	0.007	1.29	1.002	0.906	1.038	1.072
		(0.001)	(0.003)	(0.005)	(0.004)	(0.006)	(0.007)
Low Processed	0.959	0.059	0.704	0.812	0.652	1.101	1.371
		(0.002)	(0.002)	(0.008)	(0.006)	(0.013)	(0.021)
High Processed	0.987	0.03	0.789	1.052	0.853	1.197	1.333
		(0.001)	(0.002)	(0.008)	(0.006)	(0.01)	(0.014)
Home	0.977	0.042	0.868	0.873	0.631	1.07	1.264
		(0.002)	(0.002)	(0.008)	(0.005)	(0.011)	(0.017))
Away	0.723	0.181	0.71	0.946	0.652	1.659	2.495
		(0.004)	(0.002)	(0.013)	(0.008)	(0.027)	(0.047))
Panel (b): Fixed Effects Estimation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Transformation	Incidence	Any (0/1)	Level	Log(Z)	Log(1+Z)	IHST(100xZ)	IHST(10000xZ)
Unprocessed	0.997	0.012	1.382	1.017	0.918	1.066	1.12
		(0.001)	(0.003)	(0.007)	(0.006)	(0.008)	(0.010)
Low Processed	0.959	0.052	0.666	0.775	0.615	1.014	1.255
		(0.003)	(0.002)	(0.01)	(0.007)	(0.018)	(0.029)
High Processed	0.987	0.031	0.672	0.931	0.759	1.073	1.213
		(0.002)	(0.003)	(0.01)	(0.007)	(0.013)	(0.019)
Home	0.977	0.044	0.777	0.754	0.549	0.951	1.153
		(0.002)	(0.002)	(0.01)	(0.007)	(0.015)	(0.024)
Away	0.723	0.136	0.541	0.868	0 .555	1.324	1.951
		(0.006)	(0.002)	(0.018)	(0.01)	(0.034)	(0.06)

Notes: Standard errors clustered at the primary sampling unit level in parentheses. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

First, we show that the incidence of consumption is quite high for each of the three categories of consumption (column 1); the lowest are lightly processed foods, at 95.9 percent.[12] Therefore, for the main categories, the logarithmic transformation does not drop many observations. The subcategory of food eaten away from home has a much lower incidence, at 72.3 percent.

In column (2), the outcome variable is a binary indicator and the logarithm of the total value of food consumption is used as the primary independent variable. Therefore, for this column the reported estimate for β_1 can be interpreted as follows. If the total value of food expenditures increased by 10%, the share of people consuming each food type would increase by $0.1 \times \beta_1$. The coefficients are small in magnitude in most categories, as we would expect given high overall incidence of positive consumption. The exception is the subcategory of food eaten away from home in the fifth row of column (2); a 10 percent increase in the value of food consumption would increase the incidence of households reporting to have consumed food away from home by 1.8 percentage points.

Columns (3)-(7) present elasticity estimates from different econometric specifications for each of the three categories of processing and two subcategories of highly processed foods. Examining the cross-sectional elasticities in Panel (a), we first note that the estimates vary substantially. Perhaps most importantly, the highest elasticity among categories switches depending upon the econometric specification. For example, in column (3) unprocessed foods are the most elastic; if we believed that model best reflected the data generating process, then we would associate a 10% increase in overall food consumption value with a 13% rise in unprocessed foods, while highly processed foods would increase by less than 8%. However, if we had some reason to prefer using IHST with a scaling factor of 100, as in column (6), we would instead conclude that highly processed foods were the most elastic, and that food away from home was highly elastic with a 17% increase associated with a 10% increase in overall food expenditures.

As previously discussed, the cross-sectional elasticity estimates confound the way that

[12] Note that households are purchasing unprocessed foods, but then might process them in the homes.

changes in food demand respond to changes in overall food expenditures with other cross-sectional factors. The regressions in Panel (a) attempt to account for those cross-sectional factors through region by survey round fixed effects, but infrastructure or food availability are still likely to vary within regions and may substantially bias elasticity estimates. Moreover, (unobserved) prices likely vary meaningfully within these regions as well. To better control for these unobservable factors, Panel (b) includes household fixed effects along with cluster by survey round fixed effects as detailed in equation (3).

In Panel (b), we find that elasticity estimates tend to increase slightly for unprocessed foods, but decrease for low and highly processed foods. However, we still see troubling ordinal inconsistency across specifications. Whether or not household fixed effects are included in the estimates, the ordering of elasticities from highest to lowest is highly dependent upon functional form and the transformation applied to the data.

A second problem with the functional forms in columns (5) through (7) of both panels is that the group of elasticities generated are collectively infeasible. Using total food expenditures as the independent variable, it must be that among the three, comprehensive categories at least one is elastic and one is inelastic; if food expenditures increase by 10 percent, then the three categories making up food consumption cannot all increase by less than 10 percent, nor can all increase by more than 10 percent. Yet the $\log(Z+1)$ adjustment generates elasticities that are inelastic for all three categories, and both IHST adjustments lead to elasticities that are all above one. As discussed above, the scaling addresses issues that arise from low levels of the underlying variables; without scaling, the estimates all suggest inelastic demand.

Therefore, we are left with two plausible specifications. The specification in logarithms is preferred as it can be derived from a well-behaved utility function, and the estimates are simply the estimated parameter; for levels, the elasticities must be measured at the average level of consumption. Though the log transformation drops all respondents with zero demand in the category, we can at least observe and characterize this distortion.

In this preferred specification, in panel (b) we find that unprocessed foods are the most elastic; in other words, if household incomes increase, then unprocessed food demand rises fastest, followed by highly processed foods, and then low processed foods. Between highly processed foods consumed at home and those consumed away from home, the foods consumed away from home are more elastic. This finding corresponds with the aggregate trends observed in section 2, with increasing consumption of food away from home but decreases in processed foods prepared and consumed at home. However, we qualify this observation by emphasizing that it is conditional on already consuming some food away from home and omits variation on the extensive margin. In the next section, we further explore heterogeneity in this particular result.

3.3 Food Away from Home/(In)stability of Elasticities

Given the rapidly rising levels and budget shares of food away from home documented in section 2, we focus on this food group to better understand variation in these trends within Nigeria. Using the $\log(Z)$ form, we next study elasticities of food away from home using subsamples of the data. Although this specification invites issues of selection discussed in the previous section, it is both consistent with consumer theory and does not invite the inconsistency in rankings that we observe when making arbitrary adjustments to include observations without food away from home.[13]

Table 4 shows the estimated elasticities of food away from home on overall food value. Elasticities using the scaled IHST form can be found in Appendix Table A.1. Each row indicates the sample used for the regressions while each column indicates the specification used for the estimation, defined further at the bottom of the column. Each cell is the estimated elasticity from a separate regression. Column (1) includes household size, survey round, and regional fixed effects. Column (2) replaces survey round and region fixed ef-

[13]Moreover, one might argue that the elasticity when going from zero to positive consumption is meaningless as the percent change is infinite.

fects with survey round by region fixed effects, allowing for different time trends across these areas. Column (3) uses finer, survey area fixed effects. Column (4) switches from survey area fixed effects to household fixed effects. And finally, Column (5) uses local survey area by survey round fixed effects along with household fixed effects. The elasticities in Column (2) use the same specification as in Panel (a) of Table 3 whereas Column (5) uses the same functional form as that in Panel (b).

Table 4: Food Away from Home Elasticities from Log(Z) Estimation

Sample:	Obs	(1)	(2)	(3)	(4)	(5)
All	17584	0.952	0.946	0.898	0.812	0.868
		(0.013)	(0.013)	(0.014)	(0.015)	(0.018)
Rural North	7450	0.837	0.837	0.789	0.749	0.828
		(0.018)	(0.018)	(0.02)	(0.023)	(0.026)
Rural South	4208	0.927	0.927	0.822	0.682	0.775
		(0.028)	(0.028)	(0.03)	(0.032)	(0.038)
Urban North	2119	1.092	1.092	1.045	0.954	0.906
		(0.038)	(0.038)	(0.041)	(0.048)	(0.051)
Urban South	3800	1.122	1.122	1.139	1.032	1.072
		(0.03)	(0.03)	(0.03)	(0.033)	(0.039)
Poor	8218	0.836	0.827	0.807	0.789	0.839
		(0.021)	(0.021)	(0.021)	(0.023)	(0.029)
Non-Poor	9366	0.909	0.905	0.905	0.833	0.923
		(0.019)	(0.02)	(0.02)	(0.021)	(0.025)
Scarcity	8190	1.086	1.085	1.063	0.982	1.016
		(0.021)	(0.021)	(0.024)	(0.036)	(0.048)
Abundance	9394	0.892	0.886	0.814	0.623	0.687
		(0.021)	(0.021)	(0.023)	(0.031)	(0.04)
Household Size FEs		Yes	Yes	Yes	Yes	Yes
Survey Round FEs		Yes	No	No	No	No
Region FEs		Yes	No	No	No	No
Region x Srv Rd FEs		No	Yes	Yes	Yes	No
Local Area FEs		No	No	Yes	No	No
Local Area x Srv Rd FEs		No	No	No	No	Yes
Household FEs		No	No	No	Yes	Yes

Notes: Standard errors clustered at the primary sampling unit level in parentheses. Note that sample includes multiple observations of the same set of 5,000 original households and, in this specification with the log transformation, only include respondents with positive consumption values. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

The results in Table 4 demonstrate two important points. First, we find that as we control for more and more variation from columns (1) to (5), the elasticity estimates typically fall. For example, among all households we estimate an elasticity of 0.952 in column (1) that falls to 0.868 in column (5). Therefore, as above when we are better able to control for more variation in the data before estimating the relationship between food expenditures away from home and total food expenditures, we find that elasticity estimates are not as high. This pattern cautions against over-interpreting results from single cross-sections.

Second, the results highlight considerable heterogeneity that is relatively consistent across specifications. Food away from home appears considerably more elastic in urban areas and in the South. Food away from home is consistently elastic in the urban South. While the different specification choices shift the levels of these estimated elasticities, there is an economically meaningful gap in these elasticities by regions. Splitting the sample by households above and below the median of overall food expenditures, we see that those above the median, labelled as non-poor, have systematically higher elasticities than those below the median, labelled as poor. We do also show these results using the IHST specification in Appendix Table A.1. Notably, by using the IHST transformation and including the extensive margin in the estimation, food away from home is elastic across all groups. However, the heterogeneity between the North and the South disappears once household fixed effects have been incorporated.

3.4 Within Household Variation

While cross-sectional heterogeneity has considerable policy relevance for understanding areas and populations with more rapidly changing diets, within household variation may also impact consumption choices on food away from home. Summary statistics in section 2 suggested that households experience considerable season to season variability in food consumption value levels. The average household had a gap of 1.126 times their mean level of consumption between their period of highest and lowest consumption. The full distribution of these fluctuations are shown in Appendix Figure A.1, demonstrating that these fluctuations are large, even when looking only within year (by season) or across years (controlling for season).

If households are able to perfectly smooth consumption, we would assume relatively stable demand elasticities for food consumption, independent of fluctuations in income. However, without perfect smoothing, households are likely to adjust their food expenditure allocations in response to tighter or looser budgetary constraints. In using the Cobb-Douglas framework in our estimating equations, recall that the elasticity is assumed to be constant for all food expenditure levels. It does not, therefore, shed any light on whether there are asymmetric responses to income fluctuations. If households prioritize the consumption of different types of foods at times of relative scarcity or abundance, the Cobb-Douglas framework will miss it. To explore this type of heterogeneity further, we rank the six observations for each household and split them at the median, labeling the top three periods, times of relative abundance within the household, and the bottom three as times of relative scarcity. We then re-estimate the elasticities using those two subsamples. Results are shown in Table 4, rows (8) and (9).

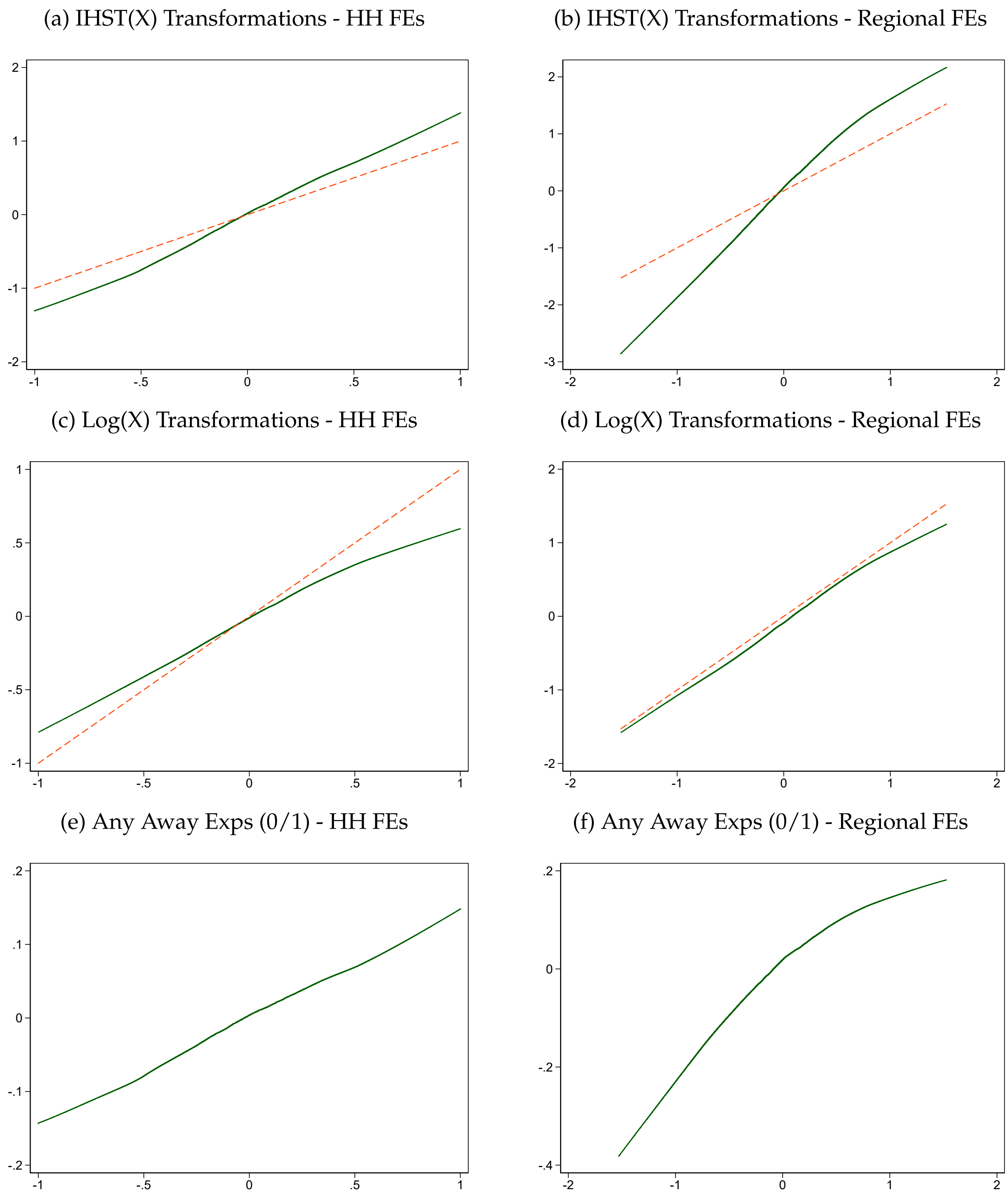
Perhaps surprisingly, we estimate higher elasticities in times of scarcity as in times of abundance across all specifications. In other words, when households' overall level of food expenditures go from average to constrained, they appear to economize by aggressively cutting out food away from home. When expenditures increase from the household median, additional food expenditures are less likely to be allocated towards food away from home. This result is particularly notable because it demonstrates the opposite dimension of heterogeneity as we saw in poor versus non-poor households. suggesting that *within* household fluctuations in income may explain more variation in observed consumption of food away from home than average, cross-sectional levels of income.

Splitting the sample provided an initial indication that a single average elasticity may be problematic. To explore this further we show semi-parametric relationships between overall food consumption value and that of food away from home. With semi-parametric estimation, we relax the Cobb-Douglas assumption of a constant relationship between the share of expenditures on food away from home and total food expenditures and allow the slope of the relationship to vary. To do so, we start by separately regressing our independent variable, the logarithm of total food expenditures, on our two preferred specifications using cross-sectional fixed effects and using household fixed effects. We repeat this estimation on three different forms of food away from home: 1) the IHST form with a 100 scaling factor, 2) the log form, with no adjustment, and 3) the binary outcome of any food away from home consumed. Finally, we take these two pairs of residuals and plot their relationship using a non-parametric kernel-weighted local polynomial smoothing algorithm as shown in Figure 5.

The panels on the left use household fixed effects whereas those on the right use only region by survey round fixed effects. First, we see confirmation of the regression results in both logarithmic and IHST forms and for both specifications. We find that elasticities are higher– or steeper in the graphs– when overall consumption levels are below the mean. This finding suggests that food away from home is likely to be cut out of household consumption more rapidly than other foods when overall consumption levels fall, faster than they are added in times of relative abundance.

Second, comparing columns, we see that the use of household fixed effects leads to lower estimated elasticities. Again, this is consistent with omitted geographic variables biasing our initial elasticity estimates upwards. And finally, we note that the IHST form is consistently steeper than that of the basic log transformation. The bottom two panels provide the explanation: There is substantial activity on the extensive margin where more households purchase food away from home in times of relative abundance, meaning that households are more likely to begin to purchase food away from home at times of abundance. The log transformation, by dropping the extensive margin, is therefore underestimating the actual elasticities.

Figure 5: Non-Parametric Regressions of Away Expenditures on Total Food Value



Notes: Dotted line shows a slope of 1 for unit elastic. Residuals of both dependent and independent variable are taken and the green lines show semi-parametric estimates of the value of food consumed away from home on on overall value of food consumption. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

4 Discussion and Conclusion

In this paper, we have used the Nigerian LSMS surveys to study the way that demand for food, by processing level, have changed within a set of households over time. First, we find that using the definition of highly processed food that is standard in the literature, demand for highly processed foods has not increased over the period of this study in real terms. Furthermore, this defintion masks a *decrease* in the value of highly processed foods being eaten at home, and an increase in the food being eaten away from home. For this latter category, most surveys do not capture the level of processing undergone by these meals; there is a clear need for better data collection on the types of food away from home that are eaten in developing countries, so that a better understanding of their growing role in people's diets can be better understood. Nonetheless, assuming all food away from home is highly processed leads to an over-estimate of consumption of highly processed foods and, in this case, mis-characterizes the trends taking place over time.

Second, we explore food demand elasticities of food groups by their level of processing. We find that our conclusions of relative elasticities depend on the methodology used to estimate this relationship. Different specifications result in elasticities where different different categories of food are most elastic. Therefore, we argue that it is preferable to use functional forms for demand that can be derived from a preference relation. Ultimately, issues related to functional form do not address other important issues, such as the endogeneity of overall food expenditure allocations with other unobservable factors. To deal with endogeneity, it might be interesting to re-analyze data from recent randomized trials on cash transfer programs for the composition of food expenditures by level of processing, and purchases of food away from home, as we have done here. Consequently, our results should still be considered with some caution, as the specification used may not fully represent the data generating process either.

Third, we find that as incomes increase, in our preferred specification, unprocessed foods are most elastic, followed by highly processed foods. This conclusion is the opposite of Tschirley et al. (2015); when we use a specification that controls neither for fixed household preferences over time nor cluster-time fixed effects for prices, our conclusion about the relative demand would change. Such results should caution against putting too much stock in estimates of elasticities that do not control for these types of variation, as they can lead to conclusions biased in unknown directions. Controlling for household-level fixed effects and for changing prices and dynamic factors at the cluster level implies that the remaining threats that could bias our estimates would have to be unobservable dynamic changes at the household level.

Somewhat surprisingly, we show that household-level income varies substantially in Nigeria across surveys, with average variation in expenditures exceeding the average household expenditures over the study period. We take this variation as an opportunity, and measure how demand for food away from home changes when household food expenditures are above and below median. We find that demand is more elastic when expenditures are below average than above average. We interpret this to mean that in times of hardship, households cut food away from home out of their diets more rapidly than they increase this expenditure share in times of relative abundance.

Finally, we find that demand for food away from home is more elastic in urban areas than rural areas and among those with higher incomes. To the extent that growing demand for food away from home could be leading to unhealthy outcomes, these results combine to suggest that policy weight should be placed on trying to ensure that restaurants or street food vendors offer foods with healthier ingredients, and further weight be placed on prioritizing changes in the South.

References

Alston, J.M., J.P. MacEwen, and A. M. Okrent (2016) "The Economics of Obesity and Related Policy," *Annual Review of Resource Economics*, Vol. 8, pp. 443–465.

Bahar, Dany and Hillel Rapoport (2018) "Migration, Knowledge Diffusion, and the Comparative Advantage of Nations," *Economic Journal*, Vol. 128(July), pp. F273–F305.

Bellemare, Marc and Casey Wichman (2018) "Elasticities and the Inverse Hyperbolic Sine Transformation," *Working Paper, University of Minnesota*, pp. 1–20.

Bouis, Howarth and Lawrence Haddad (1992) "Are Estimates of Calorie-Income Elasticities Too High? A Recalibration of the Plausible Range," *Journal of Development Economics*, Vol. 39, pp. 333–364.

Butta, Zulfiqar A. and Rehana A. Salam (2012) "Global Nutrition Epidemeology and Trends," *Annals of Nutrition & Metabolism*, Vol. 61(suppl 1), pp. 19–27.

Clemens, Michael and Erwin R. Tiongson (2017) "Split Decisions: Household Finance when a Policy Discontinuity Allocates Overseas Work," *Review of Economics and Statistics*, Vol. 99(3), pp. 531–543.

Cockx, Lara, Liesbeth Colen, and Joachim De Weerdt (2018) "From Corn to Popcorn? Urbanization and Dietary Change: Evidence from Rural-Urban Migrants in Tanzania," *World Development*, Vol. 110, pp. 140–159.

Deaton, Angus and Christina Paxson (1998) "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, Vol. 106, pp. 897–930.

Deaton, Angus and Salman Zaidi (2002) "Guidelines for Constructing Consumption Aggregates for Welfare Analysis," *World Bank LSMS Working Paper*.

Food and Agriculture Organization (2018) *The State of Food Security and Nutrition in the World*: Rome:FAO.

International Food Policy Research Institute (2016) *Global Nutrition Report 2016: From Promise to Impact: Ending Malnutrition by 2030.*: Washington, DC: International Food Policy Research Institute.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas (2017) "Cash for Carbon: A Randomized Trial of Payments for Ecosystem Services to Reduce Deforestation," *Science*, Vol. 21(6348), pp. 267–273.

McKenzie, David (2017) "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition," *American Economic Review*, Vol. 107(8), pp. 2278–2307.

Monteiro, Carlos Augusto (2009) "Nutrition and health. The issue is not food, nor nutrients, so much as processing," *Public Health Nutrition*, Vol. 12, pp. 729–731.

Monteiro, Carlos Augusto, Renata Bertazzi Levy, Rafael Moreira Claro, Ines Rugani Ribeiro de Castro, and Cannon Geoffrey (2010) "A New Classification of Foods Based on the Extent and Purpose of Their Processing," *Saude Publica*, Vol. 26(11), pp. 2039–2049.

Pingali, Prabhu (2015) "Agricultural Policy and nutrition Outcomes: Getting Beyond the Preoccupation with Staple Grains," *Food Security*, Vol. 7(3), pp. 583–591.

Popkin, Barry, Linda Adair, and Shu Wen Ng (2012) "Global Nutrition Transition and the Pandemic of Obesity in Developing Countries," *Nutrition Reviews*, Vol. 70(1), pp. 3–21.

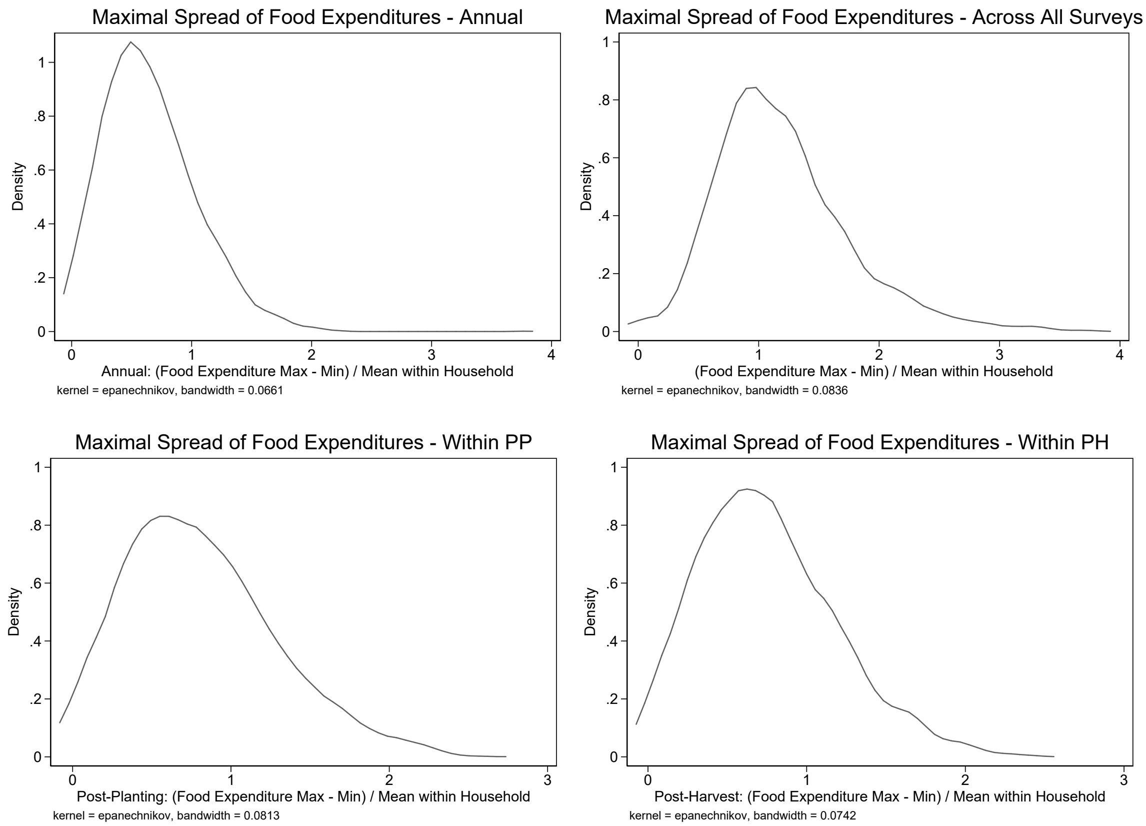
Reardon, Thomas, C. Peter Timmer, Christopher B. Barrett, and Julio Berdegué (2013) "The Rise of Supermarkets in Africa, Asia, and Latin America," *American Journal of Agricultural Economics*, Vol. 85(5), pp. 1140–1146.

Skoufias, Emmanuel, Sailesh Tiwari, and Hassan Zaman (2012) "Chrises, Food Prices, and the Income Elasticity of Micronutrients: Estimates from Indonesia," *World Bank Economic Review*, Vol. 26, pp. 415–442.

Tschirley, David, Thomas Reardon, Michael Dolislager, and Jason Snyder (2015) "The Rise of a Middle Class in East and Southern Africa: Implications for Food System Transformation," *Journal of International Development*, Vol. 27, pp. 628–646.

Appendix A: Additional Tables and Figures

Figure A.1: Within Household Food Expenditure Variabilty



Notes: Data included up to six total interviews per household. Three different years, each with two seasonal visits: post-harvest (PH), and post-planting (PP). Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

Table A.1: Total Food Value Elasticities from IHST(100 x Z) Estimation

Sample:	N	(1)	(2)	(3)	(4)	(5)
All	24323	1.654	1.659	1.524	1.336	1.324
		(0.027)	(0.027)	(0.028)	(0.031)	(0.04)
Rural North	10325	1.507	1.507	1.4	1.287	1.313
		(0.041)	(0.041)	(0.042)	(0.046)	(0.04)
Rural South	6526	1.637	1.637	1.438	1.153	1.201
		(0.053)	(0.053)	(0.056)	(0.06)	(0.04)
Urban North	2563	1.773	1.773	1.652	1.608	1.422
		(0.082)	(0.082)	(0.085)	(0.101)	(0.04)
Urban South	4902	1.932	1.932	1.838	1.592	1.528
		(0.06)	(0.06)	(0.061)	(0.068)	(0.04)
Poor	12212	1.452	1.458	1.349	1.3	1.241
		(0.042)	(0.042)	(0.042)	(0.044)	(0.04)
Non-Poor	12110	1.579	1.582	1.512	1.358	1.426
		(0.042)	(0.042)	(0.042)	(0.043)	(0.04)
Scarcity	12027	1.922	1.931	1.816	1.477	1.522
		(0.042)	(0.042)	(0.047)	(0.066)	(0.04)
Abundance	12296	1.564	1.566	1.395	1.033	1.034
		(0.046)	(0.046)	(0.049)	(0.064)	(0.04)
Household Size FEs		Yes	Yes	Yes	Yes	Yes
Survey Round FEs		Yes	No	No	No	No
Region FEs		Yes	No	No	No	No
Region Srv Rd FEs		No	Yes	Yes	Yes	No
Local Area FEs		No	No	Yes	No	No
Local Area x Srv Rd FEs		No	No	No	No	Yes
Household FEs		No	No	No	Yes	Yes

Notes: Standard errors clustered at the primary sampling unit level in parentheses. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

Table A.2: Categorization of Food Items from Nigeria LSMS

Food	**Perishability**	**Processing**	**Food Category**
Guinea corn/sorghum	Not Perishable	None	Grain
Millet	Not Perishable	None	Grain
Maize	Not Perishable	None	Grain
Rice-local	Not Perishable	Low	Grain
Rice-imported	Not Perishable	Low	Grain
Maize flour	Not Perishable	Low	Grain
Yam flour	Not Perishable	Low	Grain
Cassava flour	Not Perishable	Low	Grain
Wheat flour	Not Perishable	Low	Grain
Maize (unshelled/on the cob)	Not Perishable	None	Grain
Maize (shelled)	Not Perishable	None	Other
Maize (shelled/on the cob)	Not Perishable	None	Grain
Maize (shelled/off the cob)	Not Perishable	None	Grain
Other grains and flour	Not Perishable	None	Grain
Bread	Perishable	High	Other
Cake	Perishable	High	Other
Buns/Pofpof/Donuts	Perishable	High	Other
Biscuits	Perishable	High	Other
Meat pie/Sausage roll	Perishable	High	Other
Cassava-roots	Perishable	None	Grain
Yam-roots	Perishable	None	Grain
Gari - white	Perishable	Low	Grain
Gari - yellow	Perishable	Low	Grain
Cocoyam	Perishable	None	Grain
Plantains	Perishable	None	Grain
Sweet potatoes	Perishable	None	Grain
Potatoes	Perishable	None	Grain
Other roots and tubers	Perishable	None	Grain
Soybeans	Not Perishable	None	Pulse
Brown beans	Not Perishable	None	Pulse
White beans	Not Perishable	None	Pulse
Groundnuts	Not Perishable	None	Pulse
Groundnuts (unshelled)	Not Perishable	None	Pulse
Groundnuts (shelled)	Not Perishable	None	Pulse
Other nuts/seeds/pulses	Not Perishable	None	Pulse
Coconut	Not Perishable	None	Fruit
Kola nut	Not Perishable	None	Pulse
Cashew nut	Not Perishable	None	Pulse
Palm oil	Not Perishable	High	Fat
Butter/margarine	Perishable	High	Dairy
Groundnuts oil	Not Perishable	High	Fat
Other oil and fat	Not Perishable	High	Fat
Sheabutter	Not Perishable	High	Fat
Coconut oil	Not Perishable	High	Fat
Animal fat	Perishable	High	Fat
Bananas	Perishable	None	Fruit
Orange/tangerine	Perishable	None	Fruit
Mangoes	Perishable	None	Fruit
Avocado pear	Perishable	None	Fruit
Pineapples	Perishable	None	Fruit

Fruit canned	Perishable	Low	Fruit
Other fruits	Perishable	None	Fruit
Pawpaw	Perishable	None	Fruit
Watermelon	Perishable	None	Fruit
Apples	Perishable	None	Fruit
Tomatoes	Perishable	None	Veg
Tomato puree (canned)	Perishable	Low	Veg
Onions	Perishable	None	Veg
Garden eggs/eggplant	Perishable	None	Veg
Okra-fresh	Perishable	None	Veg
Okra-dried	Perishable	None	Veg
Pepper	Perishable	None	Veg
Fresh pepper	Perishable	None	Veg
Dry pepper	Perishable	None	Veg
Leaves (cocoyam, spinach, etc.)	Perishable	None	Veg
Other vegetables (fresh and canned)	Perishable	None	Veg
Chicken	Perishable	None	Meat
Duck	Perishable	None	Meat
Other domestic poultry	Perishable	None	Meat
Agricultural eggs	Perishable	None	Dairy
Local eggs	Perishable	None	Dairy
Other eggs (not chicken)	Perishable	None	Dairy
Beef	Perishable	None	Meat
Mutton	Perishable	None	Meat
Pork	Perishable	None	Meat
Goat	Perishable	None	Meat
Wild game/bush meat	Perishable	None	Meat
Canned beef/corned beef	Perishable	None	Meat
Other meat (excl. poultry)	Perishable	None	Meat
Fish - fresh	Perishable	None	Meat
Fish - frozen	Perishable	Low	Meat
Fish - smoked	Perishable	Low	Meat
Fish - dried	Perishable	Low	Meat
Snails	Perishable	None	Meat
Seafood (lobster, crab, prawns, etc.)	Perishable	None	Meat
Canned fish/seafood	Not Perishable	High	Meat
Other fish or seafood	Perishable	None	Meat
Fresh milk	Perishable	High	Dairy
Milk powder	Not Perishable	High	Dairy
Baby milk powder	Not Perishable	High	Dairy
Milk tinned (unsweetened)	Not Perishable	High	Dairy
Cheese (wara)	Perishable	High	Dairy
Other milk products	Perishable	High	Dairy
Coffee	Not Perishable	High	Other
Chocolate drinks (incl. Milo)	Not Perishable	High	Other
Tea	Not Perishable	High	Other
Sugar	Not Perishable	High	Other
Jams	Perishable	High	Other
Honey	Not Perishable	Low	Other
Other sweets and confections	Not Perishable	High	Other
Condiments (salt,spices,pepper,etc.)	Not Perishable	None	Other
Salt	Not Perishable	Low	Other
Unground Ogbono	Not Perishable	Low	Pulse
Ground Ogbono	Not Perishable	Low	Pulse

Ground pepper	Not Perishable	Low	Other
Melon (shelled)	Not Perishable	Low	Fruit
Melon (unshelled)	Not Perishable	Low	Fruit
Melon (ground)	Not Perishable	Low	Fruit
Bottled water	Not Perishable	Low	Other
Sachet water	Not Perishable	Low	Other
Malt drinks	Not Perishable	High	Other
Soft drinks (Coca Cola, spirit, etc.)	Not Perishable	High	Other
Fruit juice canned/pack	Not Perishable	High	Other
Other non-alcoholic drinks	Not Perishable	High	Other
Beer (local and imported)	Not Perishable	High	Other
Palm wine	Not Perishable	High	Other
Pito	Not Perishable	High	Other
Gin	Not Perishable	High	Other
Other alcoholic beverages	Not Perishable	High	Other
Guava	Perishable	None	Fruit

Appendix B: Inverse Hyperobolic Sine Transformation Instability

As discussed in Section 3, the inverse hyperbolic sine transformation (IHST) has been increasingly used in empirical economics as a way to address issues that arise when measuring elasticities of variables that contain zeros. A recent working paper by Bellemare and Wichman (2018) demonstrate how to calculate elasticities from regressions using IHST forms of variables. The authors illustrate that IHST transformations and logarithmic transformations perform well when values are "high" (as a rule of thumb, they suggest that values above ten are probably only minimally affected). They then make two potential suggestions.

First, a mathematical adjustment can be made to the estimated coefficients, taken at their mean values such that the elasticity of two variables can be found from an IHST-IHST regression:

$$\hat{\xi}_{yx} = \hat{\beta} \times \frac{\sqrt{y^2+1}}{y} \times \frac{x}{\sqrt{x^2+1}} \tag{4}$$

An alternative adjustment acknowledges that because $lim_{y\to\infty} = 1$ for the second term on the right hand side, and $lim_{x\to\infty} = 1$ for the third term, with large values of x and y the estimate for $\hat{\xi}$ of an IHST-IHST form regression will converge to β, or the elasticity from a log-log functional form. The suggestion is therefore to rescale the x and y variables by a sufficiently large factor to facilitate this convergence.

However, the motivation for using IHST adjustments in place of logs is to address issues of selection that emerge when there are values of zero in your dependent variable. To explore this further, we estimate a number of elasticities using a range of different specifications both with and when excluding observations with values of zero for the dependent variable. The independent variable, total food expenditures, always takes a positive value in this data.

Panel (a) uses the food away from home category from the main paper analysis, which has a relatively low positive incidence of 0.723, while Panel (b) focuses on a food category with relatively high incidence of 0.963. For each column, the same transformation is made on both the dependent and independent variables. Focusing on Panel (a), the first column uses the regular logarithmic transformation, allowing zeros to be dropped from the estimation. The next two columns show two conventional adjustments frequently used, adding 1 and adding 0.01 to x before taking the logarithm of the variable. This seemingly arbitrary choice has large consequences for our elasticity, resulting in a highly inelastic measure of 0.652 and highly elastic measure of 1.532 respectively. Column (4) uses IHST along with the first suggested adjustment from Bellemare and Wichman (2018). Columns (5)-(8) show results from the IHST form regressions with a range of different scaling factors. And finally, Columns (9) - (12) show this progression of IHST forms estimated exclusively on positive valued data, to remove selection issues in a comparison of these elasticities with the "correct" elasticities estimated in column (1).

Although the adjustment in column (9) does not closely match that estimated in Column (1), we may not expect them to be the same, since the average percent change in the dependent variable with respect to the independent variable may not be the same as at the mean of Z. What we do see is that scaling up the IHST transformation in columns (10)-(12) appears to work. Higher scaling factors converge to the values in Column (1).

What is striking and concerning is that columns (5)-(8) do not show any signs of convergence to an elasticity we would reasonably expect from the data. Higher and higher scaling factors are effectively giving increasing weight to the binary switches to and from zero. Comparing Panel (a) and Panel (b) we see that this problem is more severe for food away from home, where the incidence of zeros in the data is higher. Given that usage of IHST has been adopted precisely to address issues that arise from having zero values in the data, the problems that come with its use may undo the benefits of retaining zeros.

Table B.1: Elasticity Instability with Food Away and Low Processed Perishable Foods

Panel (a) Food Away: Positive Incidence = .723

									Positive Values Only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log(Z)	Log(1+Z)	Log(.01+Z)	IH(Z)-adj	IH(Z)	IH(100xZ)	IH(1000xZ)	IH(10000xZ)	IH(Z)-adj	IH(Z)	IH(100xZ)	IH(1000xZ)
Food Exps	0.946	0.652	1.532	0.814	0.758	1.659	2.077	2.495	0.794	0.765	0.945	0.946
	(0.013)	(0.008)	(0.024)		(0.009)	(0.027)	(0.037)	(0.047)		(0.010)	(0.013)	(0.013)
N	17575	24313	24313	24313	24313	24313	24313	24313	17575	17575	17575	17575
r2	0.343	0.289	0.200	0.285	0.285	0.191	0.170	0.158	0.366	0.366	0.343	0.343
Positive Only	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes

Panel (b) Low Processed, Perishable Foods: Positive Incidence = .963

									Positive Values Only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log(Z)	Log(1+Z)	Log(.01+Z)	IH(Z)-adj	IH(Z)	IH(100xZ)	IH(1000xZ)	IH(10000xZ)	IH(Z)-adj	IH(Z)	IH(100xZ)	IH(1000xZ)
Food Exps	1.150	0.940	1.281	1.056	1.045	1.310	1.390	1.470	1.04	1.029	1.149	1.150
	(0.007)	(0.005)	(0.009)		(0.006)	(0.010)	(0.012)	(0.015)		(0.006)	(0.007)	(0.007)
N	24063	24428	24428	24428	24428	24428	24428	24428	24063	24063	24063	24063
r2	0.580	0.611	0.500	0.602	0.602	0.479	0.408	0.347	0.606	0.606	0.582	0.581
Positive Only	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the primary sampling unit level in parentheses. Both dependent and independent variables adjusted using the transfrmation listed at the top of each column. Data source: Nigeria Living Standards Measurement Survey (LSMS) - 2011, 2014, and 2016.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org