# Status in sub-Saharan Africa

*Quality and lifestyle amidst urbanisation*

## 1 Introduction

Models for demand analysis rely on exogenous price observations for the consumer's basket. The prices observed in the market embody how the market, subject to supply pressures both i) addresses the needs of the consumers and ii) prices the differences in quality variants of the same commodity to the consumers. Despite market-observed prices being of a central importance to the demand analysis, it is only recently that the detailed prices observed in the market are made available in micro-data surveys. Arguing that market prices are key to an understanding of quality, we enhance a typical demand analysis to include the quality variation within commodities, the needs faced by the consumer and other household characteristics such as income and social-economic standing.

We focus on food quality to understand the choice the consumers face in an environment of rising prices, incomes and urban development. The quality in food is made relevant in the developing countries due to most of the population living in a subsistence mode and spending a significant portion of their income on food. Further, since not all consumers may face the same basket of goods due to uneven rates of economic development, the quality the consumer chooses is also subject to the long-term choices she has made (e.g. residence in a region, education, occupation etc.). The demand for quality in food is thus also likely to be segmented due to severe inequalities in varied circumstances of the population and low possibilities of a transition in the consumers' longer-term circumstances (compared to the developed countries). The market-prices of the variants of a commodity in the various regions of economy are thus a prerequisite for such an analysis of quality.

We find the segmentation of quality relevant from the point of view of status consumption as well - as status consumption is often equivalent to high-quality consumption. This is because i) status-benefit from an item is sure to be priced in the market and ii) high relative-quality of an item is likely to carry some status benefit. The consumer choice for quality given her longer-term circumstances (both encapsulated within the household characteristics related to her needs and long-term economic standing) thus provides us a way to explain conspicuous consumption in terms of status-benefits from quality and longer-term economic standing. In the analysis that follows, we employ methods to measure the needs, the socio-economic standing and the consumer choice for quality. While the needs of the consumers are proxied by the number of members (age-adjusted) in the household (and other available scales), quality is measured with the quantity consumed of high vs low-price variants of the commodity. With the AIDS formulation [1] for this model of quality, the scale of consumer needs appears as a control variable in the AIDS budget share equations whereas the quality-choices appear as dependent variables in the equations (along with the budget shares for every commodity). This is the so-called "Unrestricted Method" used by McKelvey [2] and Gibson et al. [3].

A consumer can improve quality in food or in other commodity groups as her income rises. The improvement and sustenance of quality across multiple commodity groups does constitute status-consumption - since consumers of high-quality in all groups are more likely to carry status than those who don't. This is ever more likely in the context of developing countries where differences in quality are wider amongst the consumers. That said, we distinguish the consumption of quality which we have discussed above from what is termed as conspicuous consumption. At least semantically speaking, the term conspicuous consumption encompasses only the exclusive consumption by the rich. Viewing the rapid economic development in Africa, one is tempted to draw parallels to Veblen's conspicuous consumption of the 19th century North America - noting how an economy can transition from a vastly agrarian society to an industrialised one. However, from a policy point of view, there are issues with extending the concepts identified by Veblen[4] into the developing markets without considering how deep and widespread the income and social inequalities in the developing economy (see Ustuner et al.[5] for problems faced by extending Veblen consumption in the developing economies). In the environment of nascent economic growth which we observe in the sub-Saharan Africa, the options for conspicuous consumption are limited to a very small part of the population. Also, the social fragmentation within the sub-Saharan Africa (see the historical note in the Appendix I) means that the social context which the market forces operate within are not the same as that of North America in its gilded age. Instead of focusing on demand for what may be considered visible items (which may not be consumed for a majority of the population) in the developing countries, we believe that observing the quality in consumption and possession of long-term assets - would provide a better picture of the status-related consumption - the latter being a concern for a much larger section of the population.

Our focus on quality is also motivated by a need to reduce subjective judgments from the classification of visible items. Status - which motivates conspicuous consumption - is a complex social psychological phenomenon and thus one cannot rule out a subjective judgment in deciding whether an item is a status good or not. For example, if one were to rely only on whether an item can be afforded by the rich - then electricity, being used by a small and rich part of the population in a sub-Saharan African country like Tanzania, may appear to be a status good - at par with weddings (which are a clearer way of indicating status). But whether the same amount of money spent on weddings enhances status more than when it's spent on electricity is a far more difficult question to answer. Similarly, one could argue that buying clothes indicates higher status and record consumption on clothes as conspicuous consumption. But without observing the effect that getting electricity in the house or moving to a bigger house may have on a consumer's long-term utility and the perceived status, we cannot assume a significant status-benefit to the consumer from expenditure on clothing - since the possibility that those spending on recently installed electricity or a newly built house may actually be cutting down on clothing expenditure while deriving status from electricity or house ownership cannot be ruled out. To solve these fundamental issues, the literature on conspicuous consumption has often proceeded with a snapshot consumer poll that obtains the appropriate classification directly from the consumer. However, drawing inferences on long-term consumer behaviour from such analyses still remains a challenge due to the cross-sectional nature of such surveys. Further, status goods are indeed hardly static - as the demand for newly introduced items - e.g. a smart-phone or a TV - changes quickly. The types of status items one may wish to model (e.g. snob / bandwagon depending on whether the consumer wants to separate herself from others with the item or catches-up with those around her) are subject to change as the consumer incomes increase over time and the markets evolve to address consumer needs by launching cheaper substitutes etc.. It is to provide a dynamic model which encompasses such classifications that we view status as part of the consumer's long term utility - and attempt to explain it using the concepts from the literature on quality. Such a model would also allow us to refine the concepts of conspicuous consumption that may be used to compare the consumer behaviour across the developing and developed markets.

If status can be seen as part of the consumer's long-term expected utility, it follows that the status that a consumer derives from consumption must come out of either elite non-durable consumption or the long-term socio-economic stnading. Thus, tracking quality over time (i.e. high-quality consumption) against the socio-economic standing including asset ownership can help us understand the demand for status. Instead of attempting to find the criteria for separating status-items from bare-necessities, we view status consumption as the consumer's short-term needs (i.e. quality or expensive non-durable consumption - henceforth referred to as ENDC) as opposed to her long-term needs that are fulfilled with the longer term socio-economic standing (e.g. with skills, occupation, family wealth etc.).

The econometric methods we suggest are meant to achieve two goals. First is to explain the differences in quality across the population - given a certain household characteristics. The second is to incorporate the discrete choices that a consumer faces for sparsely available or costlier items (such as electricity in Tanzania or urban housing). In the current chapter, we attempt to address only the former by enhancing the AIDS formulation. The latter goal - which can be seen as an instance of intertemporal substitution between improving quality and acquiring costlier long-term status - is to be addressed in a subsequent chapter. We believe that the econometric methods developed in the current chapter are generic enough to be applied to surveys from other economies. More generally, the reason why we set ourselves to explain long-standing inequalities or regional disparities in terms of gaps in incomes, available quality and asset ownership is to be able to translate the wide-ranging social observations on conspicuous consumption in the literature onto empirical measurements that may work under disparate social contexts.

The AIDS methods used in the current chapter take quality and socio-economic standing as inputs. For the measurement of quality, we take a price-based view. As quality is priced in the market, the quality for a commodity is indicated by the prices of the variants of a commodity. In the Tanzanian data which we have used, therefore, eggs and milk would be deemed of lower quality than goat and beef for proteins. Similarly for transport, bicycle would be of lower quality than car and so on. To measure the socio-economic standing of a consumer[1], we consider a vector of income, education and region/locality/neighborhood as well as the long-term durable goods possessed by the consumer. This is inspired by the use of socio-economic status (SES) indicators in the literature on health and psychology[6].

Section 2 surveys the literature on quality of consumption and the relevance of often overlooked market prices. A selection equation approach for the conventional

[1] The terms social capital - or status capital are also used in the literature [5]

AIDS formulation and an approach to view aggregate quality are presented in Section 3. The details of the data used in the study are provided in Section 4. A discussion of the distribution of quality for food and non-food items follows in Section 5.

## 2 Literature Survey

Although not as the main focus of demand studies, quality elasticities have been derived from consumption data since at least as far back as Houthakker [7]'s post-war analysis of family surveys. In a context of the cost-of-living index problem, the later "simple-repackaging" method used by Fisher-Shell [8] compares higher-quality consumption with the lower-quality equivalent - carrying forward the notion of equivalence scales used by Houthakker. In the simple repackaging method, with the quality for a (composite) good parametrised as $\theta$ and its base-price $p_1^1$, an adjusted-price $p_1^*$ is sought so that:

$$x(p_1^*, p_2, ...) \equiv x(p_1^1, p_2, ..., b) \tag{1}$$

The base-price is chosen so that one can view the demand for items with price $p_1^2$ and $p_2^3$ in terms of the base-price $p_1^1$ and the quality parameter $b_1$ (where $b = \{b_1, b_2, ...\}$). Here, an implicit assumption is that the variants in every category are perfect substitutes. With that assumption, setting $\theta$ to be the quality indicator and $h(\theta)$ a function of the quality parameter, one can write the following in terms of a cost function:

$$\tilde{c}(u, p_1, p_2, ..., \theta) = c(u, \frac{p_1}{h(\theta)}, p_2, ..., 1)$$

As the idea of simple repackaging requires perfect substitution, its applicability to all commodity groups is somewhat limited. Consider for example, a consumer switching from kerosene usage to electricity usage - which is far more than an improvement of quality in "energy" and makes the expression of electricity consumption in terms of the quantity and price of kerosene consumed somewhat of a stretch. To proceed, one may consider how using electricity would change the utility for food (with refrigerator), household consumption (hair-dryers) etc.. The classical approach taken by Fisher and Shell [8] recommends adjusting the prices of the final goods i.e. by decomposing utilities in a way that quality-change can be traced to all composite goods whose quality is affected by the particular commodity (e.g. electricity). This may sound rather cumbersome, but the issue can in fact be mitigated if there were a recursive tree-like hierarchy of the items such as refrigerators that improve the quality of other goods. With a tree-like structure, one needs to only look at the nodes of the trees i.e. adjust the price of all food items whose quality is improved. In addition to adjusting the price for energy, therefore one would also adjust the price for food. It is indeed possible to write separable utility functions for all such node items whose quality is enhanced. It is clear, however, that price-adjustment for electricity would need to go beyond the energy-commodity-group. Further, a detailed historical price information for all commodities in the context whose quality is changed due electricity for such a tree-structure to be set up would be required for an empirical analysis[2].

While theoretical methods for demand have been refined over time, the empirical issue of unavailability of market prices (and thereby of quality) has probably not received the attention it deserves [2, 9]. A methodology commonly used in demand studies for its robustness and simplicity is the AIDS model (see Deaton [1]) - which we adapt for quality purpose (see McKelvey [2]). Based on PIGL approaches originally popularised by Barten[10], Gorman[11] and Muellbauer [12], the method is simple to implement and yet readily conforms to the demand function restrictions. Still, AIDS implementations - like any other - are often ailed by the lack of observed market prices. An approach used quite often in empirical studies is to stick the implicit price-deflator (or unit-values) as prices in the AIDS regressions (see Tafere et al [13] for a recent implementation). While this may be the best way to proceed when market prices are not observed, their use can result in a severe measurement error since the quantities consumed by the consumer appear both in the dependent variable (as budget share's numerator) of the AIDS equations (see Equation 4) and in their control variables (as the denominator for unit-value used as price). Since observing market prices prevents such measurements errors (see the discussion by McKelvey [2] and Gibson [9]), we prefer their use for measuring the consumption of quality.

The lack of market prices in consumer surveys hardly makes the AIDS methodology less usable - just that the correction of such measurement errors is not always trivial (see Deaton[14, 15] for one such approach and McKelvey's assessment [2] of the Deaton method). AIDS method can be readily extended to other panel data scenarios. For example, the demographic variables in the AIDS model can be accommodated with the so-called demographic translation where the intercept $\alpha_i$ in the equation $w = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} ln(p_{jh}) + \beta_i ln(\frac{x_h}{P})$ is expanded as

$$\alpha_i = \rho_{io} + \sum_{k=1}^{s} \rho_{ik} d_k$$

Here, $i \in [1,n]$ and $d_k$ are the demographic variables (see Pollack and Wales [16]). Similarly, the structural changes in the parameters over time is addressed with a dynamic AIDS formulation that is often based on differencing the budget shares over the time-periods (see Anderson and Blundell [17, 18] for a first-order dynamic model).

Heien and Wessells [19] adjust AIDS to account for cases when there is no consumption on certain items - by incorporating an inverse Mills ratio into the AIDS regressions. This approach is particularly relevant for the analysis of quality since the consumption on high quality goods is often "truncated" for a significant part of the population. This approach performs a first-stage probit where the dependent variable

[2] Notice that we are concerned with the change in quality - rather than change in taste in the current study. Another way of stating this assumption is that all changes in tastes that are relevant to our study must arise out of a change in quality in one of the commodities.

is an indicator variable - denoting whether the household $h$ consumes the item $i$ (given $s$ demographic variables) or not.

$$Y_{ih} = f(p_{1h}, p_{2h}, ..., p_{nh}, m_h, d_{1h}, d_{2h}, ...d_{sh}) \tag{2}$$

The inverse mills ratio is thus calculated for every household using the probit regression equation 2. This ratio i.e. the quantity $R_{ih}$

$$R_{ih} = \begin{cases} \frac{\phi(p_h, m_h, d_h)}{\Phi(p_h, m_h, d_h)} & q_{ih} > 0 \\ \frac{\phi(p_h, m_h, d_h)}{1-\Phi(p_h, m_h, d_h)} & q_{ih} = 0 \end{cases} \tag{3}$$

is used as an instrument in the following equation (second stage) to estimate the item-elasticities:

$$w_i = \rho_{io} + \sum_{k=1}^{s} \rho_{ik} d_{kh} + \sum_{j=1}^{n} \gamma_{ij} ln(p_{jh}) + \beta_i ln(\frac{x_h}{P}) + \delta_i R_{ih} \tag{4}$$

Here, we typically have

$$ln(P) = \alpha_0 + \sum_{i=1}^{n} \alpha_i ln(p_i) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} ln(p_i) ln(p_j)$$

However, a simpler linear-AIDS formulation is often used so that $P = \sum_{i=1}^{n} w_i ln(p_{ih})$ (the Stone's index). Notice that the inverse-mills instrument must also comply with the symmetry and homogeneity-restrictions and thus one must ensure that $\sum_{i=0}^{n} \delta_i R_{ih} = 0$ in addition to $\sum \alpha_i = 0$, $\sum_{i=1}^{n} \gamma_{ij} = 0 \forall j \in [1, n]$ , $\sum_{j=1}^{n} \beta_i = 0$ and $\gamma_{ij} = \gamma_{ji}$. Heien-Wessells[19] choose a more general approach for the instrument variable by adding an equation to the system with term $\delta_i R_{ih}$ replaced by $-\sum_{i=0}^{n-1} \delta_i R_{ih}$so that the additivity condition is always met. Unfortunately, they are compelled to use the unit-values instead of (unavailable) market prices - an issue probably made worse by having to stick in the unit-values (obtained from the consumers' using an item) as prices on to those who don't consume the item. This is needed because the price isn't available when the indicator variable is zero i.e. when the consumers don't consumer an item[3].

Deaton method[14] gets around the price-deflator (unit-value) issues by using the cluster-averages fitted budget shares and unit-values (which are used instead of the unobserved prices). Even though Deaton uses these methods for a cross-sectional analysis, the intuition can be easily extended to panel data (see McKelvey [2] for this extension). We reckon that the so-called RDMP ("real" deviations from regional/quarterly mean prices) approach (see Cox-Wohlgenant [20]) - used for studying quality variation in commodities when market prices are not available - would also suffer with the issue of measurement error that is introduced with the use of unit values (see Gibson and Kim [9] for further details).

[3]In fact, Heien-Wessells use a regression with unit-values obtained from users who do consume the item as the dependent variable and with control variables as their region, income etc. . The values thus obtained are used as price for the consumers who don't consume the item.

# 3 Model

## 3.1 Separability and Visible Consumption

In the empirical studies on demand, the separability of demand for commodities that are grouped on the basis of their functional utility is often assumed. The separability is essentially a structure imposed on consumer demand by the researcher, but what make them ubiquitous in the demand literature is high-level interpretations they offer for consumer demand. A high-level interpretation is what makes a comparison with the past states of demand or the state of demand in other economies possible.

Yet these functional groups are largely ignored in the literature of status consumption. While measuring the consumption of"conspicuous" items, the demand elasticities are often estimated in the literature on conspicuous consumption using a collection of visible/conspicuous items as if they were a separate commodity group themselves. The literature justifies the use of a separate group of conspicuous items with the assumption that conspicuous consumption has a zero-utility in the usual (private) sense (see Ireland model[21] and Hopkins et al. [22] for the details of this argument). This approach makes sense only because the focus of such studies is the visible consumption itself. Thus the demand for the items that are known to be visible *a priori* can be assumed to be also separable. On the other hand, working with consumption diary microdata in the developing (largely agrarian) countries, we find this assumption (henceforth referred to as the futility assumption) somewhat limiting. This is because in such markets, a high-quality item (priced higher in the market than its substitutes) is conspicuous not solely on grounds of its visibility - when most people don't have access to electricity in an economy, then having electricity as opposed to kerosene can also provide a status benefit (even though electricity is not a status or visible good *per se*). Similarly, when most people take public-transport to work, having a car or taking a taxi also provides a status benefit and so on. To use a framework for status consumption in a developing economy, we must go beyond the idea of visible items. In the model we develop, therefore, no distinction of private vs public/visible consumption is considered necessary as all items are potential signaling items. Because of how widespread the low-quality consumption may be, an item that is not so important from a visibility perspective (i.e. not futile) may still be relevant for status - which we see as a part of the consumer's long-term utility. This we argue is even more so the case in the developing world - where both the poverty levels and inequality are very high - and conspicuous consumption need not be futile in the usual sense at all.

Arguing that a discussion of status needs to encompass more than visibility in the developing world[4], we seek - i) a quality metric which given a set of functionally separable groups can measure the segmentation of quality in the population and ii) the factors affecting such a segmentation - as the drivers of status. We do not rely on

[4]We note that the visibility approach is often adjusted to take into account the environment of wider inequalities of the developed world. In one particular instance[23], we find that the questionnaires used for visibility surveys in the developing world also take into account the relative income that the consumer associates with a certain "visible" item. See also the inclusion of income structure in the successive study from Kaus et al [24] based on the model from Hopkins-Kornienko[22].

analysis of visible items alone whose demand needs to be considered as if they were a weakly separable group by themselves. To avoid confusion of terms in the future discussion on items and quality, we use the nomenclature provided by Cramer[25] so that "goods" are variants with different quality that all fall under a certain "commodity" (see Cox[20] for further notes on this nomenclature).

Since we don't use the futility assumption, the issue of separability requires a particular attention for any status-related interpretations. The (weakly) separable groups being based on a rather strict criteria (functionality, substitution), the high quality items from different commodity groups are not meant to be interchangeable with each other as such. However, if status is derived from items, then either status items can exist as separable groups (which is the tractable option but is not plausible in our view) or they cannot (in which case an aggregation for them is not possible). For the representative consumer in the model we develop, we choose the latter option and prevent the substitution amongst high-quality items from different commodity groups while keeping the commodity groups weakly separable.

As the consumer maintains different qualities within each group, it is still possible to talk about quality-substitution amongst high-quality items in a certain sense - even though the claims for separability of high-quality are improper - as we explain with a rejection of the two-state approach in Section 3.5. More particularly, as the representative consumer expands her needs and increases her income, she may switch to expenditure on high quality from one commodity group to another. Equivalently two consumers - *ceteris paribus* - may acquire habits of higher quality consumption in different categories. To explain this with an example, consider that even though kobe beef cannot be substituted with a luxury travel as such (the goods being in separable commodity groups), a high standard of life can be indicated by either or both of them and the consumer may make any of the possible choices. Our approach therefore gauges how consumers put some weight on higher-quality within each commodity group based on their long-term circumstances or habit as well as the market prices of higher quality in every category. We do not try to answer how overall quality is budgeted for by the consumer i.e. no attempts are made to understand how consumers may purchase a desired amount of (overall) quality (see Section 3.5 for further discussion). Instead, we focus on how high-quality consumption varies and could be segmented across a population. We argue that such a segmentation is relevant for status-related consumption in the context of the developing economies.

Since quality cannot be aggregated from across commodity groups, the econometric methods we develop employ a vector of quality metrics for every commodity group i.e. a vector $v_{G\in\mathscr{G}}$ where $G$ is a commodity group $G \in \mathscr{G}$ and $\mathscr{G}$ is the set of all commodities so that the representative consumer can prefer higher quality in one group $G_1$ over that in the another $G_2$. The quality metric in each group $G$ is calculated using the local prices of the goods within that commodity. More simply put, a consumer who buys more of the high-priced item within a commodity group is viewed as consuming higher quality of that commodity. The econometric methods then attempt to view both how the demand for quality is distributed across the population and how it varies in response to change in prices.

## 3.2 Quality and Prices

Before we explain the properties of the quality-metric we use for the econometric analysis let's first establish that quality-related inferences cannot be drawn from a model that does not use an explicit measure of quality in the utility function. Without an explicit measure of quality, the commonly used Stone-Geary utility for all commodity groups - for example - would be inadequate to explain demand for quality. Here, the indirect utility function obtained with the optimisation of Stone-Geary direct utility function $u = \sum \beta_i log(q_i + \gamma_i)$ (with $\sum \beta_i = 1$ under the conditions $p \cdot q = \sum p_i q_i = x$ ) can be written as [5]

$$u \equiv \psi(x, p) = \sum \beta_i log(x\beta_i / p_i + \gamma_i)$$

With the assumption that a consumer observes the price index of a commodity (rather than individual prices), an econometric method based on the above indirect utility can be employed for a set of commodity groups with items within. However, since a better quality item is set to be more expensive than the less quality item, a switch to higher quality by the consumer is inseparable from an increase in the price-index of the commodity (whose multiple variants are observed). The increase in the price-index of an item (regardless of whether it's a "parallel" shift -so that the relative price structure of the variants remains the same or not) is going to be always interpreted as the consumer reducing the quantity consumed of the item. Indeed it may still be possible to use the model as long as one assumes that the individual's quality preferences are fixed, but a change in quality as a response to price cannot be studied with the model as such. An explicit measure of quality is therefore necessary in the direct utility (or equivalently in the cost-function) for us to interpret meaningful results for quality using the corresponding econometric method[6].

In our model, a quality metric is therefore associated to the aggregate of goods seen as a commodity (see Houthakker [7], Nelson[26]and Deaton[14, 27] for the related formulations). Here, the quantities consumed for two (or more) items within a commodity (e.g. tea and carbonated drinks under beverages that can be interchanged by the consumer) are viewed as an aggregate-quantity of a composite good (i.e. beverages) along with the value consumed of a quality metric (proportion of tea and carbonated drinks). In other words, the quantity-price tuples $(q_1, p_1), (q_2, p_2), (q_3, p_3)$... within a commodity could be expressed as a quantity and quality pair for a commodity aggregate - $(Q_G, P_G), (\nu_G, p_{\nu_G})$ where $P_G, Q_G$ are prices and quantity consumed for a composite commodity while $\nu_G$ represents the quality of the commodity - treated as a separate good. The quantity and quality pair of goods correspond to the real quantities consumed for several items.

It follows that in the above aggregation, the price of quality (along with the quality metric) should also appear in the model for consumer choice. Further, the demand must also follow the Hick's commodity theorem. This is made possible - indirectly - with a constant relative price structure of the variants of a commodity. The constant relative

[5]The first-order conditions for the Lagrangean $L = u - \lambda(x - p \cdot q) = \sum \beta_i log(q_i + \gamma_i) - \lambda(x - p \cdot q)$ lead to $\frac{\partial L}{\partial q_i} = 0 \Rightarrow \frac{\beta_i}{q_i} = \lambda p_i \forall i \in [1, n] \Rightarrow \frac{\beta_i}{p_i q_i} = \lambda \forall i \in [1, n]$. Further, $\sum \beta_i = 1 \Rightarrow \lambda = 1/x$ and $q_i = \frac{\beta_i}{p_i} x$ .

[6]All of the usual limitations of an additive utility function such as Stone Geary indeed would also apply.

structure is a condition necessary for aggregation and allows the price of quality to be interpreted in terms of a composite price $P_G$ for the commodity aggregate $G$ and the price-structure $p_G^*$. Thus, if $E_G$ represents the expenditure on group $G$, then using a constant relative price-structure allows us to write the demand function $g(E_G, p_G)$ as :

$$g(E_G, p_G) \equiv g(\frac{E_G}{P_G}, p_G^*) \tag{5}$$

Thus, instead of viewing demands for {wheat, rice, corn, eggs, beef, chicken, fish, charcoal, kerosene, electricity} - for example - with a list of 10 functions, one can choose to view the demands for commodities that encompass the 10 goods - {carb,protein and energy} with 3 functions corresponding to each commodity group. The econometric methods can thus be developed to measure the elasticities of the quality and the quantity consumed for each of the three commodity groups. If the AIDS method is used, the constraints on the symmetry of cross-elasticities, homogeneity and non-negativity are relevant for both the cases (10 items or 3 commodities). Since the constant-relative price structure prevents the need for cross-price derivatives related to quality in Equation 5, we do not require an explicit price-of-quality in the cost-function (or the budget share regressions) [7]. The cross-price-derivatives in discussion are therefore with respect to $P_{G\in\mathscr{G}}$ where $\mathscr{G}$ is the collection of all commodities (e.g. {food, energy, household, transport} in the LSMS Tanzania data).

## 3.3 Econometric model

Apart from the prices of items, which we view in terms of a relative price structure and a base price ($P_G$), the quality demanded by the representative consumer in the model is sensitive to needs of the consumer and her household characteristics. Here, what is implied by needs are the means of subsistence that are standardised to be the same for a wealthy and a destitute household. No qualitative distinction is drawn - for example - between the status needs of a rich and a poor family (through their disparate educational needs etc.). The notion of status in the model comes only from the quality and household characteristics that are associated with the common need for the rich and poor households - who all are meant to maintain high quality as possible across all commodity groups - each of which are consumed by every consumer in the population. In other words, only the quality and security associated with these common needs differ amongst population and the higher quality/security is better afforded with higher income (or other relevant household characteristics relevant in the population). The first assumption that helps parametrise factors affecting quality is therefore that the essential needs of the family unit (member) are the same for all consumers. The second

[7] Another way to establish this equivalence is to interpret the quality in the market as the gradient of price of preferred items (sorted by price or otherwise) - an approach that is more akin to an aggregation problem. Consider for example that with energy, such a gradient would be calculated with the price one has to pay to get more of electricity. Similarly, for food, this would be the price one has to pay to get more of meat etc. The formulation effectively relies on the consumed basket aggregate - and is thus equivalent to the composite aggregated method (which relies the price of a quality aggregate).

assumption which makes the consumer utility a non-decreasing function of this desired quality metric - is that all consumers hope to increase either their long-term standing (household characteristics) or quality (ENDC) at all times. These assumptions allow us to write the direct utility for the consumer as:

$$u = \upsilon(\eta, \nu, \rho) \tag{6}$$

Here, $\eta$ represents the needs of the family (a scaled number of family members), $\nu$ is the quality consumed and $\rho$ are the household characteristics that are held for the long-term. Contrasting this with the standard direct utility $u = \upsilon(q)$, we essentially view $q$ as $\eta, \nu$ and include $\rho$ in the direct utility only to model the segmentation of the long-term prevailing conditions under which quality($\nu$) and needs ($\eta$) are availed by the consumer. Ignoring the price changes for a moment, consider how a consumer might address her essential needs and improve quality at the same time. In the environment of stable prices, it follows that the distribution of ENDC weights - which the quality metric measures - would be different for a richer household that has access to higher income, better markets (which higher quality goods are available) and more assets (whose upkeep and maintenance costs cause the demand for quality to be different from someone either does not own or have given up on hopes of acquiring assets or addresses). We include all such factors within $\rho$.

Now if $\rho$ was to be held as the same for two consumers, quality $\nu$ essentially competes with the needs $\eta$ of the consumer. Further, due to the very definition of needs, the consumer cannot sacrifice needs in order to purchase quality. Thus, when prices are stable, a rise in ENDC (quality) must come either from an increase in income or from change in circumstances i.e. socio-economic standing - $\rho$ or $\eta$. Likewise, a drop in quality must be due to change in circumstances(socio-economic standing which we assess with a vector of occupation, education, region/locality/neighborhood/assets-$\rho$) or addressing of more urgent needs ($\eta$).

Since $\rho, \eta$ are likely to be fixed in the short-run, we can now consider the effect of price changes. Here, the fact that quality in one group may be preferred to quality in another becomes relevant (hence the use of $\rho$). As we have discussed before, quality across commodity groups may not be aggregated into an overall quality metric but a consumer could prefer quality in one commodity (e.g. transport) to another (food or vice versa) - driven by the change in prices and her long-term standing (household characteristics, ownership of long-term durable good or needs). For example, a middle-income family that cuts down on expenses (due to loss of income, change in circumstances or due to price changes) to move to a cheaper region (where the rents/housing prices are lower but the restaurants are fewer) effectively changes its long-term quality habit. After the move, the consumption on expensive dining or similar options - which signaled status and improved self-image - may be compromised for the sake of improved assets or improved quality in some other category (e.g. better jewelry, better schooling for children). If the needs of the family $\eta$ (indicated by the number of family members in the example) haven't changed, a change in long-term habit of quality could be due to $\nu$ (quality prices) or $\rho$ (circumstances other than needs) . Notice how the difference between basic needs and status-needs is not a concern for the model - which only observes the differences between high-price and low-price consumption in the same group - assuming that the essential needs are independent of the wealth of the consumer (an important condition for Equation 6). All non-essential needs are interpreted solely in terms of quality (ENDC) and the socio-economic standing. Further, the constant-price structure assumption along with Equation 6 allows the price elasticities with respect to the base-price $P_G$ to enter into the econometric equations side-by-side with the needs $\eta$ and personal characteristics $\rho$.

The model in Equation 6 would need to be more general if one considers that a substitution exists between consuming quality now vs improving the long-term standing. For example, both having a car and hiring taxi could carry some status benefit to the consumer (as both are in the transport commodity group). If all else being the same, a consumer is indifferent to saving for a car (and forgoing taxi consumption) and paying for taxis, then the model based on $\rho, \eta$ and $\nu$ is no longer sufficient. Fortunately, the real world offers us a simpler context - where only either one of i) non-durable consumption and ii) long-term assets may dominate within a commodity group. For example, only one of i) car ownership and ii) expensive-taxi would dominate in data from a particular economy. The empirical implications of this assumption are therefore not as serious and allow us to view quality variation only for commodities where assets do not play a role. In the Tanzanian data for example, no assets are meant to exist for food and energy categories given the non-durable nature of their consumption. Further, the quality variation is immaterial in the data for non-durable consumption in transport and household categories. Thus if every commodity group is either dominated by long-term asset or it isn't, there are only two ways a consumer can spend on quality - either by switching between the variants of a composite commodity (choosing higher or lower quality of a composite commodity where long-term assets do not exist) or spending more on an elementary commodity (where long-term assets dominate and quality variation doesn't exist) [8]. The implication of the assumption is only that both elementary consumption and composite commodity consumption - compete for the consumer's money. This offers a significant reduction in complexity for the empirical analysis.

Since the segmentation of habits is a key purpose of our analysis, we use a metric simple enough to be applied to any commodity group where the prices of items is available. A metric that achieves this goal is the relative-price-weighted sum of expenditure per physical unit within a commodity (which is argued for in Nelson [26]). Given a group $G$ when the quantities consumed are observed, the metric is simply a sum of quantity $q_{i,G}$ weighted by price-structure ($p^*_{i,G}$) :

$$\nu_G = \sum_{i \in G} \left(\frac{q_{i,G} p^*_{i,G}}{q_G}\right) \tag{7}$$

[8] Recall that an elementary commodity is one where no quality variation is supposed to exist and higher quality can be achieved only through more purchase of the commodity. A composite commodity - on other hand - is a commodity where a sufficient quality variation exists and a consumer can use a mix of higher quality and lower-quality goods within the commodity.

Here, $p^*_{i,G}$ represents the price-structure ratio corresponding to the item $i$ in a group $G$ and $q_G = \sum_{j \in G} q_{j,G}$. $q_G$ is the physical sum of quantities consumed. Recall that regardless of how the items are aggregated, such a formulation always requires the aggregation conditions of the composite commodity theorem conditions to be met (see Nelson[26] and Deaton [27] for details). The ratio $p^*_{i,G}$ is effectively a price-multiple with respect to the base-price $P_G$ i.e. the price of the base-commodity in the group $G$ with respect to which the quality in $G$ is being considered. In other words, $p^*_{i,G}$ represents the price structure vector such the prices $p_{i,G}$ observed in the market for item $i$ in the group $G$ can be written as follows:

$$p_{i,G} = (P_c \times p^*_{i,G}) \forall i \in G \tag{8}$$

Let's consider an example to illustrate how the measure of quality would relate with the total expenditure $E_G$ in the group $G$ and the group aggregate price $P_G$ (see Nelson[26] for a more detailed theoretical discussion). Let's say we observe the consumption on lamb and chicken within a protein commodity with prices observed as 20£, 10£ (resp.) and the quantities consumed by the consumer are 1.5 kg and 2 kg (resp.). One approach is to the treat the physical sum of quantities i.e. 3.5 kg as the quantity of the composite good "protein". However, it is difficult to argue that the consumers optimise the weight of the composite good - rather than comparing the disparate quantities of items (see Nelson [26] for the technical details of this argument)[9]. A quality-adjusted quantity is a more appropriate formulation for this purpose. Since quality is to be priced in the market, we can consider this price-structured weighted quantity as the quality-adjusted quantity and express the pair$\{(20,1.5).(10,2)\}$ as $(10,5(=2\times 1.5+1\times 2))$ and $(0,1.14=\frac{4}{1.5+2})$ - where the minimum of $20,10$ is used as $P_G(=10)$ and the price-structure $p^*_G$ is (2,1) for {lamb,chicken}. In a certain way, the price of composite quality can be viewed as exactly zero since the total expenditure is already accounted for in the composite good's consumption i.e. 10×4 = 40. If the quantities consumed were slightly different for another consumer e.g. $\{(20,0.5),(10,3)\}$ then too $(E = P_G Q_G = \sum_G p_i q_i = 40)$ and we have the latter pair interpreted as$(10,4(=2\times 0.5+1\times 3))$ and $(0,1=\frac{4}{4})$. It is easy to see that the quality is lower for the latter case even though the physical quantity is the same. The quality adjusted quantity accounts for the total expenditure and as such there is no price needed for the quality value calculated in the above fashion. Even though quality is treated as a separate good, the use of a quality-adjusted quantity instead of the physical quantity gives us a clearer view of the composite price of the commodity[10].

The above quality metric readily lends itself to estimation of price derivatives (with respect to base-price $P_G$) related quality and the household characteristics relevant for quality. The needs of the consumer $\eta$- which we estimate with an

[9] If the sum of physical quantities is denoted as $q_G = \sum_{i\in G} q_i$then $Q_G = \frac{E_G}{P_G}$ where $E_G$ is the expenditure and $P_G$ is the price-index. The core of Nelson's argument is that $Q_G = q_G$ only if quality effects are absent since $Q_G = q_G \Rightarrow \frac{E_G}{P_G} = \frac{E_G}{V_G} \Rightarrow P_G = V_G = P_G v_G$ (unit quantity is interpreted as $V_G = \frac{E_G}{q_G}$ ).

[10] This measure of quality is slightly different from Deaton's approach [14] which uses the sum of physical quantities.

age-adjusted number of family members - can be treated as the intercept in the budget share (AIDS) regressions. Lastly, the assets ownership indicator variables (1 if owned 0 otherwise) used in the AIDS intercept can be viewed either as split across commodity groups (i.e. as dummy variables indicating assets owned in transport, in household or in energy etc.) or as variables indicating which particular cluster of assets is owned by the consumer (agricultural, electric, household-related etc.). Since our model does not specify how asset ownership in one commodity group (e.g. transport or household) may influence asset ownership and/or quality in another (e.g. household) we avoid the use of indicator variables altogether in the assets and use the logarithm of the total cost of assets owned instead. This approach also allows us to consider the quality differences within assets (i.e. an expensive house or an expensive car can be separated from cheaper options).

## 3.4 Implementation

The quality metric discussed above helps measure the relative consumption quality for commodities in the population. To illustrate how the quality-aggregations provide a structure to the demand for various items in our discussion, we compare i) an item-wise AIDS implementation with ii) the AIDS implementation of quality-aggregates. The details for these methods have already been provided in the Section 2 above. The first method is a detailed item-wise analysis including the cross-price elasticities (Heien-Wessells [19]) while the second method uses the above defined measure of quality ( except for electricity - which is simply an indicator variable - for reasons we discuss in Section 3.5). In the latter method, we also include the ownership of assets (logarithm of the costs of the asset) and the demand for items we don't calculate the quality metric for.

### 3.4.1 A selection approach to item-wise AIDS

This first AIDS variant uses non-consumption corrected estimates of demand elasticities to explain the effect of prices and household income on the demand for every item. The method comprises of two parts - the first being a selection step - needed to account for the non-consumption of a certain items. The second part is a pooled static-AIDS method (see Equations 2 and 4 in Section 3) that is run over the consumption microdata from 2010, 2012 and 2014.

The first-part i.e. selection is a probit that provides refinement to the second-part item-wise AIDS (which would have been appropriate in a more homogeneous dataset with less instances of non-consumption for items). The dependent variable in the probit is an indicator variable which is unity when the item is consumed and zero otherwise. The control variables used in the first-stage probit are the prices of the item and market-characteristics (region etc. - see Equation 2) . The inverse-mills ratio calculated from the probit (as explained with the Equation 3 - see Heien-Wessells [19]) is in turn used as an instrument in the AIDS regressions of the second part (see Equation 4). Notice that the instrument is calculated for every household and every item - and is part of the additivity constraint that is used for estimation.

### 3.4.2 An LA-AIDS model for Aggregate Quality

In the AIDS equations used for this second "Unrestricted method" variant, we measure quality as described with the Equation 7 (Section 3.3). These AIDS equations used are evidently fewer than those in the item-wise censored model of the first AIDS variant. This is because every commodity aggregate - comprising of several items - corresponds to one aggregate quantity/budget equation and a quality metric corresponding to the commodity aggregate. The aggregation also circumvents the issue of non-consumption since it is much less likely for no items to be consumed within a functional commodity group (fruits-veg, starch, protein etc.). Therefore no correction related to non-consumption is needed in this approach.

Since we assume that a commodity group is either dominated by asset possession or it isn't (see Section 3.3), the quality variation is only meant to exist in a certain commodity groups. Out of the commodity groups - food, energy, household and transportation - which we use in the current chapter, we observe quality variation only in food and energy - where base-prices derived from the local market prices for goods within a commodity are used. The expenditure on household and transportation commodity groups - on the other hand - are treated as if they were (each) just one commodity i.e. no quality differences are observed in the non-durable consumption within household commodity group (toilet rolls, clothing etc.) and the transportation commodity group. Here we use the CPI (base year 2010) for household goods and transportation as prices in the respective budget share regressions for household and transportation (the way base-prices would be used for food and energy commodity groups).

The base-price $P_G$ for every aggregate $G \in \mathcal{G}$ (obtained from the price analysis detailed in Section 4) is simply the minimum price of the goods within the commodity group. This serves as price of the commodity whose budget share is set as the dependent variable in the following AIDS equations:

$$w_G = \rho_{1,G} + \sum_{k=1}^{s} \rho_{1,Gk} d_k + \sum_{j=1}^{n} \gamma_{1,Gj} ln(P_j) + \beta_{G,1} ln(x) + \varepsilon_{G,1} \tag{9}$$

$$ln\ V_G = \rho_{2,G} + \sum_{k=1}^{s} \rho_{2,Gk} d_k + \sum_{j=1}^{n} \gamma_{2,Gj} ln(P_j) + \beta_{G,2} ln(x) + \varepsilon_{G,2} \tag{10}$$

There are $s$ demographic variables (including the asset ownership related dummy variables) in the above equation. The composite unit value $V_G = P_G v_G$ serves as the dependent variable in the latter quality equation [11](see McKelvey[2] and Gibson et al [3, 9] for details on the implementation of McKelvey's "Unrestricted method"). Since the price for quality is not required, the price-coefficients in the "additional" $ln\ V$ equations are not subject to any additional symmetry restrictions (see Appendix for the details of the implementation). The usual constraints of adding up, symmetry and non-negativity of Slutsky matrix do apply to the above demand equations.

[11] Quantity elasticitities are related with quality elasticities with the following relationship: the income elasticity of composite quantity is the sum of income elasticity of quality and the income elasticity of physical quantity (see [26]).

### 3.5 Discussion

We now discuss the conditions under which the quality aggregation should be interpreted and detail the reasons for some choices made in the model for quality. The first condition for the quality metric described in Section 3.3 is the rather stringent condition of a constant relative price structure for aggregation (i.e. the Hicks aggregation commodity theorem). This is the reason why our classification of items into commodity groups is driven by the price-structure (and not just on separability of demand).

The second condition is that the quality metric is aggregated over a commodity group only as long as the goods within the group can be ordered based on a certain criteria - an order that is the judgment of the researcher - not the consumer (see Nelson[26] for this clarification). If we take the meat commodity group as an example, there is no reason to think that goat is of higher quality than beef if the former is only slightly costlier. However, quality differences are more plausible if the price differences are larger - e.g. with the choice of pulses over chicken under the protein category. Such a judgment of quality (imposed by the order of price in our case) is thus provided by the researcher at an appropriate level of granularity - a justification that is a prerequisite to interpret any sort of quantity aggregation.

Due to the obvious limitations to the universality of such an aggregation criterion, the utility of the above metric in understanding consumption of quality is restricted by the very aggregated view that makes it desirable[12]. We illustrate this with a comparison of item-wise and aggregated approaches in the results (see Section 5). While the item-wise approach gives us a more detailed sense of elasticities, an interpretation that is more desirable for an understanding of the segregation of quality across the segments in the population is provided more readily with a quality-aggregated view.

Notice that since we don't attempt to aggregate high-quality items from across commodity groups, a two-state budgeting approach (commonly used in the literature on household demand) can be conveniently discarded. Recall that the two-state budgeting represents a hierarchy of separable goods. If the futility argument were valid for high-quality products, one could have aggregated quality consumed in different commodity groups into an overall quality metric. Without the futility argument, on the other hand, such an aggregation is improper. This is because a key condition for the tree-like structure of a two-state budgeting model is that the relative demand between two items in any two different categories changes only in response to the total expenditure (i.e response to price changes and total expenditure changes must use the same channels of communication in the commodity tree[27, 28]). Evidently, this is impossible to guarantee if we claim that quality can be substituted from across the commodity groups i.e. if consumers may be indifferent to choosing quality in one category to that in another.

To illustrate the issue with aggregating quality from across commodity groups in a two-state budgeting model with an example, consider a two-group world of food and energy - within which we have items : beef-bread-milk and electricity-kerosene-charcoal respectively. If consumers don't cut back on quality in

[12]Notice also that one cannot talk about the quality metric $v_G$ defined above unless the quantity of consumption is observed in data.

one group to gain quality in another, the items - food and energy - can be assumed (and tested empirically) to be separable groups - since it is plausible for a consumer to allocate some money $x_F, x_E$ on food and energy respectively in the first-stage (based on her knowledge of the price-indices for food and energy). Once the expenditures $x_F, x_E$ are fixed, the (representative) consumer can subsequently decide in the second-state how much of kerosene or electricity to buy with $x_E$ and how much beef or bread to buy with $x_F$. Individual prices of every item within the category would decide which food varieties and which energy sources are to be purchased. Under the conditions of weak separability, it is implied that cheaper bread (price-sensitivity for bread consumption) would not make a consumer choose a more expensive fuel (price-sensitivity for electricity, charcoal etc. - see [29, 27] for details on the theory). This condition - however - falls apart if we introduce the substitution for quality across the commodity groups. More particularly, if we assume that a consumer gets quality utility from using electricity and from eating beef interchangeably, then having decided on $x_F, x_E$ does not guarantee that she would not switch to more consumption on electricity if kobe beef becomes cheaper. This is because overall quality - due to our assumption - is aggregated from across the commodity groups. Thus unlike in the case of energy and food without quality substitution across commodity-groups, cheaper food can be compensated with the switch to electricity. If we have set ourselves to test the presence of such a substitution across commodity groups, we cannot assume that it simply doesn't exist. Since price sensitivities and total expenditure sensitivities don't follow the same channel of communication in the two-state budgeting tree, the necessary condition for two-stage separability is not met (see [29, 27]). Thus even though a tree-structure could present a relevant view of the hierarchy of commodity groups, we cannot accept the model due to the weak separable condition on overall quality.

The grounds for rejection of the two-state model - i.e. the fact that quality cannot be aggregated at a level coarser than the separable categories (commodity groups) we impose on the demand - is the third condition for our quality-aggregation. The question that follows next is how fine can the commodity group over which the quality is aggregated can be divided (the finest granularity of groups being one where every commodity group is composed of exactly one item). As would be amenable to the separability assumption, we're guided by the substitution within the group while we form the commodity groups from the individual goods. Thus if there is substitution possible between two items, they ought to be in the same commodity group.

An approach that strictly follows this rule is the so-called "simple-repackaging" method - which assumes a perfect substitution across goods in a commodity group - so that a higher quality product can be viewed in terms of a lower quality product. However, the perfect substitution assumption is problematic when the goods within the commodity vary widely in quality (or fundamental other characteristics). As an example, the demand of a better good in a commodity group as a multiple of the worse good may be plausible for meat/food - but considering electricity in terms of coal into one aggregate is far less plausible (see the discussion on consumer indices in Fisher-Shell [8] on how this issue could be remedied). At the consumer's (behavioural) level, the substitution between kerosene and electricity is not as straightforward as amongst the quality variants of protein - so while a metric for quality does make sense in a

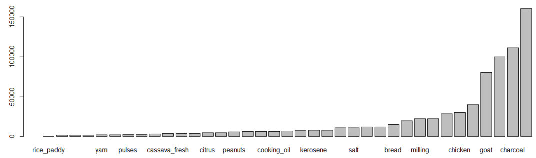


Figure 1: Items sorted according to maximum prices recorded in the survey

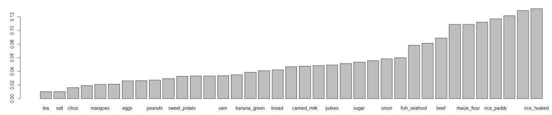


Figure 2: The items sorted according to the mean budget shares

setting where a mixture of high and low quality items are consumed (e.g. in food) it fails to account for quality differences when choices vary widely. Reflecting this observation, the exception we make for wide quality differences is the fourth condition for quality aggregation. In other words, wide differences in quality or fundamental characteristics are not to be viewed as a continuous measure for aggregate quality. Thus while we do form groups based on possible substitution within the group for all the food-categories, we use indicator variables instead of the aggregate-quality measure for energy commodity (where consumption on kerosene, charcoal, electricity and gas are observed in the data).[13].

In the sections that follow, we use the data on possession of assets in the Tanzanian consumption microdata to seek differences in the quality habits in the different asset ownership bands (and other characteristics). For consumption of quality, the commodity groups: food, energy, household and transport are identified. The quality of protein, carbohydrates, beverages etc. - each of which is treated as a composite good - are measured with the metric presented in Section 3.3. On the other hand, household and transport are treated as elementary goods (i.e. no quality variation is assumed and only the cost of the items is considered). The cross-price effects for the items in the survey for 2010, 2012 and 2014 are also presented to provide an insight into how prices of non-food items may have an effect on food quality and vice versa.

[13]Going a step further in the direction provided by the fourth condition, a discrete view of quality differences may be appropriate for all non-food categories (e.g. transport - which is the subject of a seminal work on discrete choice by Domencic-McFadden[30]) where the consumers either consume a certain quantity of an item or they don't at all (more so in the developing countries where income inequalities are starker). The discrete utility models are to be explored in the subsequent chapters.

# 4 Data

With recent economic growth and a history of nationalization, Tanzania serves as a typical case-study of consumer markets in the developing sub-Saharan Africa. In more recent times, since the investments from South Africa and Western countries have significantly expanded in the country, the trends of urbanisation appear clearly in its consumption data. The data used for our studies is from the Living Standard Measurement Study (LSMS) - which is conducted by the World Bank.

The substitution between items has driven the "functional" classification of items. A functional category is one where there is a sufficient degree of substitution possible amongst its members. As an example, goat and beef are put under the same group - but not with tea - as there is little sense in substitution between tea and meat. The food category is further divided into sub-categories so that demands within the category can be considered separable in both the functional sense and the price-movement sense (which is the basis of the computation of quality as presented in Section 3.3 ). Due to reasons cited in the Section 3.1, the items of high quality are not explored as a separate category by themselves.

LSMS survey uses diary method for food and recall methods for non-food items. While we are interested in how sparsely some items are consumed, we are careful not to interpret the absence of an item in the recall record for a household as the item not being used by the household. This is because not all items in the recall diary are recorded at the same frequency - some items in the survey are recorded as they're recalled from a month before the interview - while others may be recalled from the year before the interview. Similarly, the an interpretation of the differences in the budget share over snapshots (2010, 2012 and 2014) for a particular household and item- cannot be interpreted as a consumption pattern of a household since we may not have the records in the recall diary for all three instances of the survey (2010, 2012 and 2014)[14]. Such matters can be made worse in cases of expenditure where seasonality (time of the year) is also an important factor. To avoid such issues, therefore, we rely only on the average expenditure by a group of households in a particular region/district rather than interpreting the changes in consumption over the years of an individual household.

The survey does not include the data on tribal/social identities of the participants. Other proxies that may be correlated with social identities (e.g. language spoken, regional origin etc.) that are available at the village or equivalent level in the survey are used for the purpose. Other household characteristics we consider of importance are education, urban residence and income levels. Both occupation and education are converted into ordered variables for use in the econometric analysis.

## 4.1 Descriptive Statistics

The consumption microdata from LSMS records food expenditure with a greater

[14] In fact, we find that for cases such as electricity, the survey field for "main lighting fuel" seems more reliable for the test of electricity consumption than the last expenditure on electricity.

level of detail than other expenditures. The summary statistics from the microdata for 2010 (base year) are shown in Table 1.

The market observed prices available in the survey facilitate a comparison of items that are associated with higher budget shares with those that have higher prices in the local markets (see Figures 1 and 2) . When observing items that may be of "visible" value ( or those that are rare or expensive), the descriptive statistics do not suggest a common trend for their consumption. On one hand, we have tea and coffee that are expensive but are used in very low quantities and on the other hand, we have skin-cream that despite their apparent importance for visibility purposes are not as expensive and show no variation in quality (as inferred from unit-values). Further, the consumption on these items may depend very much on what they're used for i.e. the household characteristics that drive their consumption vary depending on their functional utility. For example, one finds the housing-status (rented/owned/free-house) relevant for expenditure on carpets, residence-area relevant for electricity and so on.

The box plots for market observed prices (per standard unit that are liters or kilograms) are shown in Figure 3. As one would expect, the variation in prices for meat and seafood is much higher than those in the prices of the rest of the food items. Processed foods as well as energy sources (kerosene) seem relatively high in price compared to other food items. The prices thus observed are matched with the consumption microdata based on the district where every household is located - a process whose details are presented in the Section 4.2.

For the empirical analysis related to quality-related aggregations discussed in Section 3, the consumption microdata was obtained by merging and normalising the item expenditure data collected through recall and diary methods. As the weekly data is mixed with yearly data, an extrapolation of past week's consumption becomes necessary. Since larger families are more likely to underestimate their purchases when recalling, such an extrapolation can overestimate food costs. To address this issue, we use adjustment factors both for the multiples used in extrapolation of weekly/monthly data (based on their usage pattern) and for the number of family members (i.e. older

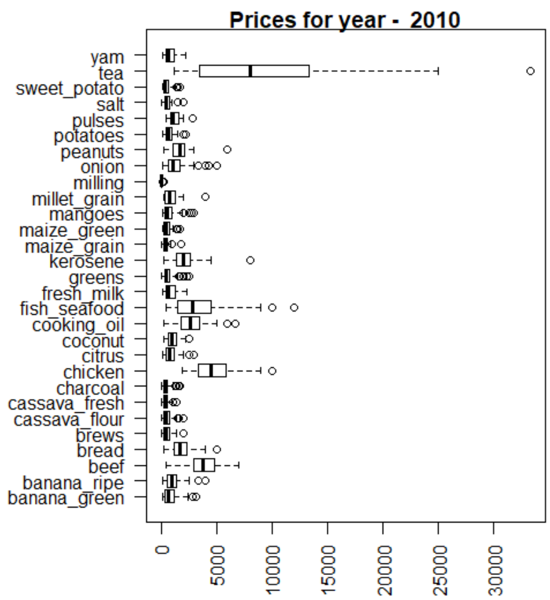


Figure 3: Recorded prices in 2010

family and children are not counted as exactly 1 family member).

| | |
|---|---|
| Mean Household size | 5.27 |
| Mean age of household head | 46.36 |
| Average number of rooms per household | 3.33 |
| Percentage with household head educated secondary or higher | 16.14 |
| Mean Total Expenditure (Tanzanian Shillings) | 2471122 |
| Percentage Employed in Agriculture | 47.76 |
| Total Number of Households | 2979 |

Table 1: Descriptive statistics for LSMS Tanzania 2010

We find that the income spectrum is heavily skewed in the data - as is characteristic of a developing country. Out of the households who consumption diary is available, only around 30% have any reported income. Amongst those who have a reported income, having two jobs and owning multiple self-owned (small) businesses is not uncommon and the mode of payment is often non-cash. The amount of income recorded for the household is thus frequently based on the person's estimate of the item provided as income. All of these can make the inference of income noisy at best. The income levels themselves seem poorly correlated with expenditure levels. Given the sparsity of available income data and its poor quality, we choose to use the occupation ranks (which are correlated with the observed income) for the current study.

## 4.2 Matching of Prices with Items consumed

As noted before, out of the food, energy, transport and household commodity groups, we only consider quality differences in food and energy subgroups - whereas the other commodity groups are viewed as groups where asset possession (or ownership of long-term durable goods) matters more than quality variation.

The prices observed in the market drive our measure of quality and the later discussion of cross-price effects. Since we don't use price-deflators in our analysis (due to measurement errors that arise with quantity being in both in the denominator of the control variables and in the numerator of the dependent variable), a matching of prices of the items observed in the market (identified by region and district) with the prices of the items purchased by the consumer (with known region and district) is of key importance to our analysis.

LSMS Tanzania surveys record the prices observed in the markets spread across all regions in the country. These prices are available for items in the food and energy commodity groups. The items within these commodities are grouped according to the *local* relative structure of the price time-series observed from 2008 to 2014 shown in Figure 5 (see McKelvey[2] for a similar analysis of rice varieties).

To classify the goods based on the movement of their prices, we gather the market prices for the relevant items during the survey periods (2010, 2012 and 2014). Following is a list of issues that were resolved to obtain the price data to be used for subsequent classification of items and matching with consumption microdata.

1. A lot of items in the year 2008 were recorded with the unit as pieces/"idadi" (e.g. 5 pieces of chicken or 4 pieces of banana etc.). For year 2008, there are 588 observations that record prices in terms of pieces. Since it is difficult to associate a physical measure to a piece, the year 2008 was removed from our analysis[15].

[15] These items were : maize_green, bread, bunscakes, pasta, cassava_fresh, cassava_flour, sweet_potato, yam, potatoes, banana_green, othervegstarch, pulses, peanuts, coconut, onion, greens, dried_canned_veg, banana_ripe, citrus, mangoes, sugarcane, goat, beef, chicken, eggs, fish_seafood, dried_canned_fish, fresh_milk, canned_milk, firewood, cooking_oil, salt, brews, kerosene and charcoal. The use of pieces, however, has been discouraged since 2008. For 2010, we only see the prices on bread, banana_ripe, goat, beef, chicken, eggs and fresh_milk items where they are recorded in pieces. These price recording are ignored for our analysis. In 2012, the items where pieces are used are limited to - bunscakes, goat, chicken, eggs and fish_seafood.

With the asset ownership details not being available for 2008, the year being too far back to use the 2010 household and transport CPI baskets and it being difficult to find prices of electricity for the year, the year 2008 offers very little to be included in our analysis. That said, we do use the 2008 prices for when the units recorded are not “pieces”, in order to determine the price-structure that aids the classification of food groups (with similar price-structure). See Figure 5 for the prices observed from 2008 to 2014.

2. A few outlier price observations (<5%) are rejected.

3. While a household in the LSMS consumption data is identified down to the village and enumeration area level - we do not expect the surveys to record market-prices at every possible enumeration area and village. We thus associate a district level price to a household’s consumed item. In other words, the price of an item consumed by a household is linked to all households in the same district where the household is located. We make no attempts to arrive at a price granularity lower than that of a district. When there are multiple price observations at the district level, we arrive a district-level price using a regression (discussed below). Further, when the prices are not available at the district levels, we use the regional averages (i.e. the next level of granularity available). For items such as electricity and gas that are limited to a few urban areas, we simply use the price at a national level[16].

The market-price observed for items are recorded in several units. To assimilate all the price data from across the quantities and regions, we run a regression of marked-price against quantity it was recorded for. The coefficients of the regression can thus provide us the price for an arbitrary quantity of the item. For example, if we had observed the marked prices of rice (let’s say) for 1 kg, 2kg, 5kg and 10kg in the market. A typical set of marked prices could be 4£, 7.5£, 18£ and 35£ respectively - see Figure 4 for the plot of these example prices. In this example as well as in the real-world, such unit-prices grow at a less than linear rate (see the unit-price corresponding to the quantities in Figure 4). Therefore a linear regression of $mp$ vs $q$ (where $mp$ is the marked price of the item and $q$ the quantity consumed) is hardly appropriate. Instead, we use a quadratic expression to regress the unit-prices observed against the quantities they’re recorded. Given this regression, the price of any arbitrary quantity can thus be obtained. What we need from the regression however, is merely the price of the unit-quantity (e.g. the price per kg for rice in the current example) rather than matching the appropriate marked price with the quantity purchased by the consumer [17]. The regression we use therefore is $\frac{mp}{q} = \beta \frac{1}{q^2} + c$ (assuming that the marked price is lower than a linear increase - see Figure 4 - by a factor that itself decreases with increasing $q$ i.e. $mp = aq - b/q$). Since we use the price set $q = 1$, the price to be used is simply

[16] Prices for electricity were obtained from Peng at al.[31].

[17] If we were to really look for a perfect match for prices, then we should not only match the quantities consumed by the consumer with the prices recorded in the market but also the time of the household interview with that of the price recorded. This fine matching is both impractical (survey prices being hardly available at such fine levels of granularity) and unnecessary since our view of price is that of a value that is faced by a sufficiently large number of consumers in the survey.

$\beta + c$ for the regression: $\frac{mp}{q} = \beta/q^2 + c$. This regression is run only when more than two prices are available. When the regression can't be used a simple median is used as the price for such cases[18]. The price-inference is therefore a 2-stage process - in the first-stage we obtain prices for every district with the regression method. If there are no marked prices available in a district, we use the regional value of the price. For some items, the region-prices are not available for some regions. Here, the price from obtained by the regressions over the country is used for the item. The prices for some items whose consumption is relatively rare in the diary data (e.g. wild-birds and honey) are not available for any of the regions. Such items (e.g. alcohol) are ignored for the analysis.

With the prices obtained after the above regression, we match the price for every commodity consumed by the household based on the district it is located in. Once we have the observed market prices for every item by every household, we can aggregate the costs across an arbitrary group of items and calculate quality for the group. Being driven by the conditions discussed in Section 3.5, the classification of items into groups is hardly arbitrary - we are guided both by separability and the constant relative price-structure assumption (i.e. first condition). Had we considered separability alone, the classification of all food-items within a food-group and kerosene-electricity into the energy group would have proven sufficient. But the structure of the price movement observed as the annual time series (see Figure 5 for a plot of the average prices for every good) for every item is also relevant for our classification. Thus the sub-groups within food are created based on the correlated changes in price from 2008 until 2014. For example, the price changes for sugar and tea are more similar to each other they are to the changes in prices for rice or cooking-oil[19]. We therefore put sugar and tea within a group complements and rice-cooking_oil within a group nonfresh (to indicate the comovement of prices). Similarly, within the food commodity we observe protein, fruitsveg, densefoods and nonfresh groups as the rest of the food-subgroups. For quality elasticities to make sense, the classification of commodity groups based on the market price-structure is essential due the composite commodity theorem restrictions (i.e. the constant relative price structure).

One may argue that a selection bias is introduced when we use the national-averages for some items and region averages for other. This could be a problem if the unavailability of regional averages occurred for a significant portion of price entries. However, since only about $\sim 4\%$ of the total price observations have no regional averages (only a few items suffer with the issue), the bias thus introduced would be rather insignificant. Another issue could be taken with our grouping of commodities - as it relies on the annual time-series data. It is indeed possible that the price of some items may be more volatile than the others - and thus the volatility being ignored in the annual time series data would bring some bias. But since we combine the household log diary with the recall diary (which is recorded on a monthly or yearly basis), a finer time-series for prices is not enough to address this issue. Such

[18] This does not happen as frequently as one may imagine. For 2010, only 9 identified regions for just 6 items have less than three marked-prices available.

[19] A clustering of regression coefficients was used to arrive at these groups.

a bias could be alleviated only through finer price data *and* a uniformly collected household consumption microdata.

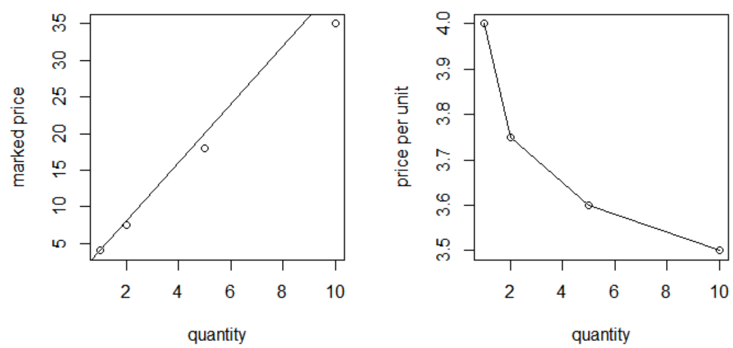


Figure 4: Typical marked prices and Price per unit of an item

# 5 Results

## 5.1 A selection approach to item-wise AIDS

The constraints of symmetry, additivity and homogeneity are essential in facilitating a discussion of the effect of cross-elasticities on demand (measured as budget shares). The results thus obtained with a seemingly-unrelated regression (SUR) after imposing the constraints are presented in the Table 2, 3 and 4. The own-price and cross-price elasticities presented in the results provide us an overview of the demand across rare and commonly consumed items.

### 5.1.1 Variables

The control variables used are `consu` (family size adjusted based on the age of the member), `educrank` (education rank interpreted from the highest education level of head of every household), `invmills` (inverse-mills ratio as explained in the Section 3.3), `lntotexp` (log of the total expenditure) and `occupationrank` (occupation rank - an ordered variable derived from the occupation of the head of the household). Notice that we use occupation rank as the proxy of income (for which the survey data is quite sparse) along side with the total expenditure. The other control variables starting with lp- are logarithm of price corresponding to every commodity in the second-stage AIDS regression (see Equation 4). The columns in the table (names starting with q) indicate the budget shares that are used as the dependent variables in the second-stage regression (see Equation 4).

### 5.1.2 Results

We base our interpretation on both the the statistical significance (value of $z$) and the economic significance (i.e. the size of the coefficient) of the control variables in the results. Observing the signs of the coefficients, we see that mangoes (less consumed) and rice (commonly consumed) are substitutes while fish is a complement to cassava-flour ([qfish_seafood]lpcassava_flour <0 etc.) rice. Further, fresh milk is a complement to beef, beer, cassava_flour ([qcassava_flour]lpfresh_milk = -0.001) and greens but a substitute to bread and mangoes.

The item-wise analysis suggests a split between the items that are consumed in rural households (where the greens and fruits seem more abundant) against those consumed in urban households (where beef, charcoal and beer are more readily available). The elasticities against income and education also show that the items available in the industrialised regions of the Tanzania seem to have positive elasticities for income and education rank whereas the items of agrarian production seem to matter less to the highly educated, middle-age populations huddled in the urban areas. While many of such insights could be offered with a simpler cross-sectional analysis using a regression of expenditure on the item against the permanent-income or total expenditure by the household etc. (an approach that is followed in much of the literature on conspicuous consumption[23, 32, 33]), an AIDS formulation is favoured due to two reasons - first it allows us to understand the cross-price effects in a more robust framework and second, it can account for the non-consumption of items.

Going further with the cross-elasticities analyses, a comparison of how consumers respond to changes in kerosene and electricity prices further confirms the concentrated availability of industrial items. A contrast between the items that are influenced by electricity prices and those that are influenced by kerosene prices demonstrates this split - since the demand for items such as rice, beef is less sensitive to electricity than it is to kerosene. That is to say, it isn't simply that the electricity consumers are more likely to consume the more expensive items (due to their higher incomes) but that a certain items become less relevant in the urban settings or are simply unavailable in the rural settings. We argue that such a segregation of consumer contexts is relevant for a discussion of status-related consumption in the developing countries.

There is more we seek from the consumption data than one item's preference relative to another - as there are limitations to a cross-elasticities analysis. The detailed cross-elasticities are more likely to fluctuate over time and would require us to use a dynamic approach to ascertain model stability (an inertia in choices could be more explicitly modeled to address this). But more importantly, the cross-price elasticities fail to model the functional separation in consumption e.g. the lack of substitutability between energy and food items. For example, the cross-elasticities of kerosene with respect to buns or cassava cannot be interpreted unless one imposes other assumptions on the closely observed consumption patterns (e.g. more buns would means less cooking in the household etc.) or stricter separability assumptions (see Deaton[34] for a discussion on the need to impose such structures). Since we are concerned with the choice of quality exercised by the consumer when faced with multiple varieties of several qualities and price, an aggregated view of the basket seems more appropriate. Such an aggregation ought to be based on a well-defined criteria - which for our current approach is the price of the items as well as the functionality of commodities. The availability of price information between the years 2008 and 2014 prevents the problems that could be faced by other grouping criteria - such as visibility - the changes to which are not observed over a sufficiently long period of time.

## 5.2 An LA-AIDS model for Aggregate Quality

The quality-aggregated view from the second AIDS method provides the functional view to the consumer choice that is not available with the previous item-wise approach. Observing consumption at a higher level with separability assumptions also facilitates a better interpretation of the cross-price effects - but most of all, the aggregation (subject to restrictions) allows us to interpret the consumer choice for quality - while at the same time representing the substitution between broader categories and the price variants within the commodities in the econometric model.

We also use the aggregation as an intermediate step in merging data from recall and diary methods. In most surveys, the items recorded with more details for the diary items than for the recall items (the types of phones, clothes purchased by the consumers are hardly recorded in the surveys for example). Viewing the diary data at a higher (aggregated) level - i.e. at a coarser granularity that matches that of the recall methods - allows us to have a combined view of data from recall and diary methods.

Thus the price-based aggregation in this second AIDS variant - which we have adapted to our requirements - allows us to have a discussion on quality that was somewhat adhoc with an item-wise approach of the previous variant. The inclusion of assets and the expenditure data from recall methods on household-transportation further helps us explain how the expenditure on quality differs across various commodity groups within the population.

### 5.2.1 Variables in Use

The main control variables in the LA-AIDS SUR equations are `ln_tot_exp` (logarithm of the total expenditure in the period) and those starting with lp- (which represent the logarithm of the prices corresponding to every commodity in the AIDS regression - see Equations 9 and 10). As the quality differences are ignored for the expenditure on transport and household commodities, the AIDS SUR equations use household-CPI and transport-CPI as prices for these commodity groups i.e. `lphousehold` and `lptransport` (respectively). Other control variables that are used in the intercept of the AIDS SUR equations (see Equations 9 and 10) are `consu` (number of members adjusted by age), `age` (age of the household-head), `occupation_rank` (an ordered variable inferred from the occupation of the head-household), `educ_rank` (a ternary variable inferred from the highest education level of the head of the household) and `expensiveregion` ( indicator variable for densely populated urban areas). The variables in the intercept are subject to the additivity constraints (see Section 3.3 for details).

Since we view the accumulated assets worth as the household characteristics (i.e. as a part of $\rho$ in the Equation 6), the assets owned for the two groups - household and

Table 2: sureg for all items after the first-stage non-consumption probit (1/3)

| | (1) qbeef | (2) qbeer | (3) qbread | (4) qbunscakes | (5) qcassavaflour | (6) qcassavafresh | (7) qcharcoal | (8) qcoconut |
|---|---|---|---|---|---|---|---|---|
| age | 0.0005 | 0.0000 | 0.0001 | -0.0002 | -0.0001 | 0.0001 | -0.0002 | -0.0001 |
| | ( 0.0000$^{***}$ ) | ( 0.0020$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| cons | -0.3348 | -0.0985 | -0.0514 | -0.0326 | 0.0639 | -0.0334 | -0.0831 | -0.0495 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| consu | -0.0033 | -0.0002 | -0.0001 | 0.0005 | 0.0003 | -0.0001 | -0.0006 | 0.0000 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0040$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.7610 ) |
| educrank | 0.0174 | 0.0018 | 0.0230 | -0.0686 | -0.0127 | -0.0073 | -0.0249 | -0.0169 |
| | ( 0.0000$^{***}$ ) | ( 0.0600 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| invmills | 0.0001 | 0.0011 | 0.0003 | -0.0032 | 0.0007 | 0.0013 | -0.0005 | -0.0007 |
| | ( 0.9140 ) | ( 0.0000$^{***}$ ) | ( 0.0560 ) | ( 0.0000$^{***}$ ) | ( 0.0250$^{*}$ ) | ( 0.0000$^{***}$ ) | ( 0.1260 ) | ( 0.0010$^{**}$ ) |
| lntotexp | 0.0285 | 0.0065 | 0.0046 | 0.0073 | -0.0044 | 0.0021 | 0.0114 | 0.0056 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpbeef | -0.0548 | -0.0074 | 0.0161 | 0.0080 | 0.0025 | -0.0015 | -0.0089 | 0.0337 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpbeer | -0.0074 | -0.0023 | 0.0024 | -0.0021 | 0.0019 | -0.0017 | 0.0034 | -0.0027 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpbread | 0.0161 | 0.0024 | 0.0008 | -0.0029 | 0.0071 | -0.0028 | 0.0001 | 0.0042 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0010$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.7670 ) | ( 0.0000$^{***}$ ) |
| lpbunscakes | 0.0080 | -0.0021 | -0.0029 | -0.0118 | -0.0043 | 0.0016 | 0.0060 | -0.0040 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpcassavaflour | 0.0025 | 0.0019 | 0.0071 | -0.0043 | 0.0001 | -0.0016 | 0.0062 | 0.0032 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.8650 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpcassavafresh | -0.0015 | -0.0017 | -0.0028 | 0.0016 | -0.0016 | 0.0008 | -0.0021 | -0.0036 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpcharcoal | -0.0089 | 0.0034 | 0.0001 | 0.0060 | 0.0062 | -0.0021 | 0.0227 | 0.0032 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.7670 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpcoconut | 0.0337 | -0.0027 | 0.0042 | -0.0040 | 0.0032 | -0.0036 | 0.0032 | -0.0078 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpcookingoil | -0.0057 | 0.0009 | -0.0010 | 0.0014 | 0.0005 | 0.0053 | -0.0225 | 0.0055 |
| | ( 0.0000$^{***}$ ) | ( 0.0060$^{**}$ ) | ( 0.0020$^{**}$ ) | ( 0.0040$^{**}$ ) | ( 0.1230 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpdriedcannedfish | 0.0038 | -0.0016 | -0.0023 | 0.0032 | -0.0070 | -0.0009 | 0.0015 | -0.0007 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0010$^{**}$ ) |
| lpelectricity | 0.0006 | 0.0022 | -0.0039 | -0.0028 | 0.0059 | -0.0007 | 0.0078 | -0.0056 |
| | ( 0.5150 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0240$^{*}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpfishseafood | 0.0051 | 0.0041 | 0.0001 | -0.0039 | -0.0005 | 0.0031 | 0.0049 | -0.0099 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.7470 ) | ( 0.0000$^{***}$ ) | ( 0.0760 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpfreshmilk | -0.0018 | -0.0022 | 0.0043 | 0.0018 | -0.0006 | -0.0028 | 0.0111 | 0.0083 |
| | ( 0.0010$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0070$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpgreens | -0.0060 | 0.0010 | 0.0016 | -0.0027 | 0.0022 | 0.0004 | -0.0069 | 0.0074 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0010$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpkerosene | -0.0140 | 0.0029 | -0.0056 | 0.0033 | 0.0054 | -0.0013 | -0.0027 | 0.0020 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpmangoes | -0.0036 | -0.0017 | 0.0003 | -0.0017 | -0.0006 | -0.0005 | 0.0038 | 0.0038 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0260$^{*}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lponion | -0.0052 | -0.0037 | 0.0007 | 0.0042 | 0.0019 | 0.0028 | 0.0028 | 0.0008 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0030$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0030$^{**}$ ) |
| lppeanuts | -0.0032 | -0.0005 | -0.0025 | -0.0027 | -0.0008 | 0.0006 | 0.0013 | -0.0004 |
| | ( 0.0000$^{***}$ ) | ( 0.0010$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0220$^{*}$ ) |
| lppotatoes | 0.0092 | -0.0020 | -0.0001 | 0.0009 | 0.0002 | 0.0001 | -0.0030 | 0.0017 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.4210 ) | ( 0.0000$^{***}$ ) | ( 0.2440 ) | ( 0.2530 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lppulses | 0.0000 | -0.0002 | -0.0023 | 0.0020 | -0.0073 | -0.0005 | 0.0041 | -0.0057 |
| | ( 0.9880 ) | ( 0.5550 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0460$^{*}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpricehusked | 0.0503 | -0.0013 | -0.0050 | 0.0004 | -0.0092 | -0.0008 | -0.0139 | -0.0080 |
| | ( 0.0000$^{***}$ ) | ( 0.0080$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.5770 ) | ( 0.0000$^{***}$ ) | ( 0.0390$^{*}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpsalt | -0.0081 | 0.0016 | -0.0007 | -0.0013 | -0.0002 | 0.0013 | -0.0022 | -0.0009 |
| | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.1760 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpsugar | -0.0003 | 0.0094 | -0.0071 | 0.0073 | -0.0045 | 0.0048 | -0.0190 | -0.0236 |
| | ( 0.7210 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| lpsweetpotato | -0.0090 | -0.0005 | -0.0014 | 0.0002 | -0.0004 | -0.0001 | 0.0026 | -0.0007 |
| | ( 0.0000$^{***}$ ) | ( 0.0020$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.3670 ) | ( 0.0270$^{*}$ ) | ( 0.3110 ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |
| occupationrank | -0.0015 | -0.0003 | 0.0011 | 0.0011 | -0.0005 | 0.0015 | 0.0139 | 0.0034 |
| | ( 0.0000$^{***}$ ) | ( 0.0010$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0020$^{**}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) | ( 0.0000$^{***}$ ) |

*p*-values in parentheses

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 3: sureg for all items after the first-stage non-consumption probit (2/3)

| | (1) qcookingoil | (2) qdriedcannedfish | (3) qelectricity | (4) qfishseafood | (5) qfreshmilk | (6) qgreens | (7) qkerosene | (8) qmangoes |
|---|---|---|---|---|---|---|---|---|
| age | -0.0002 | -0.0002 | 0.0001 | 0.0001 | 0.0000 | -0.0001 | 0.0004 | -0.0001 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| cons | 0.3757 | 0.1711 | -0.1125 | 0.0766 | -0.0623 | 0.0536 | 0.5433 | -0.0182 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| consu | 0.0014 | 0.0014 | -0.0009 | 0.0003 | -0.0004 | -0.0004 | 0.0013 | -0.0003 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| educrank | 0.0120 | -0.0022 | 0.0735 | 0.0331 | 0.0215 | -0.0146 | 0.0243 | 0.0004 |
| | ( 0.0000*** ) | ( 0.2630 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.4690 ) |
| invmills | -0.0053 | -0.0001 | 0.0067 | -0.0032 | 0.0019 | -0.0009 | 0.0008 | 0.0011 |
| | ( 0.0000*** ) | ( 0.7150 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0310* ) | ( 0.0000*** ) |
| lntotexp | -0.0217 | -0.0107 | 0.0008 | -0.0027 | 0.0035 | 0.0007 | -0.0335 | 0.0010 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpbeef | -0.0057 | 0.0038 | 0.0006 | 0.0051 | -0.0018 | -0.0060 | -0.0140 | -0.0036 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.5150 ) | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpbeer | 0.0009 | -0.0016 | 0.0022 | 0.0041 | -0.0022 | 0.0010 | 0.0029 | -0.0017 |
| | ( 0.0060** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpbread | -0.0010 | -0.0023 | -0.0039 | 0.0001 | 0.0043 | 0.0016 | -0.0056 | 0.0003 |
| | ( 0.0020** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.7470 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0260* ) |
| lpbunscakes | 0.0014 | 0.0032 | -0.0028 | -0.0039 | 0.0018 | -0.0027 | 0.0033 | -0.0017 |
| | ( 0.0040** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcassavaflour | 0.0005 | -0.0070 | 0.0059 | -0.0005 | -0.0006 | 0.0022 | 0.0054 | -0.0006 |
| | ( 0.1230 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0760 ) | ( 0.0070** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcassavafresh | 0.0053 | -0.0009 | -0.0007 | 0.0031 | -0.0028 | -0.0004 | -0.0013 | -0.0005 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0240* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcharcoal | -0.0225 | 0.0015 | 0.0078 | 0.0049 | 0.0111 | -0.0069 | -0.0027 | 0.0038 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcoconut | 0.0055 | -0.0007 | -0.0056 | -0.0099 | 0.0083 | 0.0074 | 0.0020 | 0.0038 |
| | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcookingoil | 0.0136 | 0.0019 | -0.0085 | 0.0084 | -0.0013 | -0.0049 | 0.0005 | -0.0043 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0030** ) | ( 0.0000*** ) | ( 0.3900 ) | ( 0.0000*** ) |
| lpdriedcannedfish | 0.0019 | -0.0089 | 0.0075 | -0.0111 | 0.0003 | 0.0049 | 0.0043 | 0.0024 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1750 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpelectricity | -0.0085 | 0.0075 | 0.0265 | 0.0054 | 0.0051 | -0.0075 | -0.0066 | 0.0064 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpfishseafood | 0.0084 | -0.0111 | 0.0054 | -0.0181 | 0.0017 | 0.0017 | 0.0020 | 0.0003 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0930 ) |
| lpfreshmilk | -0.0013 | 0.0003 | 0.0051 | 0.0017 | -0.0046 | -0.0005 | 0.0043 | 0.0016 |
| | ( 0.0030** ) | ( 0.1750 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0160* ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpgreens | -0.0049 | 0.0049 | -0.0075 | 0.0017 | -0.0005 | 0.0004 | -0.0007 | -0.0005 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0160* ) | ( 0.0900 ) | ( 0.0300* ) | ( 0.0000*** ) |
| lpkerosene | 0.0005 | 0.0043 | -0.0066 | 0.0020 | 0.0043 | -0.0007 | -0.0041 | -0.0028 |
| | ( 0.3900 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0300* ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpmangoes | -0.0043 | 0.0024 | 0.0064 | 0.0003 | 0.0016 | -0.0005 | -0.0028 | 0.0018 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0930 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lponion | -0.0112 | 0.0034 | 0.0037 | 0.0028 | -0.0043 | -0.0021 | 0.0001 | -0.0015 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.8780 ) | ( 0.0000*** ) |
| lppeanuts | -0.0005 | -0.0010 | 0.0062 | -0.0001 | -0.0039 | 0.0023 | -0.0008 | 0.0005 |
| | ( 0.1200 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.5280 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) |
| lppotatoes | -0.0019 | 0.0016 | -0.0061 | 0.0024 | 0.0009 | 0.0015 | 0.0005 | -0.0007 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0290* ) | ( 0.0000*** ) |
| lppulses | -0.0067 | -0.0026 | 0.0098 | 0.0011 | -0.0029 | -0.0047 | 0.0015 | -0.0023 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0240* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0150* ) | ( 0.0000*** ) |
| lpricehusked | 0.0102 | 0.0028 | -0.0093 | -0.0029 | 0.0017 | 0.0078 | 0.0108 | 0.0046 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0080** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpsalt | 0.0015 | -0.0009 | -0.0010 | 0.0028 | -0.0017 | -0.0023 | -0.0013 | -0.0004 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0230* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpsugar | 0.0213 | 0.0002 | -0.0347 | 0.0006 | -0.0124 | 0.0053 | -0.0038 | -0.0054 |
| | ( 0.0000*** ) | ( 0.6450 ) | ( 0.0000*** ) | ( 0.2340 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpsweetpotato | -0.0025 | -0.0009 | -0.0006 | 0.0003 | -0.0021 | 0.0022 | 0.0062 | 0.0005 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1750 ) | ( 0.1180 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| occupationrank | -0.0070 | -0.0044 | 0.0037 | -0.0033 | 0.0016 | 0.0007 | 0.0030 | 0.0006 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |

*p*-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: sureg for all items after the first-stage non-consumption probit (3/3)

| | (1) qonion | (2) qpeanuts | (3) qpotatoes | (4) qpulses | (5) qricehusked | (6) qsalt | (7) qsugar | (8) qsweetpotato |
|---|---|---|---|---|---|---|---|---|
| age | -0.0001 | 0.0000 | -0.0001 | 0.0001 | -0.0003 | 0.0001 | 0.0002 | -0.0001 |
| | ( 0.0000*** ) | ( 0.0450* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| cons | 0.3803 | -0.0119 | -0.0409 | 0.0248 | -0.4139 | 0.2239 | 0.1990 | 0.0138 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| consu | 0.0002 | -0.0001 | -0.0003 | 0.0012 | -0.0006 | 0.0009 | 0.0015 | -0.0001 |
| | ( 0.0010** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| educrank | -0.0108 | -0.0039 | -0.0007 | -0.0161 | -0.0210 | 0.0177 | 0.0007 | -0.0080 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3210 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.7030 ) | ( 0.0000*** ) |
| invmills | -0.0075 | 0.0011 | 0.0005 | -0.0030 | -0.0065 | 0.0001 | -0.0042 | 0.0011 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.4090 ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lntotexp | -0.0195 | 0.0002 | 0.0045 | -0.0017 | 0.0356 | -0.0151 | -0.0040 | 0.0009 |
| | ( 0.0000*** ) | ( 0.0690 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpbeef | -0.0052 | -0.0032 | 0.0092 | 0.0000 | 0.0503 | -0.0081 | -0.0003 | -0.0090 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.9880 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.7210 ) | ( 0.0000*** ) |
| lpbeer | -0.0037 | -0.0005 | -0.0020 | -0.0002 | -0.0013 | 0.0016 | 0.0094 | -0.0005 |
| | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.5550 ) | ( 0.0080** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0020** ) |
| lpbread | 0.0007 | -0.0025 | -0.0001 | -0.0023 | -0.0050 | -0.0007 | -0.0071 | -0.0014 |
| | ( 0.0030** ) | ( 0.0000*** ) | ( 0.4210 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpbunscakes | 0.0042 | -0.0027 | 0.0009 | 0.0020 | 0.0004 | -0.0013 | 0.0073 | 0.0002 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.5770 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3670 ) |
| lpcassavaflour | 0.0019 | -0.0008 | 0.0002 | -0.0073 | -0.0092 | -0.0002 | -0.0045 | -0.0004 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2440 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1760 ) | ( 0.0000*** ) | ( 0.0270* ) |
| lpcassavafresh | 0.0028 | 0.0006 | 0.0001 | -0.0005 | -0.0008 | 0.0013 | 0.0048 | -0.0001 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2530 ) | ( 0.0460* ) | ( 0.0390* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3110 ) |
| lpcharcoal | 0.0028 | 0.0013 | -0.0030 | 0.0041 | -0.0139 | -0.0022 | -0.0190 | 0.0026 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcoconut | 0.0008 | -0.0004 | 0.0017 | -0.0057 | -0.0080 | -0.0009 | -0.0236 | -0.0007 |
| | ( 0.0030** ) | ( 0.0220* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpcookingoil | -0.0112 | -0.0005 | -0.0019 | -0.0067 | 0.0102 | 0.0015 | 0.0213 | -0.0025 |
| | ( 0.0000*** ) | ( 0.1200 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpdriedcannedfish | 0.0034 | -0.0010 | 0.0016 | -0.0026 | 0.0028 | -0.0009 | 0.0002 | -0.0009 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.6450 ) | ( 0.0000*** ) |
| lpelectricity | 0.0037 | 0.0062 | -0.0061 | 0.0098 | -0.0093 | -0.0010 | -0.0347 | -0.0006 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0230* ) | ( 0.0000*** ) | ( 0.1750 ) |
| lpfishseafood | 0.0028 | -0.0001 | 0.0024 | 0.0011 | -0.0029 | 0.0028 | 0.0006 | 0.0003 |
| | ( 0.0000*** ) | ( 0.5280 ) | ( 0.0000*** ) | ( 0.0240* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2340 ) | ( 0.1180 ) |
| lpfreshmilk | -0.0043 | -0.0039 | 0.0009 | -0.0029 | 0.0017 | -0.0017 | -0.0124 | -0.0021 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0080** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpgreens | -0.0021 | 0.0023 | 0.0015 | -0.0047 | 0.0078 | -0.0023 | 0.0053 | 0.0022 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpkerosene | 0.0001 | -0.0008 | 0.0005 | 0.0015 | 0.0108 | -0.0013 | -0.0038 | 0.0062 |
| | ( 0.8780 ) | ( 0.0010** ) | ( 0.0290* ) | ( 0.0150* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpmangoes | -0.0015 | 0.0005 | -0.0007 | -0.0023 | 0.0046 | -0.0004 | -0.0054 | 0.0005 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lponion | 0.0027 | -0.0009 | 0.0005 | 0.0041 | 0.0070 | -0.0002 | -0.0101 | 0.0016 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0070** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2450 ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lppeanuts | -0.0009 | -0.0026 | -0.0004 | 0.0002 | -0.0033 | 0.0016 | 0.0095 | 0.0013 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0280* ) | ( 0.6520 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lppotatoes | 0.0005 | -0.0004 | -0.0016 | 0.0011 | 0.0023 | -0.0011 | -0.0039 | -0.0021 |
| | ( 0.0070** ) | ( 0.0280* ) | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lppulses | 0.0041 | 0.0002 | 0.0011 | 0.0108 | 0.0112 | -0.0018 | -0.0056 | -0.0034 |
| | ( 0.0000*** ) | ( 0.6520 ) | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpricehusked | 0.0070 | -0.0033 | 0.0023 | 0.0112 | -0.0654 | 0.0070 | -0.0005 | 0.0034 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.6630 ) | ( 0.0000*** ) |
| lpsalt | -0.0002 | 0.0016 | -0.0011 | -0.0018 | 0.0070 | 0.0011 | 0.0053 | 0.0021 |
| | ( 0.2450 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpsugar | -0.0101 | 0.0095 | -0.0039 | -0.0056 | -0.0005 | 0.0053 | 0.0625 | 0.0047 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.6630 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpsweetpotato | 0.0016 | 0.0013 | -0.0021 | -0.0034 | 0.0034 | 0.0021 | 0.0047 | -0.0014 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| occupationrank | -0.0023 | -0.0009 | 0.0000 | -0.0035 | 0.0000 | 0.0003 | -0.0076 | -0.0003 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.5270 ) | ( 0.0000*** ) | ( 0.9600 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0010** ) |

*p*-values in parentheses

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

transport - are also used in the AIDS SUR equations i.e. the variable `all_assets_mtm` (which is the logarithm of total cost that the asset would reportedly sell for if it were to be sold at the time of the interview). We also use the variable `has_electric` (which indicates whether the household has access to electricity or not) in place of a quality metric for energy. Other formulations discussed in the Section 5.3 may use control variables based on the ownership of the types of assets - namely `electric_assets_mtm` (the logarithm of total cost of electrical assets - costs being what the asset woulds reportedly sell for if they were to be sold at the time of the interview), `household_assets_mtm` (the logarithm of total cost of land or the house that they would reportedly sell for if they were to be sold at the time of the interview) and `transport_assets_mtm` (the logarithm of total cost of the car, motorcycle, bicycle and other means of conveyance that they would reportedly sell for if they were to be sold at the time of the interview).

The columns in the tables 5 and 6 either start with q- or with qV-. The columns starting with q- indicate the budget shares ($w$) that are used as the dependent variables in AIDS regressions (see Equation 9 ) and the columns starting with qV- indicate the logarithm of the quality metric as the dependent variable (i.e. $ln\ V$ in the Equation 10). As explained earlier in Section 3.5, we use an indicator variable `has_electric` (see Table 5) for whether the consumer uses electricity or not instead of a quality metric for energy in the intercept of the AIDS equations. This is due to the wide differences in the characteristics of goods (kerosene, electricity and charcoal) within the energy commodity. In an alternate formulation in Section 5.3, a variable showing the total cost of electric assets is used as one of the household characteristics in lieu of the `has_electric` variable.

The $ln\ V$ equations are used only for the food commodity group. Using the classification based on prices and functionality, the commodities chosen within the food group are - densefoods, non-fresh, fruits-veg, protein and complements (see Table 9 all the goods that belong to these commodity groups). Since quality measure is not used for energy, household and transport no quality equations are needed for these commodity. The total number of equations to estimate with the SUR is therefore $2 \times 5 + 3 \times 1 = 13$. The results obtained from the SUR that are shown in Table 5 confirm many of the findings from the itemwise-probit but also suggest the relevance of asset ownership in the demand for non-durable items - their quality in particular.

#### 5.2.2 Results

As would be amenable to a permanent income approach, the consumer habits do seem to be influenced by the financial standing of the consumer that we have attempted to gauge with the occupation rank and assets owned by the consumers. The life-style differences that we observe in non-durable consumption - with respect to the quality of items (indicated with variables starting with `qV-`) do seem associated with household characteristics related to the long-term economic standing (and particularly on whether they use electricity or not).

To be specific, it is worth differentiating between the effect that ownership of assets has on quality vs the effect of total expenditure[20]. Those with electrical assets (the logarithm of whose total cost is used as the control variable) tend to spend more on fruits-veg quality. With higher expenditure (ln_tot_exp), the quality of non-fresh increases but the quality of fruits-veg decreases. Those with a stronger long-term economist standing spend relatively less on quality of protein and transport (public_transport, car maintenance etc.) - which are commodities where those who have electricity seem to spend more. It seems that just starting to use electricity (but not having lots of electric assets) can increase the quality of protein consumed for the average consumer (see the effect of the has_electric variable in the Table 6).

Running another SUR with assets split into electrical, household and transport assets confirm the same results (see Table 6). Here, as one would expect, the consumption on energy is high for electricity users due to the high prices of electricity in Tanzania. The owners of electrical assets also consume more fruits-veg quality and have more expenditure on household products. Here too, slightly lower fruitsveg quality is associated total expenditure (ln_tot_exp). While those with more assets spend less on some food quality categories, they do spend more on the household-related non-durable consumption.

The variable expensive_region is significant for non-fresh quality, protein quality, household commodity group, transport commodity group and the energy commodity group. Within urban areas, having electricity is associated with higher protein and fruits-veg quality. The urban consumers may have less overall assets compared to their counterparts in rural areas but they spend more on quality on protein and fruits-veg (less on dense-foods and household asset). The above observations suggest the life-style differences in the densely populated urban areas brought about by availability of industrial goods.

## 5.3 Robustness Checks

To ensure that the relevance of high asset ownership in differences of quality is a robust observation, we rerun the SUR with an aggregated food category. This formulation no longer uses the price-based groups and views all food-subcategories as one food group[21] . Using the CPI of food basket as prices (just the way transport CPI and household CPI are used) the above analysis is repeated with food, energy, transport and household as separable categories. In the results shown in Tables 7 and 8, a preference for higher quality for electricity users and household asset owners is observed when we use a food aggregate (a collection of the subcategories). Food quality increases with total expenditure and with acquisition of household assets. As one would expect, transport costs are higher in the urban areas. The non-durable expenditure on household items and overall food quality both decrease for urban areas.

As a further test of the robustness of results, we calculate the food-quality using the market-provided prices. More specifically, the results shown in Tables 5 and 6 we have used diary-costs on goods as reported by the consumer. However, the LSMS records this cost only from the items that were consumed from purchases in the last

[20]The literature often uses total expenditure as a measure of permanent income.
[21]The quality is still calculated based on the prices.

week. If a consumer didn't purchase anything in the past week, there would only be quantities consumed (not costs) that would be reported for the consumer in the diary. Since we have the market-prices available, another way to calculate costs, budget-shares and quality would be to infer costs as the multiple of market-prices from the quantities consumed in the past week (which is reported for all consumers regardless of whether they had purchased something in the last week or not). Recalculating quality and budget shares still shows us the same economic and statistical significance (results available on request) of electricity and household assets in quality of fruitsveg, proteins and other subcategories.

It is worth reiterating that the differences in quality should be seen as more than just response to high prices in the urban areas. As our quality metric is derived from locally available prices (the base price $P_G$ can be different in every district), it already accounts for regional differences in prices. One possible explanation is that the focus on industrial goods and meat categories is stronger in urban areas and is maintained as consumer habits. A comparison of quality differences amongst the population in other countries could verify this argument i.e. if quality differences for a certain commodity groups are found significant with respect to the long-term socio-economic standing of the consumer in other markets.

## Table 5: sureg with electricity usage indicator variables and cost of all assets owned

| | (1) qcomplements | (2) qdensefoods | (3) qenergy | (4) qfruitsveg | (5) qhousehold | (6) qnonfresh | (7) qprotein | (8) qtransport | (9) qVcomplements | (10) qVdensefoods | (11) qVfruitsveg | (12) qVnonfresh | (13) qVprotein |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| _cons | 0.0156 | 0.0646 | 0.4814 | 0.0807 | 0.5407 | 0.1278 | 0.0573 | -0.5270 | 5.7687 | 6.5246 | 6.8667 | 6.3753 | 8.3453 |
| | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) |
| age | 0.0000 | 0.0000 | 0.0001 | 0.0001 | -0.0007 | 0.0001 | 0.0001 | -0.0001 | -0.0010 | -0.0006 | -0.0003 | -0.0005 | 0.0001 |
| | (0.0110*) | (0.0000***) | (0.2460) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0230*) | (0.2990) | (0.0220*) | (0.2130) | (0.5360) | (0.1430) | (0.8340) |
| all_assets_mtm | 0.0001 | 0.0000 | 0.0030 | 0.0003 | 0.0024 | 0.0005 | 0.0001 | -0.0050 | 0.0187 | 0.0030 | 0.0009 | -0.0007 | -0.0091 |
| | (0.0000***) | (0.5560) | (0.0000***) | (0.0000***) | (0.0080**) | (0.0000***) | (0.5670) | (0.0000***) | (0.0000***) | (0.4130) | (0.7750) | (0.7690) | (0.0420*) |
| consu | 0.0000 | 0.0003 | 0.0002 | 0.0002 | -0.0043 | 0.0003 | -0.0003 | -0.0030 | -0.0039 | 0.0100 | 0.0013 | -0.0034 | 0.0051 |
| | (0.0000***) | (0.0000***) | (0.5570) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0150*) | (0.0000***) | (0.0290*) | (0.0000***) | (0.4260) | (0.0080**) | (0.0250*) |
| educ_rank | 0.0000 | 0.0024 | 0.0166 | 0.0041 | -0.0124 | 0.0010 | 0.0039 | -0.0165 | -0.2086 | -0.1167 | -0.0438 | 0.0316 | -0.1846 |
| | (0.8730) | (0.0730) | (0.1790) | (0.0050**) | (0.4140) | (0.6550) | (0.2600) | (0.2160) | (0.0000***) | (0.0290*) | (0.3750) | (0.4290) | (0.0040**) |
| expensiveregion | -0.0003 | -0.0004 | 0.0441 | -0.0018 | -0.0313 | 0.0003 | 0.0006 | 0.0108 | -0.0338 | -0.0345 | 0.0119 | 0.0473 | 0.0465 |
| | (0.0000***) | (0.2460) | (0.0000***) | (0.0000***) | (0.0000***) | (0.5380) | (0.4210) | (0.0000***) | (0.0130*) | (0.0130*) | (0.3480) | (0.0000***) | (0.0080**) |
| has_electric | 0.0000 | -0.0005 | 0.0593 | 0.0002 | -0.0220 | 0.0001 | 0.0002 | -0.0084 | 0.0444 | 0.0283 | 0.0748 | 0.0764 | 0.0693 |
| | (0.5570) | (0.0870) | (0.0000***) | (0.4830) | (0.0000***) | (0.8990) | (0.8090) | (0.0070**) | (0.0020**) | (0.0470*) | (0.0000***) | (0.0000***) | (0.0000***) |
| ln_tot_exp | -0.0012 | -0.0046 | -0.0328 | -0.0062 | -0.0167 | -0.0088 | -0.0034 | 0.0482 | 0.0681 | -0.0039 | -0.0267 | 0.0148 | -0.0268 |
| | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.5800) | (0.0000***) | (0.0050**) | (0.0010**) |
| lpcomplements | 0.0003 | 0.0003 | -0.0001 | -0.0001 | -0.0005 | 0.0001 | -0.0001 | 0.0001 | 0.2108 | -0.0100 | -0.0253 | 0.0352 | -0.0578 |
| | (0.0000***) | (0.0000***) | (0.2850) | (0.0670) | (0.0290*) | (0.3030) | (0.3790) | (0.6940) | (0.0000***) | (0.5330) | (0.0790) | (0.0020**) | (0.0040**) |
| lpdensefoods | 0.0003 | 0.0012 | 0.0008 | -0.0004 | 0.0025 | -0.0001 | -0.0006 | -0.0037 | 0.0569 | 0.2620 | 0.0269 | 0.0025 | 0.0209 |
| | (0.0000***) | (0.0000***) | (0.0010**) | (0.1030) | (0.0100*) | (0.6270) | (0.0520) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0390*) | (0.8030) | (0.2460) |
| lpenergy | -0.0001 | 0.0008 | -0.0092 | 0.0008 | 0.0074 | -0.0006 | -0.0007 | 0.0015 | -0.0208 | -0.0166 | -0.0392 | -0.0353 | -0.0615 |
| | (0.2850) | (0.0010**) | (0.0000***) | (0.0060**) | (0.0060**) | (0.1880) | (0.2800) | (0.5220) | (0.0690) | (0.1540) | (0.0000***) | (0.0000***) | (0.0000***) |
| lpfruitsveg | -0.0001 | -0.0004 | 0.0008 | 0.0004 | 0.0006 | 0.0020 | 0.0007 | -0.0040 | 0.0046 | 0.0070 | 0.1444 | 0.0209 | -0.1053 |
| | (0.0670) | (0.1030) | (0.0060**) | (0.2570) | (0.5540) | (0.0000***) | (0.0360*) | (0.0000***) | (0.7420) | (0.6240) | (0.0000***) | (0.0370*) | (0.0000***) |
| lphousehold | -0.0005 | 0.0025 | 0.0074 | 0.0006 | 0.0791 | 0.0072 | -0.0013 | -0.0950 | 0.2656 | 0.5604 | 0.7699 | 0.6729 | 0.7941 |
| | (0.0290*) | (0.0100*) | (0.0060**) | (0.5540) | (0.0000***) | (0.0000***) | (0.5120) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) |
| lpnonfresh | 0.0001 | -0.0001 | -0.0006 | 0.0020 | 0.0072 | -0.0005 | -0.0008 | -0.0073 | -0.0383 | -0.0507 | 0.0318 | 0.1653 | -0.0495 |
| | (0.3030) | (0.6270) | (0.1880) | (0.0000***) | (0.0000***) | (0.4250) | (0.1040) | (0.0000***) | (0.0250*) | (0.0040**) | (0.0430*) | (0.0000***) | (0.0230*) |
| lpprotein | -0.0001 | -0.0006 | -0.0007 | 0.0007 | -0.0013 | -0.0008 | -0.0012 | 0.0039 | 0.0582 | 0.1135 | 0.0892 | 0.1099 | 0.3244 |
| | (0.3790) | (0.0520) | (0.2800) | (0.0360*) | (0.5120) | (0.1040) | (0.1650) | (0.0390*) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) |
| lptransport | 0.0001 | -0.0037 | 0.0015 | -0.0040 | -0.0950 | -0.0073 | 0.0039 | 0.1046 | -0.5371 | -0.8657 | -0.9977 | -0.9715 | -0.8653 |
| | (0.6940) | (0.0000***) | (0.5220) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0390*) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) | (0.0000***) |
| occupation_rank | 0.0001 | 0.0000 | 0.0118 | -0.0001 | -0.0064 | -0.0003 | -0.0001 | 0.0001 | 0.0419 | 0.0364 | 0.0131 | 0.0152 | 0.0158 |
| | (0.0260*) | (0.9530) | (0.0000***) | (0.4280) | (0.0000***) | (0.1920) | (0.6760) | (0.9140) | (0.0000***) | (0.0000***) | (0.0190*) | (0.0000***) | (0.0400*) |

*p*-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: sureg with price-based food subcategories and long-term assets

| | (1) qcomplements | (2) qdensefoods | (3) qenergy | (4) qfruitsveg | (5) qhousehold | (6) qnonfresh | (7) qprotein | (8) qtransport | (9) qVcomplements | (10) qVdensefoods | (11) qVfruitsveg | (12) qVnonfresh | (13) qVprotein |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| _cons | 0.0155 | 0.0594 | 0.4358 | 0.0732 | 0.7115 | 0.1205 | 0.0477 | -0.5875 | 6.0705 | 6.5670 | 6.8900 | 6.2708 | 7.9214 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| age | 0.0000 | 0.0000 | 0.0003 | 0.0001 | -0.0007 | 0.0001 | 0.0001 | -0.0002 | -0.0008 | -0.0005 | -0.0001 | -0.0004 | 0.0001 |
| | ( 0.0020** ) | ( 0.0000*** ) | ( 0.0060** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0120* ) | ( 0.0420* ) | ( 0.0620 ) | ( 0.2750 ) | ( 0.8050 ) | ( 0.2560 ) | ( 0.9100 ) |
| consu | 0.0000 | 0.0003 | 0.0003 | 0.0002 | -0.0040 | 0.0003 | -0.0003 | -0.0033 | -0.0029 | 0.0105 | 0.0012 | -0.0033 | 0.0038 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3940 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0120* ) | ( 0.0000*** ) | ( 0.1010 ) | ( 0.0000*** ) | ( 0.4590 ) | ( 0.0100* ) | ( 0.0960 ) |
| educ_rank | 0.0002 | 0.0024 | 0.0638 | 0.0054 | -0.0429 | 0.0017 | 0.0054 | -0.0231 | -0.1736 | -0.0710 | 0.0013 | 0.1072 | -0.0998 |
| | ( 0.5080 ) | ( 0.0630 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0040** ) | ( 0.4540 ) | ( 0.1110 ) | ( 0.0760 ) | ( 0.0000*** ) | ( 0.1740 ) | ( 0.9780 ) | ( 0.0060** ) | ( 0.1050 ) |
| electric_assets_mtm | 0.0000 | -0.0002 | 0.0025 | -0.0002 | 0.0035 | -0.0003 | -0.0002 | -0.0021 | 0.0167 | 0.0086 | 0.0081 | 0.0046 | -0.0030 |
| | ( 0.8760 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0080** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0960 ) |
| expensiveregion | -0.0003 | -0.0003 | 0.0545 | -0.0017 | -0.0350 | 0.0007 | 0.0006 | 0.0098 | -0.0372 | -0.0426 | 0.0222 | 0.0514 | 0.0531 |
| | ( 0.0000*** ) | ( 0.2500 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1530 ) | ( 0.4340 ) | ( 0.0010** ) | ( 0.0060** ) | ( 0.0020** ) | ( 0.0780 ) | ( 0.0000*** ) | ( 0.0020** ) |
| household_assets_mtm | 0.0001 | 0.0000 | 0.0011 | 0.0002 | 0.0003 | 0.0003 | 0.0000 | -0.0026 | 0.0126 | 0.0025 | -0.0013 | -0.0001 | -0.0044 |
| | ( 0.0000*** ) | ( 0.7660 ) | ( 0.0660 ) | ( 0.0050** ) | ( 0.6600 ) | ( 0.0090** ) | ( 0.8240 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3940 ) | ( 0.6210 ) | ( 0.9470 ) | ( 0.2250 ) |
| ln_tot_exp | -0.0011 | -0.0041 | -0.0304 | -0.0055 | -0.0281 | -0.0080 | -0.0026 | 0.0512 | 0.0465 | -0.0074 | -0.0318 | 0.0205 | 0.0007 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3270 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.9340 ) |
| lpcomplements | 0.0003 | 0.0002 | 0.0000 | -0.0001 | -0.0006 | 0.0001 | -0.0001 | 0.0003 | 0.2170 | -0.0058 | -0.0278 | 0.0319 | -0.0638 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3390 ) | ( 0.0760 ) | ( 0.0040** ) | ( 0.3280 ) | ( 0.1740 ) | ( 0.2180 ) | ( 0.0000*** ) | ( 0.7180 ) | ( 0.0530 ) | ( 0.0040** ) | ( 0.0010** ) |
| lpdensefoods | 0.0002 | 0.0013 | 0.0008 | -0.0003 | 0.0020 | 0.0000 | -0.0006 | -0.0034 | 0.0434 | 0.2451 | 0.0255 | -0.0066 | 0.0158 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0020** ) | ( 0.1290 ) | ( 0.0370* ) | ( 0.9750 ) | ( 0.0650 ) | ( 0.0000*** ) | ( 0.0020** ) | ( 0.0000*** ) | ( 0.0520 ) | ( 0.5160 ) | ( 0.3850 ) |
| lpenergy | 0.0000 | 0.0008 | -0.0089 | 0.0007 | 0.0080 | -0.0006 | -0.0006 | 0.0007 | -0.0151 | -0.0143 | -0.0371 | -0.0351 | -0.0633 |
| | ( 0.3390 ) | ( 0.0020** ) | ( 0.0000*** ) | ( 0.0120* ) | ( 0.0030** ) | ( 0.1800 ) | ( 0.3130 ) | ( 0.7720 ) | ( 0.1830 ) | ( 0.2160 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpfruitsveg | -0.0001 | -0.0003 | 0.0007 | 0.0004 | -0.0005 | 0.0021 | 0.0008 | -0.0030 | 0.0040 | 0.0066 | 0.1431 | 0.0183 | -0.1070 |
| | ( 0.0760 ) | ( 0.1290 ) | ( 0.0120* ) | ( 0.2580 ) | ( 0.6380 ) | ( 0.0000*** ) | ( 0.0190* ) | ( 0.0040** ) | ( 0.7720 ) | ( 0.6450 ) | ( 0.0000*** ) | ( 0.0680 ) | ( 0.0000*** ) |
| lphousehold | -0.0006 | 0.0020 | 0.0080 | -0.0005 | 0.0828 | 0.0060 | -0.0029 | -0.0948 | 0.2654 | 0.5445 | 0.7743 | 0.6612 | 0.7689 |
| | ( 0.0040** ) | ( 0.0370* ) | ( 0.0030** ) | ( 0.6380 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1400 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpnonfresh | 0.0001 | 0.0000 | -0.0006 | 0.0021 | 0.0060 | -0.0004 | -0.0005 | -0.0066 | -0.0412 | -0.0481 | 0.0317 | 0.1700 | -0.0423 |
| | ( 0.3280 ) | ( 0.9750 ) | ( 0.1800 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.4810 ) | ( 0.2670 ) | ( 0.0000*** ) | ( 0.0150* ) | ( 0.0060** ) | ( 0.0430* ) | ( 0.0000*** ) | ( 0.0520 ) |
| lpprotein | -0.0001 | -0.0006 | -0.0006 | 0.0008 | -0.0029 | -0.0005 | -0.0012 | 0.0052 | 0.0451 | 0.1015 | 0.0904 | 0.1063 | 0.3251 |
| | ( 0.1740 ) | ( 0.0650 ) | ( 0.3130 ) | ( 0.0190* ) | ( 0.1400 ) | ( 0.2670 ) | ( 0.1450 ) | ( 0.0060** ) | ( 0.0030** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lptransport | 0.0003 | -0.0034 | 0.0007 | -0.0030 | -0.0948 | -0.0066 | 0.0052 | 0.1016 | -0.5187 | -0.8295 | -1.0003 | -0.9459 | -0.8333 |
| | ( 0.2180 ) | ( 0.0000*** ) | ( 0.7720 ) | ( 0.0040** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0060** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| occupation_rank | 0.0000 | 0.0001 | 0.0144 | 0.0000 | -0.0095 | -0.0002 | -0.0001 | 0.0009 | 0.0325 | 0.0296 | 0.0123 | 0.0152 | 0.0177 |
| | ( 0.0910 ) | ( 0.5610 ) | ( 0.0000*** ) | ( 0.8960 ) | ( 0.0000*** ) | ( 0.5100 ) | ( 0.8060 ) | ( 0.5000 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0280* ) | ( 0.0000*** ) | ( 0.0210* ) |
| transport_assets_mtm | 0.0000 | 0.0000 | 0.0003 | 0.0000 | 0.0005 | 0.0002 | 0.0000 | 0.0000 | -0.0035 | -0.0058 | 0.0009 | -0.0035 | -0.0019 |
| | ( 0.3340 ) | ( 0.4450 ) | ( 0.1100 ) | ( 0.1670 ) | ( 0.0230* ) | ( 0.0000*** ) | ( 0.5880 ) | ( 0.8210 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.3110 ) | ( 0.0000*** ) | ( 0.1240 ) |

*p*-values in parentheses

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 7: sureg with food categories aggregated and with all assets

| | (1) qenergy | (2) qfood | (3) qhousehold | (4) qtransport | (5) qVfood |
|---|---|---|---|---|---|
| _cons | 0.4764 | 0.3326 | 0.5381 | -0.5351 | 1.9663 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| age | 0.0003 | 0.0002 | -0.0006 | -0.0002 | 0.0012 |
| | ( 0.0040** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0480* ) | ( 0.0810 ) |
| all_assets_mtm | 0.0033 | 0.0020 | 0.0020 | -0.0049 | 0.0261 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0120* ) | ( 0.0000*** ) | ( 0.0000*** ) |
| consu | 0.0007 | 0.0009 | -0.0049 | -0.0033 | 0.0185 |
| | ( 0.0700 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| educ_rank | 0.0389 | 0.0131 | 0.0000 | -0.0277 | 0.0417 |
| | ( 0.0010** ) | ( 0.0040** ) | ( 0.9990 ) | ( 0.0260* ) | ( 0.1180 ) |
| expensiveregion | 0.0393 | -0.0026 | -0.0286 | 0.0111 | -0.1164 |
| | ( 0.0000*** ) | ( 0.0120* ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| has_electric | 0.0555 | -0.0013 | -0.0206 | -0.0073 | -0.0315 |
| | ( 0.0000*** ) | ( 0.2630 ) | ( 0.0000*** ) | ( 0.0120* ) | ( 0.1300 ) |
| ln_tot_exp | -0.0341 | -0.0241 | -0.0157 | 0.0486 | 0.0252 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpenergy | -0.0089 | -0.0003 | 0.0083 | 0.0006 | -0.0500 |
| | ( 0.0000*** ) | ( 0.7020 ) | ( 0.0010** ) | ( 0.7860 ) | ( 0.0030** ) |
| lphousehold | 0.0083 | 0.0206 | 0.0910 | -0.0993 | -0.0807 |
| | ( 0.0010** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.1550 ) |
| lptransport | 0.0006 | -0.0202 | -0.0993 | 0.0988 | 0.1307 |
| | ( 0.7860 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0210* ) |
| occupation_rank | 0.0123 | -0.0009 | -0.0056 | -0.0008 | -0.0227 |
| | ( 0.0000*** ) | ( 0.0680 ) | ( 0.0000*** ) | ( 0.5350 ) | ( 0.0100* ) |

*p*-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

## 5.4 Discussion

We strive to understand the effect of supply side differences in the market before implying what factors may influence the status consumption by the consumer. The availability of market-prices and the details on household characteristics allow us to understand the factors that may cause the differences in the consumer baskets (i.e. what we refer to as the supply-side concerns) and influence consumption of quality (and hence status consumption). We believe that observing the differences in the consumer baskets spread across an unevenly developing economy is necessary before one may draw conclusions about high-quality consumption in the economy - hence the inclusion of the long-term characteristics (regional differences, degree of urbanisation, ownership of long-term assets etc. or long-term economic standing) in the utility model for consumer's choice of quality.

Given the differences in consumer baskets for the urban and rural consumer, it should not be surprising that those in the urban areas spend differently on certain non-durable items. Nevertheless, the sustained quality differences for meat and fresh goods for urban consumers indicate a peculiar environmental difference in the urban areas - particularly sensitive to the distribution of electricity and availability of public transport. We find that it is not just that those with higher asset worth tend to buy more

Table 8: sureg with food categories aggregated and with separated assets

| | (1)<br>qenergy | (2)<br>qfood | (3)<br>qhousehold | (4)<br>qtransport | (5)<br>qVfood |
|---|---|---|---|---|---|
| _cons | 0.4392 | 0.3102 | 0.6959 | -0.5893 | 1.9310 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| age | 0.0004 | 0.0002 | -0.0006 | -0.0003 | 0.0007 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0050** ) | ( 0.2880 ) |
| consu | 0.0007 | 0.0009 | -0.0046 | -0.0035 | 0.0170 |
| | ( 0.0770 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| educ_rank | 0.0820 | 0.0156 | -0.0280 | -0.0327 | 0.0143 |
| | ( 0.0000*** ) | ( 0.0010** ) | ( 0.0450* ) | ( 0.0070** ) | ( 0.5510 ) |
| electric_assets_mtm | 0.0021 | -0.0009 | 0.0035 | -0.0020 | -0.0024 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2140 ) |
| expensiveregion | 0.0493 | -0.0017 | -0.0335 | 0.0109 | -0.1160 |
| | ( 0.0000*** ) | ( 0.0940 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| household_assets_mtm | 0.0015 | 0.0014 | 0.0003 | -0.0030 | 0.0294 |
| | ( 0.0070** ) | ( 0.0000*** ) | ( 0.5920 ) | ( 0.0000*** ) | ( 0.0000*** ) |
| ln_tot_exp | -0.0319 | -0.0215 | -0.0270 | 0.0519 | 0.0284 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) |
| lpenergy | -0.0087 | -0.0004 | 0.0086 | 0.0001 | -0.0496 |
| | ( 0.0000*** ) | ( 0.6630 ) | ( 0.0000*** ) | ( 0.9570 ) | ( 0.0030** ) |
| lphousehold | 0.0086 | 0.0172 | 0.0907 | -0.0993 | -0.0689 |
| | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.2220 ) |
| lptransport | 0.0001 | -0.0168 | -0.0993 | 0.0992 | 0.1185 |
| | ( 0.9570 ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0000*** ) | ( 0.0350* ) |
| occupation_rank | 0.0150 | -0.0005 | -0.0092 | 0.0001 | -0.0245 |
| | ( 0.0000*** ) | ( 0.2700 ) | ( 0.0000*** ) | ( 0.9110 ) | ( 0.0060** ) |
| transport_assets_mtm | 0.0004 | 0.0002 | 0.0004 | 0.0000 | 0.0016 |
| | ( 0.0650 ) | ( 0.0010** ) | ( 0.0860 ) | ( 0.9860 ) | ( 0.2590 ) |

*p*-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- but the quality purchased for a few categories always appears higher (lower) for the consumers who own medium-cost assets and lower (higher) for consumers who don't.

Within the urban areas, the households that don't own assets would benefit from spending more on industrial goods than an equivalent consumer in the rural area who does not face the same choice of industrial goods. Those who own more assets in the rural areas may not feel the status competitions the way the asset owning consumer in an urban area would. The consumer in the urban areas may be willing to sacrifice quality in one group to avail it in another - in a way that the rural consumer cannot. The role of long-term assets and residence is thus considered in our discussion of status consumption. The segmentation of quality based on the long-term economic standing need not be due to a fundamental shift in demand for quality in the urban areas. Such a segregation of quality (in food or in other non-durable commodities) implies just that pricier meats or electrical items are made available to the urban food consumer - that may let her deprioritse the quality in other commodities. Therefore, instead of arguing whether electricity or electrical items have signaling qualities or not, we propose that the idea of relative quality of common commodities (and under common needs) be considered in a discussion of status consumption. A preference order for consumption of quality - that can exist as long term habits for the consumer - may provide us a more universal extension to the idea of signaling goods for developing markets. It also addresses the apparent incongruence of visible consumption with the functional separability of utility.

Indeed as we observe relatively higher consumption on household items for those with better occupations, a certain degree of compensation (or pecuniary emulation) cannot be ruled out in the data[22] either. However, such an observation on visible items (in the household commodity group) would easily fit within the framework of quality and the socio-economic standing which we propose. There is after all more to status consumption than compensation[23] - which can fall within a wider substitution between consuming quality and planning for longer-term assets (including education etc.). In terms of the model for quality and socio-economic standing, a consumer doesn't compensate or emulate the behaviour of the rich as much as she chooses which commodity she can improve quality on - subject to her needs and her budget.

In conclusion, instead of focusing on the so-called "signaling" items that those with perceived lesser status would consume - we have set ourselves to exploring the factors that may influence the perception of status itself. The choice for quality may be influenced by the owned long-term assets - which along with the commodity quality chosen drive the perception of status in the model. There are more reasons to consider these differences in the developing countries where inequalities are wider. Further, if it is possible to model the acquisition of assets over time, an intertemporal setting could enhance our understanding of how quality of expenditure changes over time - allowing us to provide more robust insights on how long-term utility inclusive of status is indicated through consumption.

[22]Notice that houses are included in the household-assets. The observation that home-owners spend differently on non-durable consumption in the household category is also included as part of the general observations on assets-ownership.

[23]See a recent paper by Chai[24] which suggests the consumers may be comparing themselves more often to the individuals in the same income group rather than those above them

Table 9: Items classified according to price groups and function

| group | item |
|---|---|
| nonfresh | cooking_oil |
| nonfresh | rice_husked |
| nonfresh | rice_paddy |
| nonfresh | maize_green |
| nonfresh | maize_grain |
| nonfresh | maize_flour |
| nonfresh | millet_grain |
| nonfresh | millet_flour |
| nonfresh | wheat |
| nonfresh | bread |
| densefoods | bunscakes |
| densefoods | pasta |
| densefoods | othercereal |
| densefoods | pulses |
| densefoods | cassava_flour |
| densefoods | sweet_potato |
| densefoods | yam |
| densefoods | potatoes |
| densefoods | othervegstarch |
| fruitsveg | peanuts |
| fruitsveg | coconut |
| fruitsveg | cashew_almonds |
| fruitsveg | nut_products |
| fruitsveg | eggs |
| fruitsveg | onion |
| fruitsveg | greens |
| fruitsveg | dried_canned_veg |
| fruitsveg | cassava_fresh |
| fruitsveg | banana_ripe |
| fruitsveg | citrus |
| fruitsveg | mangoes |
| fruitsveg | sugarcane |
| fruitsveg | banana_green |
| protein | milk_products |
| protein | fresh_milk |
| protein | canned_milk |
| protein | goat |
| protein | beef |
| protein | pork |
| protein | chicken |
| protein | wild_birds |
| protein | wild_meat |
| protein | fish_seafood |

Table 9: Items classified according to price groups and function

| group | item |
|---|---|
| protein | dried_canned_fish |
| protein | packaged_fish |
| alcohol | beer |
| alcohol | brews |
| alcohol | winespirits |
| complements | sugar |
| complements | tea |
| complements | coffee |
| complements | miscdrinkpowder |
| complements | canned_drink |
| complements | readymade_tea_coffee |
| complements | sweet |
| complements | honey |
| complements | spices |
| complements | salt |
| energy | kerosene |
| energy | charcoal |
| energy | gas |
| energy | electricity |
| energy | petrol |
| transport | public_transport |
| transport | petrol |
| household | cigarettes |
| household | cellphone_voucher |
| household | bar_soap |
| household | clothes_soap |
| household | toothpaste |
| household | toilet_paper |
| household | skin_cream |
| household | other_personal |
| household | misc_cleaning |
| household | household_products_repair |
| household | consumer_durables_repair |
| household | mensclothes |
| household | womensclothes |
| household | childrensclothes |
| household | mensshoes |
| household | womensshoes |
| household | childrensshoes |
| household | services |
| household | bride_price |
| household | marriage |
| household | funeral |

Table 9: Items classified according to price groups and function

| group | item |
|---|---|
| household | light_bulbs |
| household | carpet |
| household | linen |
| household | mat |
| household | mosquito_net |
| household | mattress |
| household | building_material |
| household | insurance |
| household | bamboo |
| household | grass |

# References


[1] A. Deaton and J. Muellbauer, "An almost ideal demand system," *The American Economic Review*, vol. 70, pp. 312–326, Jun 1980.

[2] C. McKelvey, "Price, unit value, and quality demanded," *Journal of Development Economics*, vol. 95, pp. 157–169, 2011.

[3] M. Andalon and J. Gibson, "The 'soda tax' is unlikely to make mexicans lighter: New evidence on biases in elasticities of demand for soda," *Institute of Labour Economics - DISCUSSION PAPER SERIES*, vol. 10765, 2017.

[4] T. Veblen, *Theory of the Leisure Class*. 1899.

[5] T. Ustuner and D. B. Holt, "Toward a theory of status consumption in less industrialized countries," *Journal of Consumer Research*, vol. 37, pp. 37–56, June 2010.

[6] B. Galobardes, J. Lynch, George, and D. Smith, "Measuring socioeconomic position in health research," *British Medical Bulletin*, vol. 81-82, no. 1, pp. 21–37, 2007.

[7] S. J. Prais and H. S. Houthakker, *The analysis of family budgets*. Cambridge University Press, 1955.

[8] F. M. Fisher and K. Shell, *Value, Capital and Growth*. Transaction Publishers, 1968.

[9] J. Gibson and B. Kim, "Quality, quantity, and spatial variation of price: Back to the bog," *Journal of Development Economics*, vol. 137, pp. 66–77, 2019.

[10] A. P. Barten, "Consumer demand functions under conditions of almost additive preferences," *Econometrica*, vol. 32, Jan-Apr 1964.

[11] W. M. Gorman, "Community preference fields," *Econometrica*, vol. 21, pp. 63–80, January 1953.

[12] J. Muellbauer, "Aggregation, income distribution and consumer demand," *The Review of Economic Studies*, vol. 42, pp. 525–543, October 1975.

[13] K. Tafere, A. S. Taffesse, S. Tamiru, N. Tefera, and Z. Paulos, "Food demand elasticities in ethiopia: Estimates using household income consumption expenditure (hice) survey data," *International Food Policy Research Institute - ESSP2 Discussion Paper 011*, 2010. https://www.africaportal.org/publications/food-demand-elasticities-in-ethiopia-estimates-using-household-income-consumption-expenditure-hice-survey-data/.

[14] A. Deaton, "Quality, quantity, and spatial variation of price," *The American Economic Review*, vol. 78, pp. 418–430, June 1988.

[15] A. Deaton, “Estimation of own and cross-price elasticities from household survey data,” *Journal of Econometrics*, vol. 36, 1987.

[16] R. A. Pollak and T. J. Wales, “Demographic variables in demand analysis,” *Econometrica*, vol. 49, pp. 1533–1551, Nov 1981.

[17] G. Anderson and R. Blundell, “Testing restrictions in a flexible dynamic demand system: An application to consumers’ expenditure in canada,” *The Review of Economic Studies*, vol. 50, pp. 397–410, Jul 1983.

[18] G. Anderson and R. Blundell, “Estimation and hypothesis testing in dynamic singular equation systems,” *Econometrica*, vol. 50, Nov 1982.

[19] D. Heien and C. R. Wessells, “Demand systems estimation with microdata: A censored regression approach,” *Journal of Business and Economic Statistics*, vol. 8, pp. 365–371, Jul 1990.

[20] T. L. Cox and M. K. Wohlgenant, “Prices and quality effects in cross-sectional demand analysis,” *American Journal of Agricultural Economics*, vol. 68, pp. 908–919, November 1986.

[21] N. J. Ireland, “On limiting the market for status signals,” *Journal of Public Economics*, 1994.

[22] E. Hopkins and T. Kornienko, “Running to keep in the same place: Consumer choice as a game of status,” *The American Economic Review*, vol. 24, pp. 1085–1107, Sep 2004.

[23] M. Khamis, N. Prakash, and Z. Siddique, “Consumption and social identity,” 2010.

[24] A. Chai, W. Kaus, and C. Kiedaisch, “Conspicuous spending and the income distribution of social groups,” *Economic Inquiry*, vol. 57, pp. 1324–1341, July 2019.

[25] J. S. Cramer, “Interaction of income and price in consumer demand,” *International Economic Review*, vol. 14, pp. 351–363, Jun 1973.

[26] J. A. Nelson, “Quality variation and quantity aggregation in consumer demand for food,” *American Journal of Agricultural Economics*, vol. 73, pp. 1204–1212, November 1991.

[27] A. Deaton and J. Muellbauer, *Economics and Consumer Behavior*. Cambridge University Press, 1980.

[28] R. H. Strotz, “The empirical implications of a utility tree,” *Econometrica*, vol. 25, pp. 269–280, Apr 1957.

[29] L. Philips, *Applied Consumption Analysis*. North-Holland, 1990.

[30] T. A. Domencic and D. McFadden, *Urban Travel Demand*. North-Holland Publishing Company, 1975.

[31] D. Peng and R. Poudineh, "Sustainable electricity pricing for tanzania," *Oxford Institute of Energy*, 2016.

[32] W. Kaus, "Conspicuous consumption and race: Evidence from south africa," 2010.

[33] K. K. Charles, E. Hurst, and N. Roussanov, "Conspicuous consumption and race," *The Quarterly Journal of Economics*, 2009. http://www.nber.org/papers/w13392.

[34] A. Deaton, *Understanding Consumption*.

[35] G. Williams, *Nigeria: Economy And Society*. 1976.

[36] K. Meagher, "Leaving no one behind?: Informal economies, economic inclusion and islamic extremism in nigeria," *Journal of International Development*, vol. 27, pp. 835–855, 2015.

[37] M. F. Lofchie, *The Political Economy of Tanzania: Decline and Recovery*. University of Pennsylvania Press, 2014.

[38] R. A. Schroeder, *Africa after Apartheid : South Africa, Race, and Nation in Tanzania*. Indiana University Press, 2012.

[39] "Tanzania country brief," *World Bank*, vol. 27, pp. 225–235, 2009.

[40] S. Hanson, "Angola's political and economic development," *Council of Foreign Relations*, 2008.

[41] M. G. Schatzberg, *The Political Economy of Kenya*. Praeger, 1987.

# 6 APPENDIX I

## 6.1 Historical developments in sub-Saharan Africa

The landscape of the political history of sub-Saharan African is dominated by anti-colonial movements in the past century. While it seems natural to hope for Adam Smith's principles to take hold in the nascent economies of the newly formed republics in the sub-Saharan Africa, neither the economic development nor the collective defense of nationalistic interests have followed a path similar to that of the developed world. The industrial class in the African countries is still poor in absolute terms and the problems of extreme poverty have remained largely unresolved in the large swathes of sub-Saharan Africa. The administrative successes and stabilities of post-colonial governments are varied and have depended on the extent of agrarian empires that had existed before. The extractive administrative frameworks of Ottoman or Moghul empires, for example, could be adapted well by European colonists in Asian countries when compared to the administrative units (much as the political boundaries themselves) created in sub-Saharan Africa. The processes of decolonization - which started in the late last century - have encompassed the loosely similar post-war political voices in Asia and Africa aspiring to establish nation-states. While centralization was attempted for decades in both Asia and Africa (often curtailing local-level status competitions and individual freedoms alike) their reach and success has been limited. in the developing world is that the goals aimed for by Adam Smith - of replacing feudalistic rent-seeking with fairness, equity and reward for hard-work - have been attempted more often with the nationalistic or socialist movements than free trade. The role of state has been an uncertain force in the developing world and has thus shaped identities at least as much as the evolution of market for industrial goods have.

**Nigeria**

The access to education is varied in Nigeria - because of the different levels of successes missionary education had in Nigeria. This in turn has created regional disparities in the current education levels between the northern and southern parts of Nigeria. The north has had a higher Islamic influence and the uniformity desired by the post-colonial government had had some initial challenges. At the start of the republic, it seemed that the market forces had been left relatively untouched - due to nature of the British rule which encouraged participation of a native authority. At the same time, the wage labour surviving on the peasantry was a trend that continued well into the post-colonial era. When the movements of African socialism in Nigeria picked up their momentum, the power of merchant class became limited. In the more recent decades, when MNCs could have brought more power to a working middle class, their presence hasn't changed the state of capital being controlled by a small minority - an environment where only the state monopolized industries and the informal sector seem to have expanded[35]. The recent Base of Pyramid (BoP) initiatives have not created a sufficient base for entrepreneurs, but they have indeed revived a focus on education and expanded the market for industrial goods. An average of 42% workforce in Nigeria have secondary education or higher. Upto 28% of those in mere survival activities have a secondary school certificate, and 12% have post-secondary qualifications [36]. Newly urbanized indigenous tribes and educated classes have taken up jobs that had earlier required a much lower level of education. The crowding hardly resolves the underlying problems with the economy - as the formal sector remains in doldrums. The state of economy, rapid population rise and the resulting migration from rural areas has given rise to problems urban overcrowding unmitigated by a large informal sector [36].

**Tanzania**

Given its engagement in the liberation movements in Mozambique in 1974 and the faith in planned economy inculcated by Nyrere, it would not be inaccurate to consider Tanzania the epicentre of the African socialism. But once the political independence was achieved in Tanzania, the membership of nationalist parties declined and the separation of civil service from political institutions became less of a priority. The faith in socialist planning under the leadership of Nyrere meant that the reduction of private sector did not experience much opposition. With a lack of support from its workers and a ban on producer-consumer societies, the socialist model seemed to have had some inconsistencies. In the decades after the independence, the industrial sector continued to suffer because of import subsidies. An import substituting industrialization (ISI) led to oversubsidizing when the need to prevent mass starvation called for an adjustment to capital flows. The rent-seeking in bureaucracy and the use of a capital-intensive technology caused the oversubsidizing to spread further across other sectors. Subsequently, an over-reliance on capital for development resulted in decline of the industrial capacity utilization [37] - many arguing that the workforce had moved to a subsistence through cultivation approach. In its modern economy, only the public officials have had the advantage of becoming entrepreneurs. Even though imports could have improved the Tanzanian exchange rate - the problems around corruption have always posed limitations to trade reforms in Tanzania (particularly in the energy sector)[37]. With state regulated economy having had no ways to expand, the growth of parallel economy has been unavoidable - leaving a disconnect between the parallel markets and a protectionist trade policy. More recently, the firms from South Africa and China have increasingly participated in Tanzania while the training programs for the local labour and the plans to develop local infrastructure languish. The conflicts between the miners and SA migrant labour are common - while many Tanzanians are sent to South Africa for training[38]. Electricity is only available to about 10% of the population. The use of internet communications is higher in Tanzania than in Africa's average but access to finance is low (albeit rising) for the private sector. Quality of life differs significantly between urban and rural regions and the size of the informal sector (60%) is significant[39].

**Angola**

Having achieved independence from Portugal in 1975, the competition between different movements that were vying to lead Angola descended the country into a civil war. The Popular Movement for the Liberation of Angola (MPLA), a Marxist-oriented group that included urban intellectuals, nominally led the country[40]. Similar to other post-colonial economies, the state-controlled companies thrived in Angola. In the view of Western economists, the state oil company Sonagol seemed to have played a quasi-fiscal role. The economy's dependence on oil revenues made economic diversification difficult[40]. Recently however, the industrial goods have proliferated in Anglola with the development of stronger business ties with China.

**Kenya**

The political power is influenced by many ethnicities in Kenya and the clan dynamics is of high importance. Due to this reason, the policy changes can quickly develop into ethnic conflicts. For example, when resettlement was attempted under Kenyatta's leadership, the discontent was predominantly from non-*Kikuyu* population. Other attempts at nationalization - e.g. taking control of foods sales or establishment of purchase centres - have met with similar confrontations. The prevalence of small-scale independent works and lack of support offered to them has not been addressed either by the socialist governments or the growing private sector[41]. Much like other sub-Saharan African countries, there are extreme difference in the urban and rural life-styles in Kenya.

# 7 APPENDIX II

## 7.1 Steps in Preparing Data

Following steps were performed towards the normalisation of LSMS data on Tanzania. The files referred to are for year 2010:

1. Read weekly diary data from Section K (a table of items with the quantities consumed and cost associated with the item for every household).

   (a) All items that had no cost associated with them were ignored (not included in total consumption)

   (b) Gift quantities were ignored for consumption ( median ratio of gift to total diary consumption was zero - only 132/3828 households had this ratio 1% or higher )

   (c) Weekly diary data was multiplied by 52 (to estimate annual consumption)

i. Weekly recall items were also multiplied by 52 (to estimate annual consumption)

(d) Monthly recall items were multiplied by 12 (to estimate annual consumption) - except for repair related cost which we only multiplied by 2 (assuming that repair frequency is ~6 months for all items to be repaired)

(e) All expenditure from (c)-(d) above were summed up as total expenditure

2. Read Assets from Section N (for year: 2010) and calculated asset scores

3. Obtained Personal Data from Section A,B,C and J files

   (a) Section C_CB was read to obtain market facilitycode and gauge the accessibility of a market in every district. The closest accessible market could be either within the district or outside the district at a given distance. If a market was within the district or less than 10 kms away it was deemed "accessible". Urban/rural classifications based on population density could be inserted at this stage (population density in not available in LSMS).

   (b) Read section B and C files

   (c) Calculated age of member by subtracting YOB (year-of-birth) from 2010 (survey year)

   (d) Read section J for housing data (total house rent, number of primary/secondary rooms)

4. Obtained income data from Section E (currently ignored for analysis for it being sparse). Here, the recorded pay frequency was in hours, days, weeks, months, fortnights, months, quarter, half year or year - while the mandatory fields corresponding to all of these units were i) number of hours worked per week ii) number of weeks worked per month and iii) number of months worked in an year .

   (a) When pay was on a per-hour basis, the number of hours worked per week (provided) was multiplied with the number of weeks worked per month (provided). This product was then multiplied with the number of months worked per year (provided) to estimate the annual income.

   (b) When pay was per-day, a 10 hour working day was assumed to obtain the effective number of work-days per week (based on the number of hours worked per week). This was then multiplied with the number of weeks worked per month in the year and then further multiplied with the number of months worked in an year to obtain the estimated annual income.

(c) When pay was per week, the number of weeks worked per month was multiplied with the number of months worked per year.

(d) When pay was in fortnights, then twice the number of months worked in an year was used to calculate the total income received over the year.

(e) When pay was per-month, then the multiplication factor was just the number of months worked per year

(f) When pay was per-quarter, then the effective number of quarters were inferred from the number of months worked per year (number_of_months/3) and multiplied with the number of months worked per year to obtain the estimated annual income.

(g) For self-employed income, the work-months in an year was similarly used to compute total income from self-employment in the year

(h) All members less than 5 year old were ignored from the income data

(i) For wage workers:

    i. summed up wages into column yearly pay
    ii. summed up values under "other forms of payment"
    iii. sum up values as secondary of payment (for wage-workers)
    iv. only primary job was used to identify the employer type of the individual
    v. added other wages from secondary job by summing up yearly-income from all sources into the yearly income

5. Ignored bad data (outliers)

    (a) Ignored 5 households with exceedingly high expenditure on marriage (more than reported annual income)

    (b) Ignored households in the income table but with zero income (number of households with income data thus ignored were under 2%)

    (c) Ignored data with more than 30 times the median cost (ensuring that no more than 3% of the data is ignored)

    (d) Ignore items that are consumed by 10 or less households (items - mortgage, rice_paddy, nut_products, wild_birds, packaged_fish, miscdrinkpowder, readymade_tea_coffee and winespirits were thus ignored in the analyses)

6. Merged all data

(a) Set education expense of houses with education expenses= NA as zero

(b) Summed up educational expense and total house rent from personal data into total expenditure (both weren't a part of diary data)

(c) Obtained personids of the house-heads and the following variables for household-head: education-level, age, years in community, language, occupation

(d) Obtained visible expenditure by summing up expenditure on visible items

(e) Merged all data into one table

While we extrapolate weekly diary to annual expense in Step 1, we must also consider that with a large size of families (40% of households have size 5 or higher), it may be common to stock items for consumption. The LSMS survey records the quantities for food items - even if they were not purchased in the past week. But for other non-food items (such as soap, skin creams whose quantities are not recorded in the survey) are likely to be purchased in bulk in large families as well. Since the frequency of purchases would be lower when the quantity of bulk purchases increases- this may cause us to overestimate consumption. To verify that this not a significant issue, we first test if such stocking (quantities purchased) may be uniform in all region in the country. We then test the significance of household factors that may affect long-term storage (e.g. household size or distance from market) of the purchased item. Observing the quantity purchased on the item $e$, the total expenditure on all items $x$ and the distance from the market $d$, we use a difference-in-means analysis to confirm the low effect of travel-costs on the ratio $log(\frac{e}{x})$ and conclude that stockpiling is not significant. It turns out that the number of family members is a far more significant factor for large purchases.

For non-food categories, where the data is less detailed, some further adjustments were made for the classification of non-durable consumption into the wide categories of energy, transport and household consumption. For example, repairs and maintenance costs - which are provided in the weekly diary - were not simply extrapolated to annual consumption (multiplication of a factor of 12) - but instead an assumption on the average life of the item being repaired or maintained was considered to perform the extrapolation to the annual expenditure.

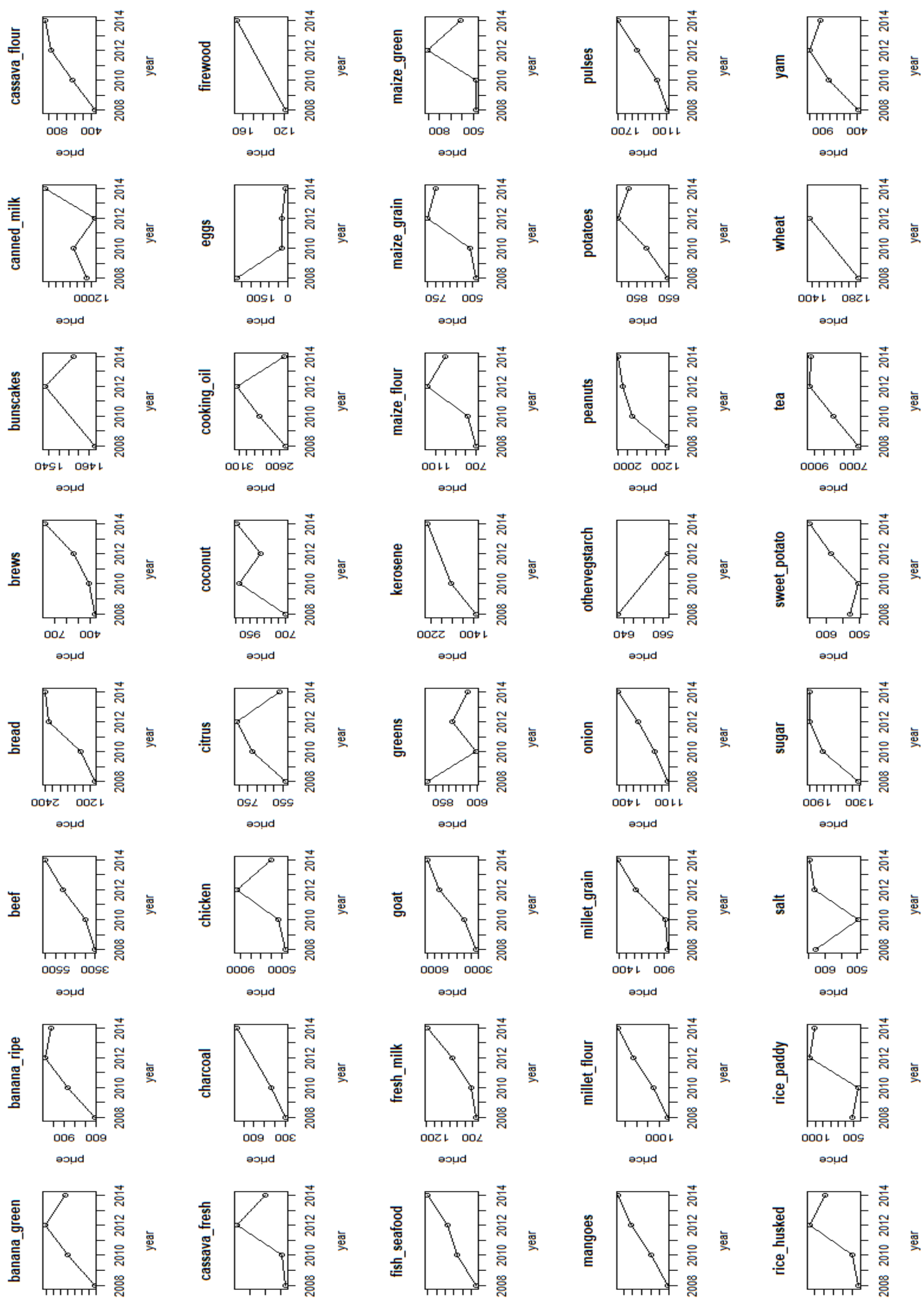


Figure 5: Market Prices recorded in 2008, 2010, 2012 and 2014 surveys