The Role of Income Inequalities in Aspirational Consumption

Summary

The presented study explores reported rises in aspirational consumption in the developing economies and asks whether such a rise may be subject to a leveling effect seen in developed economies or not. Viewing status consumption in a generic context, the study assumes an explicit link between status goals and economic goals of the consumer. The empirical analyses then assess possible rises in consumption by examining the role of household permanent income on food and education - both of which are observed to carry a status value in sub-Saharan Africa. The three chapters in the study draw a set of independent conclusions on effects of inequality on status consumption. The equilibrium conditions discussed in the first theoretical chapter for a model of status as a position hierarchy where consumers are able to move between through social capital investments reveal that positional contests can become uncompetitive with a rise in income inequality. This first chapter highlights two types of long-term equilibria where either only rich consumers participate in status competitions or both rich and poor participate. The empirical second chapter explores the role of household permanent income on food quality in Tanzania to observe that permanent income - alongside with availability of electricity in the country - may have a segregating effect on food quality. The third chapter compares education expenses in Nigeria and Tanzania - finding that education expenses in Nigeria are more significantly influenced by local wealth levels than permanent income. In light of the two long-term equilibria discussed in the first chapter, the study highlights how a rise in expenditure on status items as a common need may contribute to rise in status consumption in developing economies when the consumption of such items is linked with perceptions of mobility.

Contents

Part I

Preface

At a time when developing economies strive to improve per-capita life-standards, studies have also highlighted an increasing role of aspirational consumption (Burke, 1996; Srivastava, Mukherjee, & Jebarajakirthy, 2020; Moav and & Neeman, 2012). Aspirational consumption is commonly understood as consumption relevant for status signaling needs or social aspirations - but not only are the empirical tests of visible consumption across developing economies sensitive to extreme variation in cultural contexts, the very idea that visible consumption is wasteful may also deserve scrutiny if a rising affluence of societies transitioning away from subsistence lifestyles (or endogenous growth) raises common needs across the society.

Such a rise in common needs could be both due to a rise in quality of basic consumption (such as food, clothing) and with an introduction of newer goods into the consumer basket. It is not surprising thus that on one hand, the literature on aspirational consumption finds aspirational consumption to form a pathway to a brighter future (Srivastava et al., 2020; Jaikumar, Singh, & Sarin, 2018; Duflo & Banerjee, 2011) but on the other, it is also suggested that consumers stretch their budgets on basic needs to address compensatory aspirational consumption (Kaus, 2013; Khamis, Prakash, & Siddique, 2012; Roychowdhury, 2016). Addressing the specific concern of what may constitute status consumption in a developing economy, the presented study focuses on two shortcomings of the prevalent view of rise in aspirational consumption in order to answer how inequalities influence the rise in aspirational consumption. First, it revisits the sociological basis for status in order to consider a broader view of status-related consumption than what is permitted with the notion of visible consumption. Second, it considers mobility perceptions as a potential factor for consumption changes in an environment of wide income differences. It is by addressing both of these issues that the thesis attempts to answer how inequalities may contribute to a rise in aspirational consumption.

The two issues pertaining to aspirational consumption amidst inequalities are particularly relevant to SSA - first since a recent urbanisation of the region may allow for status indication though higher quality or newer consumption types and second, since recent economic growth may prioritise mobility and its perceptions - resulting in the seemingly "futile" (Veblen, 1899) consumption becoming a social need with material consequences for the consumer. Drawing from the sociological literature for status (Wegener, 1992), the presented study links consumption with notions of aspiration to provide theoretical arguments for how aspirational consumption could change under income inequalities in developing economies. The empirical chapters in the study then examine distribution of consumption pertaining to status with respect to permanent income in SSA.

We argue that the idea of certain goods carrying a visible appeal has often been embraced in the developing economies without examining the historical and cultural settings that might influence status perceptions in an economy (Nieftagodien, Van Der Berg, et al., 2007; Burger, Louw, Pegado, & van der Berg, 2015). The interpretations of status signaling mechanisms in the developing world also ignore that widespread poverty tends to make seemingly necessary items carry status significance as well. The expenditures such as education - for example - have been treated as a necessity despite the literature finding some evidence of rent-seeking (Buchert, 1992; Babalola, 1998; Filmer & Pritchett, 1999; Barr, 2004; Ndyali, 2016; Allais, 2017; Ebaidalla, 2018) or eliteness (Minnis, 2006) from educational degrees. Conversely, while consumption of clothing or personal products - regardless of variation in quality - are commonly considered visible consumption (Kaus, 2013; Khamis et al., 2012; Roychowdhury, 2016) in the developing economies, the literature also points to the necessity of such consumption by finding evidence of improvement in subjective well-being (Linssen, Van Kempen, & Kraaykamp, 2011; Guillen-Royo, 2011). A more general view of aspirational consumption adopted in the study attempts to highlight how some of the claims of unusual or prominent urges towards aspirational consumption among the poor developing economies (Subrahmanyan & Tomas Gomez-Arias, 2008; Srivastava et al., 2020) - may in part be due to the lack of an established standard for maximal level of common needs in an economy. With there being well-known issues with a direct measurement of necessities of consumers (Sen, 1985; Townsend, 1985; Doyal & Gough, 1991), the attribution of higher consumption to aspirational needs might be appropriate in a normative sense, but frequent attributions of higher or lower wastefulness to specific sections of society require one to reconcile the sociological basis for status needs in an economy as well - thus considering both the dynamics of status in the economy and the notion of wastefulness associated with certain consumption.

In a theoretical exploration of how income differences may influence aspirational consumption (in a more general sense than what is implied by visible consumption), we focus firstly on a particular social value that the consumer may derive from aspirational consumption. Drawing from the literature on status in sociology (Weber, 1925; Goode, 1978; Wegener, 1992), the theoretical first chapter assumes a specific link between economic and social goals of the individual to explain effect of income-inequalities on consumption - which are viewed as investments towards social capital needed to pursue social positions. The empirical chapters then proceed to examine the distribution of exclusive and high-quality consumption that have been reported to carry status advantage in SSA but do not fall under visible goods category.

It is worth elaborating how the above view of aspirational consumption contrasts against that of conspicuous consumption as futile splurges of the rich (Veblen, 1899) and how it motivates our exploration of income inequality effects. The distinction between the view of conspicuous consumption as splurges of the rich and that of a generalised status

consumption is an important one since if high-priced consumption were to be limited only to those with higher income or wealth, then there are evident limits implied for the rise of such consumption (Hirsch, 1977). In other words, the high-priced consumption - if limited to richer sections - is subject to how extreme income inequalities persist in the economy - since such consumption may rise only as much as the proportion of rich rises in a population. Therefore if aspirational consumption is strictly interpreted as trophies of the rich or as futile attempts to emulate the rich, then such consumption is necessarily limited to the rich and its purported rise cannot be widespread in an unequal society. With a more general view of aspirational consumption taken in the study on the other hand, the scope of aspirational consumption is necessarily wider as it concerns itself with how aspirational consumption may expand to the masses as a common need. The general view of aspirational consumption combined with the link between economic and social goals allows us to interpret the effects of mobility concerns on consumption as investments towards social capital towards economic gains (Goode, 1978; Hirsch, 1977) for the consumer. As we further explain, such a general view may be better suited to reconciling the sociological bases for status in an economy.

The empirical chapters in the study do not assume the existence of purely visible items and consider two ways in which high-consumption of status relevance can be availed in the cross-section of a typical sub-Saharan economy. The first type of consumption is one that can be viewed as a higher quality of a commonly consumed basic good. The second type is one that is altogether exclusive to the richer sections i.e. inaccessible to the vast majority. The primary distinction between the two-types of consumption in the study is an empirical one wherein the quality variation is relevant for the first type of consumption and it isn't for the second type. The rationale behind this distinction is that in a generic sense of status signaling (rather the one implied by futile goods alone), the available quality in the market makes it possible for a richer consumer to distinguish herself from the poorer consumer while an item with less lower-quality variants may be relevant for status purposes only through the poor being priced of purchasing such an item. One can hardly dispute that food is a basic necessity but there also seems evidence for status-indication through food quality in developing economies (e.g., for Tanzania(I. L. Ohna, 2007; I. Ohna, Kaarhus, & Kinabo, 2012)). Mass Education, on the other hand, is quite unlike food and is a relatively newer consumption type in SSA - with an evidence for its status value as well (Minnis, 2006; Porzio, Rossi, & Santangelo, 2021).

The first empirical chapter in the study inspects variation in food quality in Tanzania - which presents itself as an economy where rapid changes in consumption patterns have been highlighted under recent urban developments. The second empirical chapter compares education expenses in a rural economy with a relatively more urbanised one by selecting Tanzania and Nigeria - which are countries with disparate levels of industrial dependence. It is worth highlighting that all consumption - regardless of high or low quality - is surveyed across the cross-section in both the empirical chapters. Since consumption being limited a

rich household must evidently limit the long-term rise in overall consumption in the economy, the role of household permanent-income is explored in both the empirical chapters so as to comment on possible future rises in consumption.

As such, many different scenarios may arise with food and education considered as the two types of aspirational consumption. With food quality, one could find that budget shares for food quality do not vary much across the population (as is the case for only a few specific food items) or that high food quality is severely limited to richer households. Similarly, with respect to education - which is more exclusive relative to food - one may find that it is affordable only to rich households or that it's availed across all income levels. A comparison of education expenses in Tanzania and Nigeria suggests that education is distributed across the population in asset-rich regions whereas a view of distribution of food quality in Tanzania shows that high quality is significantly limited to richer and urban households. The role of household permanent income appears to be less important for education in urbanised areas than for high quality food - suggesting how certain types of consumption - subject to changes in economic growth - are more likely to become a growing need in the population. These observations can be viewed in light of the theoretical findings of the first chapter on how persistent household income inequalities influence demand for status-goods varying in their potential mobility value. As we detail in the study and discuss further in the epilogue (Part V), education may be different from food not only through being a newer type of consumption category in the rural economies (which also allows the variation in quality to be of a less empirical concern) but also in how it may relate to future economic gains. A view of expenditures sensitive to perceptions of mobility (e.g., education) that considers status neither a purely conspicuous concern nor a necessity may be better equipped to explaining upward pressures on certain aspects of aspirational consumption.

To sum up, a theoretical chapter in the presented study finds that long-term socio-economic income inequalities limit competitions for status even though such investments (i.e. aspirational consumption) rise with an increase in inequality in the short-term. The empirical assessments then find that while high expenditure for certain items may be limited to those with higher permanent income, the urbanisation-related changes could have an effect of reducing the role of household permanent income on expenditure for certain consumption categories (while also creating wider differences between urbanised and rural areas). That richer consumers (those with higher wealth and/or income) are likely to have better access to higher quality goods and items that are out of reach for the masses is less surprising, but whether such consumption is likely to be limited to the richer sections of the society in an environment of wide differences in wealth and income levels depends much on the type of value the particular consumption may provide to the consumer rather than on notions of wastefulness - especially when the notions are not validated as a part of the status dynamics of an economy. The presented study urges more focus on mobility perceptions as such a value - which ought to be taken into account when

examining changes in consumption in developing economies.

Part II

Status concerns and consumption under Income Inequality

Abstract

Status motivations for consumption have generated an immense interest at a time when social mobility has been relatively stagnant - but the consequences of inequality within societies on patterns of development are still poorly understood. To explain status motivations under environments of income inequality, we treat opportunities for mobility as a link between economic and social goals of the individual and view the consumption as investments towards future economic success. Considering an economy with a dynamic hierarchy of rich and poor positions that consumers are able to move between, we then ask how relative incomes associated with each position may influence the agents' consumption or labour supply decisions. With a diminishing returns to utility from wealth optimised by a consumer endowed with randomly distributed success-determining attributes and the promotion probabilities depending on her expenditure, we find that a rise in income inequality provides more incentive to participate in status competitions. With consumers bearing the participation costs for mobility, the long-term equilibrium conditions suggest two states - one where both rich and poor strive for status and the other where only the rich consumers participate in status competitions.

1 Introduction

The consumer response to relative income has appealed to economists since at least as far back as Veblen (1899). While the post-war period of relative stability and life-style improvements have renewed an interest in the positional concerns (Frank, 1993; Hirsch, 1977), there remains a disconnect between descriptive theories for relative income effects and theories of economic growth (see (Kay & King, 2020; Dhami, 2016)). In the context of aspirational consumption in the developing economies, such a disconnect makes it difficult to argue whether rises in aspirational consumption are a mere consequence of social shifts pertaining to growth and income redistribution in the economy or not. As the economies

may vary significantly across institutional efficacy and socio-political environments, a prediction of rise in aspirational consumption is problematic unless one relies upon a well-defined mechanism relating relative income concerns and growth in the economy. We argue that a view of status goals as independent of economic concerns may not offer a satisfactory view of aspirational concerns in the developing economies. Our appraoch uses a specific link between economic and social goals instead to answer the questions pertaining to effects on income inequality on aspirational consumption. The presented chapter focuses on a context of aspirational consumption where it can be treated as investments towards mobility and then asks whether such investments should increase or decrease with a rise in income inequality.

Such a view of status adds a particular dynamism to the predominant view of status as a hierarchy of ranks (Goldthorpe & Hope, 1974; Wegener, 1992; Goldthorpe, 2000). In viewing status competitions as games that provide direct gains to the consumer as probabilistic economic rewards pertaining to status, the current chapter takes cue from several attempts in both sociology and economics that have argued for a more dynamic framework for status (Weber, 1925; Shils, 1969; Parsons, 1971; Coleman, 1990; Wegener, 1992). The presented approach pins down a specific link between economic and social goals relevant for status consumption by assuming an indirect (instrumental) effect of status on consumer wealth (Cole, Mailath, & Postlewaite, 1992; Postlewaite, 1998). While such a link may not have been of relevance to many empirical studies on visible consumption (Heffetz, 2011; Ireland, 1994; Omori & Smith, 2015; Charles, Hurst, & Roussanov, 2009; Kaus, 2013; Khamis et al., 2012; Friehe & Mechtel, 2014), it receives much attention in the sociological literature (Wegener, 1992) where several approaches have addressed the equivalent problem of a reconciliation of subjective aspects of status (the effect of status differences on individual behaviour) with its objective aspects (the observed reality of status differences). In particular, there seem two broad sets of theories for status mechanisms (Wegener, 1992) - considering either a hierarchical order (Parsons, 1971; Wegener, 1992) or a conflict among social groups (Weber, 1925; Wegener, 1992) that attempt to reconcile the two aspects of status. Relying on a particular dynamism in the social order argued for by Goode (1978), the current chapter views status as a product of economic processes shaped by institutional nurturing or discouragement of individual attributes (Goode, 1978; Moldovanu, Sela, & Shi, 2007).

Given a significant variation in economic development across developing economies (such as Tanzania and Nigeria that we discuss further in Chapter 2), the endogenisation of status in such a manner provides a generic framework to compare aspirational consumption across disparate developing economies that may have differing levels of urban development. With regards to the specific problem of how variation in urban institutions could influence expenditures such as education (that carry some status but also contribute to future income mobility), the proposed theoretical chapter considers how expenditure towards status could

respond to changes in income differences and participation costs under the urbanisation-related changes. Unlike the typical attribution of status to exogenous and idiosyncratic factors such as race or caste in empirical studies, the presented approach views status as expectation of future incomes and treats higher urbanisation as an environment where purchasable social capital has a significant role to play towards towards individual success (Weber, 1925; Goode, 1978; Wegener, 1992).

We believe that such an explicit link between economic and social goals could also be key to addressing some of the debates in economics on whether inequality has a negative or positive effect on status consumption. More particularly, while the post-war literature on status competitions attributes a higher status consumption to a fall in income inequalities and improvement in life-standards (Hirsch, 1977; Galbraith, 1958, 1994, 2001), the empirical literature finds more often that status consumption increases with rising inequality. The disagreement seems at least partly due to the sociological observations of higher status needs being normative rather than descriptive concerns. More specifically, the observations of the post-war era - pertaining largely to the objective notions of status - do not suggest that consumers must feel a higher need for status as the inequality rises (or vice versa) - only that the avenues for indicating status have declined as the life-standards become uniform with decrease in income inequality (Harrod, 1958; Hirsch, 1977). When equipped with subjective welfare data, on the other hand, the recent empirical analyses - concerned with the subjection notions of status - find a positive effect of income inequality on status consumption instead. Since status competitions are tied with particular social contexts, an analysis of status consumption in a growing economy must deal both with the environment where inequality is experienced along with the growth in economy. A specific link between objective and subjective notions of status - where the former represents economic goals and the latter social goals for the individual - may thus allow us to better understand how inequality shapes consumption decisions in an economy.

To delineate the structural relationship between the economic and social goals of an individual, the current chapter uses probability functions for mobility that determine promotion to a richer class. While the probability of promotion for the consumer is driven by her expenditure and certain attributes of the individual, the utility for the consumer comes only from her accumulated wealth. Our approach where participants do not consider the rich status directly into the utility (and therefore do not directly maximise the chances to be rich) - follows Postlewaite (1998) and Cole et al. (1992) closely who also discuss a relevance of uncertainty in conspicuous consumption (also see Truyts (2010)). In the probabilistic game we use to represent status competitions, the participants compete both with raw talent and acquired skills (social capital) to get promoted to (or to maintain) a rich position. The winner of such a game is promoted to a rich position (or maintains her rich position) by having the maximum total score comprising both of raw-talent r_i (uniformly distributed in a population) and social capital s_i (enhanced with expenditure). The social-capital s_i that is enhanced

through expenditure is also subject to budget constraints imposed on the consumer[1].

It is worth highlighting that the only two types of participants in the tournament view of status used in the current chapter are the poor and rich classes - so that the investments towards status (such as club-membership fees, or higher-education) increase the chances to be promoted to the rich status. The rich consumers in the game or those who benefit with a higher disposable income (y_2) solely on grounds of their rich status while the poor consumers are those who receive a disposable income $y_1 < y_2$. Those who don't participate in the tournament certainly suffer with being poor, but those who do spend only benefit probabilistically. A useful analogy to explain such a competition is that of tennis leagues or golf tournaments - where rewards are risky but are unlike lotteries (in that - more expenditure does not linearly increase one's chance to win in a tennis championship). As with the game of tennis, participants with appropriate raw-talent and skills may exist naturally in a population but their success also depends on engagement with other tennis players. More specifically, there may exist biological characteristics such as strength, dexterity, height or psychological abilities (resilience, conscientiousness etc.) that help one's tennis, but there also exist several characteristics such as technique, rankings, knowledge of strings-compatibility or racket-customisations etc. that are only available at a significant cost to the participant. The enhancement of these latter skills (social capital) are what the expenditure-depending institutional support are meant to represent. Several reali-life status competitions are similar to tennis-like tournaments where interaction is small-scale as well as personal and social capital investments have a strong role to play - e.g., in interviews, formal introductions or business interactions. It has been argued more generally that the characteristics of status games are often made to reward curated abilities rather than naturally occurring raw-strengths (Goode, 1978). Given the specific context of new environments of higher economic growth and introduction of education as an engine of mobility in the developing economies, we focus on consumer behaviour where the participants in a status competition expend without an *a priori* knowledge of their individual raw-talent or social capital.

It is important to highlight that above status competitions are not meant to be population-wide tournaments that involve all participants of a society. Much like a tennis match, the number of participants is meant to be small in such competitions. The small-scale competitions vary locally defined criteria that may have multiple stages. Simply put, such competitions are more like an in-person job-interview than a nationwide examination for college-admissions. All such competitions are meant to be independent and may vary only in the threshold participation cost and range of distribution of r_i, s_i of the participants. As scope of the presented chapter is limited to the effect of income differences on consumption,

[1]We do not view leisure as separate from budget constraints in the model. In other words, the leisure relevant for higher stauts is interpreted simply as a higher cost for the poorer participant. So long as once can view leisure as the amount of money a participant must forgo in order to participate in status competition, the leisure differences can be thus also viewed as differences in disposable incomes.

we discuss only a single instance of a competition with known distribution of r_i and s_i among a relatively small number of participants who exercise their choice to spend on social capital without an *a priori* knowledge of r_i and s_i.

Notably, a few properties of the game follow where a wealth-optimising participant with a maximum score wins by enhancing her skills (social capital) and using her raw-talent. First, only one player with the maximum score can win a particular tournament. Second, the rank of the participant (rich / poor status) is of no direct significance part of her utility from wealth. Lastly, the raw-talent r_i and skills s_i are of no significance to expenditure decisions - which vary solely on the basis of rich or poor status. A closed-form solution with a diminishing utility from wealth and a threshold level expenditure required for participants to spend in order to avail social capital suggests that the only two equilibria relevant for status consumption are one where rich as well as poor consumer participates in the status competitions and one where only the rich consumer engage in competitions. Further, the higher income differences seem to encourage both poor and rich participants to engage in status competitions while high participation costs discourage the participation from the poor participant.

We believe that the results from the model contribute to the literature in a few different ways. First, the framework with uncertainty in consumer choice proposed in the chapter may be better suited for incorporating choices that affect mobility (including social hierarchy) that in turn correspond to material gain from status competitions. Second, the model is more useful for understanding status-related investments such education or professional club memberships that contribute probabilistically to consumer wealth. Third, it elucidates how higher income differences and participation costs in developing economies could contribute to a rise or decline in investments such as education. Within the framework used in the current chapter, it can be argued that education expenditures are subject to a discouragement through high participation costs but an incentivisation through high income differences.

In what follows, Section 2 surveys the issues with definition of status in sociology and economics, Section 3 discusses the model and its solution while Section 4 draws some conclusions about expenditure towards status in the growing economies.

2 Literature Survey

While status is often contrasted as a social goal rather than the economic goal of an individual, the difference between social and economic goals remains unclear in both sociology and economics (Wegener, 1992). The idea that status is orthogonal to economic goals - or the "pure status" view - seldom fits with the real-world where status both contributes to income and improves with rises in income (Moldovanu et al., 2007). Indeed the other extreme view that the economic goals alone motivate status competitions - so that status is derived purely from economic rewards is also unrealistic. As Moldovanu et al. (2007) argue, the reality is often somewhere in between the two extremes. Truyts (2010) go

further in saying that status would in fact of little importance if there were no interdependence between status and consumption or labour decisions.

A structural model for the interplay between economic and social goals is however yet to specified in either sociology or economics (Truyts, 2010). The dominant view for status in both sociology and economics has been based on hierarchy and social deference (Coleman, 1990; Goldthorpe, 2000) - a view that is inherent in occupational status surveys widely used in empirical studies (e.g., the EGP classes or the Goldthorpe class scheme). A fundamental problem that the sociology literature has elaborated further (Wegener, 1992) is the failure in empirical studies to reconcile the difference between objective aspects of status (the observed reality of status differences) and its subjective aspects (the effect of status differences on individual behaviour). An evident issue with addressing only the objective aspect of status using a utilitarian model while leaving the subjective aspect exogenous is that status may never be explained in economic terms when an economy for status is ruled out. As the literature in sociology warns, the idea of social-closure (aggregate or *Stände* which were first discussed by (Weber, 1925)) ought not be conflated with that of social ranks (Wegener, 1992). Besides, the issue has not been merely a theoretical one. Empirical studies suggest that the subjective (private) evaluations of status often measure an overall desirability of occupations rather than common objective notions for status (Goldthorpe & Hope, 1974). Adler and Kraus (1985) find that skills or knowledge of individuals to be a better predictor for individual judgments of status. Lastly, the scales of status used in empirical studies are observed to vary significantly across the sections of a society as well (Acker, 1980).

A link between objective and subjective aspects of status may also be key to addressing the debates in economics on whether inequality has a negative or positive effect on status consumption. On one hand, the post-war literature on status competitions argues that a fall in income inequalities and improvement in life-standards has resulted in a higher focus on status goods (Hirsch, 1977; Galbraith, 1958, 1994, 2001). On the other hand, the empirical literature finds more often that status consumption increases with rising inequality. In a study focusing on education expenses for status purposes in China, Jin, Li, and Wu (2011) find, for example, that higher inequality increases the consumption towards status. Similarly, Jaikumar and Sarin (2015) also report that consumers in India have increased conspicuous consumption with rise in inequality. An appropriate line of research has been to examine the environment where inequality is experienced (often relying on social attitudes surveys). Mijs (2019), for example, use the data from International Social Survey (Haller, Jowell, & Smith, 2009) to detail how a rise in inequality might have cemented a belief in meritocracy in the developed economies and coincided with segregation rather than a healthy competition. Since the environment where inequality is experienced comprises both of the objective and subjective notions of status, a link between the two could help us elucidate the effects of inequality on status consumption.

Several approaches have aimed to reconcile the difference between objective and

subjective aspects of status (Wegener, 1992). Parsons (1971), for example, addresses the issue by simply ensuring that the subjective and objective aspects coincide in status-related interpretations. Parsons (1971) argues that the basic categories in terms of which we describe a system of hierarchies as a structure is the same as those in terms of which individual behavior or performance should be defined. While culture places a premium on equality, Parsons (1971) argues, the distribution rules in different social subsystems are not aimed at generating equality - thus letting status exist to assert basic importance of equality of membership status while also making allowances for inequalities resulting from equality of opportunity (Wegener, 1992). The bottom line of these arguments is that societies must institutionalize some balance between equality and inequality at all times (Parsons, 1963, 1971). On the other hand, Goode (1978) discusses an economy of status where status is created as part of the economic processes while social institutions nurture or discourage certain naturally characteristics in a society. Instead of interpreting status items as those that signal hierarchical status, the dynamic view of status allows Goode (1978) to suggest that investments towards status may increase the likelihood of associating with higher-status members of society.

Other attempts to reconcile the subjective and objective aspects of status (and therefore of social-closure and hierarchy) use rational conflict-based arguments to specify the structural basis for status. Coleman (1990), for example, considers power (purportedly amounting to status) as a commodity to be exchanged in the economy - an approach that has drawn much criticism due to the zero-sum characteristics it imposes on status while treating the total status on a status society as fixed (Wrong, 1979). The current chapter instead leans towards the view from Goode (1978) who had argued that as a product of exchange processes, status is likely to be available in ample supply in a society. While the rich and poor positions are limited in an objective sense in the current chapter, the sense of status as a motivator for consumption decisions applies to all consumers. An explicit mechanism for mobility links the subjective and objective notions in the chapter.

To specify the processes that link the economic and social goals of the individual, the current chapter uses probability functions leading to hierarchical promotion that depend on an *a priori* distribution of personal attributes. Instead of using status as a wealth-determined goal, the presented chapter follows an approach suggested by Postlewaite (1998) to consider status only of an indirect role in maximisation of consumer utility from wealth. While the literature has used both preference interdepdence and constraint interdependence approaches for modeling of status (Truyts, 2010), Postlewaite (1998) argue the two to be equivalent - since a utility function with status directly as an argument could also be seen as a reduced form of instrumental treatment of relative position. The indirect role of status in utility maximising is modeled as a matching-process by Cole et al. (1992) - where the participants have a concern for relative standing because of their standing being instrumental in determining ultimate consumption levels rather than being directly contributing to their

utility. In the approach taken in the current chapter, the rank of the consumer (rich or poor status) does not directly contribute to consumer utility. Instead, the role of rank (rich or poor status) is realised only through expected wealth utility which the consumers maximise.

The above probability function approach also bears some similarity with the Tullock function used in contest theory approaches (Tullock, 1967). Further developments in the contest-theory literature relating to multiple stage contests (Parco, Rapoport, & Amaldoss, 2005; Stein & Rapoport, 2005) and budget-constraints (Jorgenson, Slesnick, & Stoker, 1988; Che & Gale, 1998, 1997, 1996; Parco et al., 2005; Stein & Rapoport, 2005; Brusco & Lopomo, 2008) have also explored contexts similar to the model proposed in the current chapter. However, the approach used in the current chapter remains much simpler than that of a multi-stage Tullock contest under budget-constraints. Unlike the setting for two-stage contests explored in the contest theory literature - where the first-stage is used to limit the number of players participating in the second-stage (Baye, Kovenock, & De Vries, 1993; Amegashie, 1999) - the game presented in the current chapter only uses second-stage to represent the last time-horizon (i.e. end of lifetime) of the participant rather than as a stage to eliminate players. The distribution of budget across the two-stages of the game is also of little concern in the setting explored in the current chapter since it's only the first-stage of the game when the expenditure can be less than the total budget. More particularly, the expenditure in the second-stage simply amounts to the disposable income y_1 for the poor participant and y_2 for the rich participant. It's only in the first-stage, strictly speaking, therefore that the consumption decisions are made by either participant.

In summary, the view of status as an economic goal that corresponds to a possible future monetary benefit to the consumer - rather than a non-monetary utility - is unlike the view of status signaling used in many studies on conspicuous consumption (Heffetz, 2011; Ireland, 1994; Omori & Smith, 2015; Charles et al., 2009; Kaus, 2013; Khamis et al., 2012; Friehe & Mechtel, 2014; Jaikumar & Sarin, 2015; Knell, 1999; Falk & Knell, 2004). This is primarily because the utility from status consumption in the empirical studies on conspicuous consumption is treated both necessarily futile and entirely separable from a private (non-visible) utility. Our specific consideration of status - an approach that equates the social goals of the consumer with her long-term economic goals - bears more similarity with the approach followed by Cole et al. (1992) instead. Unlike the utility from consumption in excess of a certain reference level of consumer needs, the current approach views status consumption as investments towards material gains that are selected to maximise the utility assets accumulated under a scheme of uncertain promotions.

3 Model

In the status game that selects the winner to be maintained at the rich position, only two characteristics matter for every individual. The first is the social capital that can be enhanced

with expenditure by the individual participant and the second is a raw-talent that cannot be changed with higher expenditure. Both skill s_i (enhanced with expenditure) and the raw talent r_i (which cannot be changed with higher expenditure) are assumed to be uniformly distributed in the population. Assuming the weight μ of raw-talent in the maximum total score considered for the win in a competition, we consider the consumer preferences for the probability of $\mu r_i + (1-\mu)s_i$ (where r_i is raw-talent and s_i are skills that are available with higher expenditure) under the assumption that the population remains constant over time (there are the same number of rich and poor in total) and that the participants are not aware of their r_i, s_i before participating in the competitions. In other words, the rich participants are replaced by either a rich or poor participant only through the tournament that selects highest combined score regardless of rich or poor background of the participant. Since all consumers can participate in every such tournament, it is possible for a rich participant to lose her rich-position and let another participant acquire her position.

We also assume that the preferences for all poor participants are the same as are the preferences of all the rich participants. Having a rich or a poor status by itself only has an indirect effect (through higher expenditure or budget) on the chance to be promoted (or maintaining one's rich position) - a probability that otherwise depends on the individual ability r_i and social-capital s_i. All rich participants and all poor participants having the same behaviour also means that the individual attributes r_i and s_i - which are not known to the participant - are not used to decide upon the expenditure (effort) towards the status game.

With r_i and s_i uniformly distributed, we can formulate the likelihood of any participant $i \in [1, N]$ as the probability for their total score $\mu r_i + (1-\mu)s_i$ being higher than any other participant - where $\mu \in (0, 1)$ is a parameter that combines social-capital s and raw-talent r so that the participant with maximum score $\mu r + (1-\mu)s$ wins the competition. The winner of the competition acquires a rich position so that she obtains the disposable income y_2 while the loser acquires the poor position receive $y_1 < y_2$.

To start with a simpler case, consider that in a competition with raw-talents alone, the winning participant is one who has the maximum r_i drawn from the sample N. In other words, the winner in a competition of raw-talents is a player j such that $\{r_i < r_j \forall i \in [1, N]\}$ and the probability that a given participant i of raw-skill r_i is the winner is essentially that of all other players $j \neq i$ having ability $r_{j \neq i}$ less than r_i. This is the probability

$$\int_0^1 P(r_i > r_1) \times P(r_i > r_2) \times ...P(r_i > r_{i-1}) \times P(r_i > r_{i+1})...P(r > r_N)dr_i$$

With a uniform distribution of skills $r \in [0, 1]$, we have $P(r_i \geq r) = r_i$ for all $i \in [1, N]$ and therefore, the probability is evaluates to $\int_0^1 (r_i)^{N-1} dr_i = \frac{1}{N}$. With more players, the chances of a particular participant winning evidently declines.

Introducing the skills s_i which are curated in social institutions, we consider a threshold

investment $\overline{\nu}$ which is necessary to be spent for the social capital to be available to a participant. While this is an arrangement that evidently favours the rich participant, there are also diminishing returns to investing $\overline{\nu}$ towards the richer position if naturally occurring raw talent r_i is also significant in the competition. To direct our focus on conditions when social capital is more important than raw-talent, we assume in the foregoing analysis that raw-talent is less valued than skills cultivated by institutions. As a result, the condition $1-\mu > \mu$ (which is equivalent to the following assumption) has been used throughout in the rest of the discussion.

Assumption 1. $\mu < \frac{1}{2}$

With the threshold investment $\overline{\nu}$ as a discrete response to costs of participation, the participants face different probabilities to win depending on whether they spend $\overline{\nu}$ or not. More specifically, the participant takes advantage of the skills with weight $1-\mu$ only when she spends an amount $\overline{\nu}$. In the all-or-nothing payment scheme, the participant might avoid the investment altogether and keep the investment towards savings instead. The behaviour of all participants depends only on ν_1, ν_2 - as they are unaware of r_i, s_i before they participate.

The result we use depends on the following distribution function of $x_i \equiv \mu r_i + (1-\mu)s_i$ under the Assumption 1 where r_i and s_i are independent random variables uniformly distributed between 0 and 1 (see Appendix 5.1) .

$$F_X(x) = P(x_k < x) = P(r_k\mu + s_k(1-\mu) < x) = \begin{cases} \frac{x^2}{2\mu(1-\mu)} & 0 < x < \mu \\ \frac{x-\mu/2}{1-\mu} & \mu < x < 1-\mu \\ 1 - \frac{(x-1)^2}{2\mu(1-\mu)} & 1-\mu < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

In the game where a participant i with the maximum score x_i is the winner, the number of rich positions is fixed to $N-P$ so that only the top-scorer gets to occupy the only rich position available to be filled in every competition. The expected behaviour of the rich and the poor participants in the long-term can be explained with four different cases.

- Case I - The poor participant does not participate due to insufficient investment $\nu_1 = 0$ while the rich participant does participate by spending $\nu_2 = \overline{\nu}$.
- Case II - The rich participant does not participate with $\nu_2 = 0$ but the poor participant does participate i.e. $\nu_1 = \overline{\nu}$.
- Case III - Both poor and rich participants do not participate so that $\nu_1 = \nu_2 = 0$.
- Case IV - Both poor rich and participant participate with $\nu_1 = \nu_2 = \overline{\nu}$.

With Case I, the poor participant can win when her raw-talent (weighted with μ) is higher than the total score of all other poor participants as well as the total score of the rich participants (who compete with acquired skills). The poor are priced out of accessing any skills s in their score in Case I so that only the rich participants with investment $\nu_2 = \overline{\nu}$ have skills $s > 0$. Thus, a given poor participant must be ahead both of all of the poor participants and the rich participants so that a participant i may win only if μr_i is higher than than raw-talent μr_k in all poor participants and also exceeds the score $\mu r_k + (1-\mu)s_k$ involving skills among the rich participants

$$P_1(\text{win}|r_i) = \prod_{k=1,k\neq i}^{P} P(\mu r_k < \mu r_i) \prod_{k=P+1}^{N} P(\mu r_k + (1-\mu)s_k < \mu r_i) = r_i^{P-1} F_X(\mu r_i)^{N-P} \tag{1}$$

Therefore,

$$P_1(\text{win}|r_i) = \begin{cases} \frac{\mu^{N-P} r_i^{2N-P-1}}{2\mu(1-\mu)} & 0 < r_i < \mu \\ r_i^{P-1}(\frac{\mu r_i - \mu/2}{1-\mu})^{N-P} & \mu < r_i < 1-\mu \\ r_i^{P-1}(1 - \frac{(\mu r_i - 1)^2}{2\mu(1-\mu)})^{N-P} & 1-\mu < r_i < 1 \\ 0 & \text{otherwise} \end{cases}$$

With the uniform density for r_i (i.e. $f(r_i) = 1$), total probability to win can be calculated as $\int_{r_i=0}^{r_i=1} P(\text{win}|r_i)dr_i$. Since the participants do not know the values of r_i,s_i before they participate in such competitions, the total probability can be evaluated with the assumption that both r_i and s_i range from 0 to 1. Notice also that since only one participant can win the skills-game, the probability $\int_0^1 \int_0^1 P_2(\text{win}|r_i, s_i)ds_i dr_i$ for a rich participant to win is simply $\frac{1}{N-P} - \frac{P}{N-P}\int_0^1 \int_0^1 P_1(\text{win}|r_i, s_i)ds_i dr_i$.

In Case II, when the rich person has no access to social capital (i.e. all rich participants have $\nu_2 = 0$), the poor participant has the skills-advantage in the game so the probability to win for the poor participant i is

$$\begin{aligned} P(\text{win}|r_i, s_i) &= \prod_{k=1,k\neq i}^{P} P(\mu r_k + (1-\mu)s_k < \mu r_i + (1-\mu)s_i) \prod_{k=P+1}^{N} P(\mu r_k < \mu r_i + (1-\mu)s_i) \\ &= F_X(x_i)^{P-1} \prod_{k=P+1}^{N} P(\mu r_k < \mu r_i + (1-\mu)s_i) \end{aligned} \tag{2}$$

In Case III, when neither the rich nor the poor participate in the game, then the game depends only on raw-talents and all participants have the same probability $\frac{1}{N}$ to win the game.

This is because $\int_{r_i=0}^{r_i=1} P(\text{win}|r_i)dr_i = \frac{r_i^N}{N}|_0^1 = \frac{1}{N}$ where $P(\text{win}|r_i) = \prod_{k=1,k\neq i}^{P} P(r_k < r_i) \prod_{k=P+1}^{N} P(r_k < r_i) = {r_i}^{N-1}$.

In Case IV, when both rich and poor participants set $\nu_1 = \nu_2 = \overline{\nu}$, then the poor participant can win only if her score x_i is above all other participants - rich or poor.

$$P(\text{win}|r_i) = \prod_{k=1,k\neq i}^{P} P(\mu r_k+(1-\mu)s_k < \mu r_i+(1-\mu)s_i) \prod_{k=P+1}^{N} P(\mu r_k+(1-\mu)s_k < \mu r_i+(1-\mu)s_i)$$
$$= F_X(x_i)^{N-1} \quad (3)$$

Using symmetry arguments, this winning probability - which is the same for all participants - also turns out to be $1/N$ for all the participants.

Notice that with the rich and poor participants having disparate incomes y_1 and y_2 (so that $y_1 < y_2$), the above setup evidently puts the poor participants at a disadvantage since they may not afford to pay the costs to enhance their skills (i.e. realise the true worth of their skills s_i due to high participation costs). However, since both raw-talent r_i and social-capital s_i (the latter enhanced by costly institutional support) are randomly distributed in the population, the competition may never rule out the chance for a poor participant to win. For the same reason, the returns from investments to enhance one's skills may also be diminishing since one cannot beat raw-talent and non-enhanced skills after a certain point.

It is worth emphasising also that the effect of expenditure on the winning probability is weakened with higher number of participants in the game where abilities (both skills s_i relevant for social capital and the raw-talent r_i) are uniformly distributed. In other words, if poor and rich participants were to be considered as two separate groups, then the effect of expenditure on their wins declines when more and more participants are sampled with varied raw-talents and skills. Figures 1 and 2 demonstrate this effect for probability calculated with Monte Carlo simulations for increasing values of the number of poor participants (P) and the total number of participants (N) while hold the ratio P/N constant ($P/N = 1/2$). Our Monte Carlo simulations suggest that the role of expenditure is insignificant for $N \geq 30$. We now explore the effect of income differences using using closed-form solution for $N = 2$ and $P = 1$ in Section 3.1. Equivalent results can be derived using Monte-Carlo simulation for higher values of N and P.

3.1 Equilibrium Conditions

We proceed with a solution by setting total number of participants $N = 2$ and the number of poor participants $P = 1$ in Equations 1, 2 and 3 above. As before, a participant takes advantage of the skills with weight $1-\mu$ only by spending an amount $\overline{\nu}$ and can decide to avoid the expenditure altogether in order to invest it towards the savings instead. We first discuss the different probabilities in Case I, II, III and IV before discussing the role of an additive

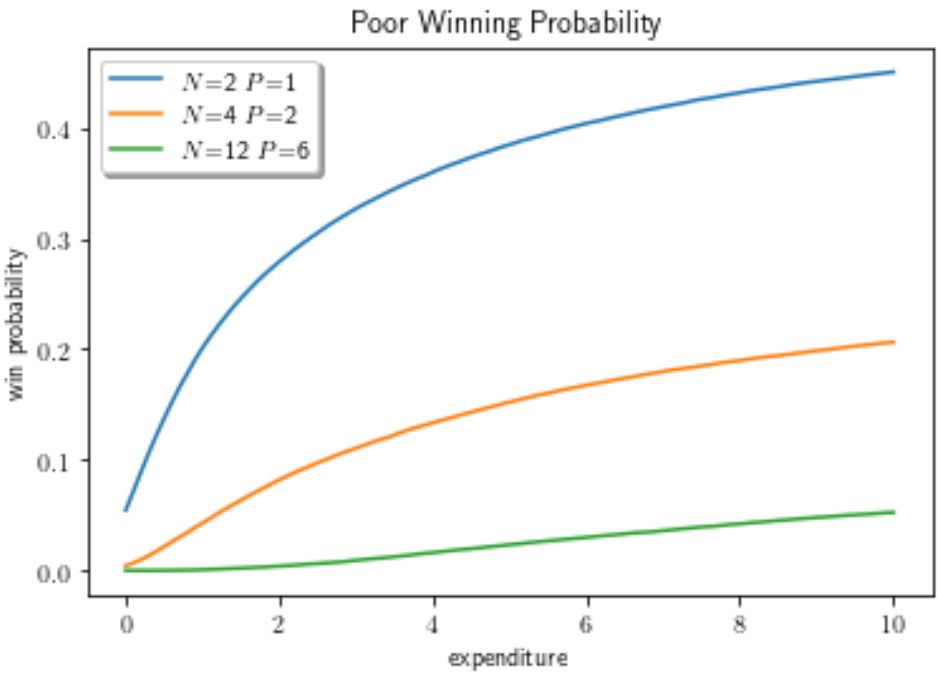


Figure 1: Probability to win for poor consumer with increasing number of participants

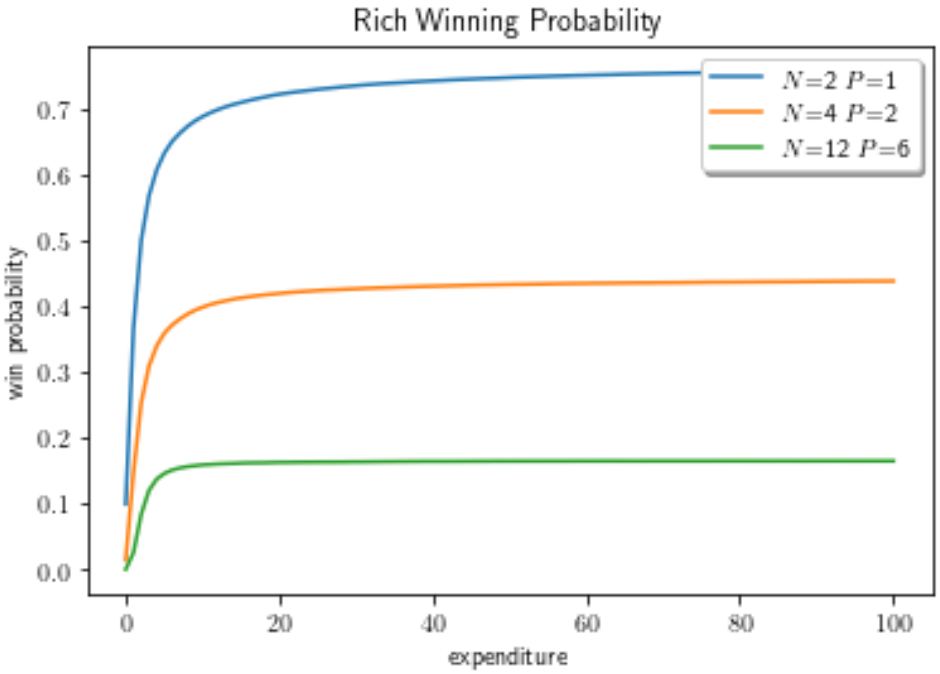


Figure 2: Probability to win for rich consumer with increasing number of participants

intertemporal diminishing utility from accumulated wealth. Recall that the behaviour of both participants depends only on ν_1, ν_2 - as they are unaware of r_i, s_i before they participate.

An important result for the closed-form solution is the probability distribution (CDF) of $x_i \equiv \mu r_i + (1-\mu)s_i$ under assumption $\mu < 1-\mu$ (for independent variables r_i and s_i that are uniformly distributed between 0 and 1) that is detailed in Appendix 5.1.

$$F_X(x) = P(x_k < x) = P(r_k\mu + s_k(1-\mu) < x) = \begin{cases} \frac{x^2}{2\mu(1-\mu)} & 0 < x < \mu \\ \frac{x-\mu/2}{1-\mu} & \mu < x < 1-\mu \\ 1 - \frac{(x-1)^2}{2\mu(1-\mu)} & 1-\mu < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

We now evaluate above probabilities under Case I, II, III and IV discussed above.

Case I

In Case I, the poor participant can win when her raw-talent r_1 is higher than of the total score of rich participant (who competes with acquired skills). The poor participant is priced out of accessing any skills s in her score so that the skills $s > 0$ is only true for the rich participant having $\nu_2 = \overline{\nu}$. Thus, the poor participant with raw-talent r_1 can win with the following probability

$$P(\text{win}|r_1) = P(\mu r_2 + (1-\mu)s_2 < \mu r_1) = F_X(\mu r_1)$$

Notice that $\mu r_1 < \mu < 1-\mu < 1$ is implied due to $\mu < \frac{1}{2}$ (and $r_1 \in (0,1)$). The probability that the total score $\mu r_2 + (1-\mu)s_2$ does not exceed μr_1 implies that only the first part of the piece-wise function for $F_X(x)$ needs to be considered. Therefore, using $P(\text{win}|r_1) = \frac{\mu^2 r_1^2}{2\mu(1-\mu)}$, we evaluate $P_1(\text{win})$ as

$$P_1(\text{win}) = \int_0^1 P(\text{win}|r_1)dr_1 = \frac{\mu}{6(1-\mu)}$$

Since only one of the participants can win, the probability $P_2(\text{win})$ for the rich participant to win is simply $1 - P_1(\text{win})$. The following thus applies to all other cases (i.e. Cases II, III and IV).

$$P_2(\text{win}) = 1 - P_1(\text{win}) \tag{4}$$

Case II

In Case II, the rich person does not participate ($\nu_2 = 0$) and both the poor participants have the skills-advantage in the game. A poor participant wins by having the total score higher than the raw-talent of the rich participant (who does not acquire skills with investment). Representing $x_i \equiv \mu r_i + (1-\mu)s_i$, we have the following probability for the poor participant to win

$$P(\text{win}|r_1, s_1) = P(\mu r_2 < \mu r_1 + (1-\mu)s_1)$$

Notice that $P(\mu r_2 < \mu r_1 + (1-\mu)s_1)$ (i.e. $P(r_2 < r_1 + (\frac{1}{\mu} - 1)s_1)$) is 1 after $r_1 + (\frac{1}{\mu} - 1)s_1 > 1$ (or $\mu r_1 + (1-\mu)s_1 > \mu$) and $\frac{1}{\mu}$ in the interval $(0, \mu)$ (this is because $r_2 \in (0, 1)$). In other words,

$$P(r_2 < r_1 + (1/\mu - 1)s_1) = \begin{cases} \frac{x_1}{\mu} & 0 < x_1 < \mu \\ 1 & \text{otherwise} \end{cases}$$

The probability to win for the poor participant can be written as (see Appendix 5.1)

$$P(\text{win}|r_1, s_1) = \begin{cases} \frac{x_1}{\mu} & 0 < x < \mu \\ 1 & \mu < x_1 < 1-\mu \\ 1 & 1-\mu < x_1 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Using double-integral properties (see Appendix 5.2), the total probability to win $P_1(\text{win})$ for a poor participant is evaluated as follows

$$\begin{aligned} P_1(\text{win}) = \int_{s_1=0}^{s_1=1}\int_{r_1=0}^{r_1=1} P(\text{win}|r_1, s_1)ds_1dr_1 &= \int_{r_1=0}^{r_1=1}\int_{s_1=0}^{s_1=\frac{\mu(1-r_1)}{1-\mu}} \frac{(\mu r_1 + (1-\mu)s_1)}{\mu} ds_1dr_1 \\ &+ \int_{r_1=0}^{r_1=1}\int_{s_1=\frac{\mu(1-r_1)}{1-\mu}}^{s_1=1-\frac{\mu r_1}{1-\mu}} 1\, ds_1dr_1 \\ &+ \int_{r_1=0}^{r_1=1}\int_{s_1=1-\frac{\mu r_1}{1-\mu}}^{s_1=1} 1\, ds_1dr_1 \\ &= \frac{\mu}{3(1-\mu)} + \frac{1-2\mu}{1-\mu} + \frac{\mu}{2(1-\mu)} \\ &= \frac{2\mu}{6(1-\mu)} + \frac{6(1-2\mu)}{6(1-\mu)} + \frac{3\mu}{6(1-\mu)} = \frac{6-7\mu}{6(1-\mu)} \end{aligned}$$

The probability $P_2(\text{win})$ for the rich participant to win follows from Equation 4.

Case III

In Case III, when neither the rich nor the poor participate in the game, then the game depends only on raw-talents and both participants have the same probability $\frac{1}{2}$ to win the game. This is because $\int_{r_i=0}^{r_i=1} P(\text{win}|r_i)dr_i = \frac{r_i^2}{2}|_0^1 = \frac{1}{2}$. Both the poor and rich participant have the same total probability $\frac{1}{2}$ to win.

Case IV

In Case IV, when both rich and poor participants set $\nu_1 = \nu_2 = \overline{\nu}$, then the poor participant can win only if her score x_i is above that of the rich participant. Therefore,

$$P(\text{win}|r_1, s_1) = P(\mu r_2 + (1-\mu)s_2 < \mu r_1 + (1-\mu)s_1) \\ = F_X(x_1)$$

$$P(\text{win}|r_1, s_1) = \begin{cases} \frac{x_1^2}{2\mu(1-\mu)} & 0 < x_1 < \mu \\ \frac{x_1 - \mu/2}{1-\mu} & \mu < x_1 < 1-\mu \\ 1 - \frac{(x_1-1)^2}{2\mu(1-\mu)} & 1-\mu < x_1 < 1 \\ 0 & \text{otherwise} \end{cases}$$

While this integral can be evaluated based on symmetry arguments (since x_i is also randomly distributed), the integral $\int_{s_1=0}^{s_1=1}\int_{r_1=0}^{r_1=1} P(\text{win}|r_1, s_1)dr_1 ds_1$ also evaluates to $\frac{1}{2}$ (see Appendix 5.2)

$$\begin{aligned} \int_{s=0}^{s=1}\int_{r=0}^{r=1} F_X(x_1)ds_1 dr_1 &= \int_{r_1=0}^{r_1=1}\int_{s_1=0}^{s_1=\frac{\mu(1-r)}{1-\mu}} \frac{(\mu r_1 + (1-\mu)s_1)^2}{2\mu(1-\mu)}\, ds_1 dr_1 \\ &+ \int_{r_1=0}^{r_1=1}\int_{s_1=\frac{\mu(1-r)}{1-\mu}}^{s_1=1-\frac{\mu r}{1-\mu}} \left(\frac{\mu r_1 + (1-\mu)s_1 - \mu/2}{1-\mu}\right) ds_1 dr_1 \\ &+ \int_{r_1=0}^{r_1=1}\int_{s_1=1-\frac{\mu r}{1-\mu}}^{s_1=1} \left(1 - \frac{(\mu r_1 + (1-\mu)s_1 - 1)^2}{2\mu(1-\mu)}\right) ds_1 dr_1 \\ &= \frac{\mu^2}{8(1-\mu)^2} + \frac{1-2\mu}{2(1-\mu)} + \frac{(4-5\mu)\mu}{8(1-\mu)^2} = \frac{4(1-\mu)^2}{8(1-\mu)^2} = \frac{1}{2} \end{aligned}$$

As is clear from Figure 3 and above results (see summaries in Tables 3 and 4), the condition $\mu < 1/2$ - indicating that curated skills are given more importance that raw-talent in the

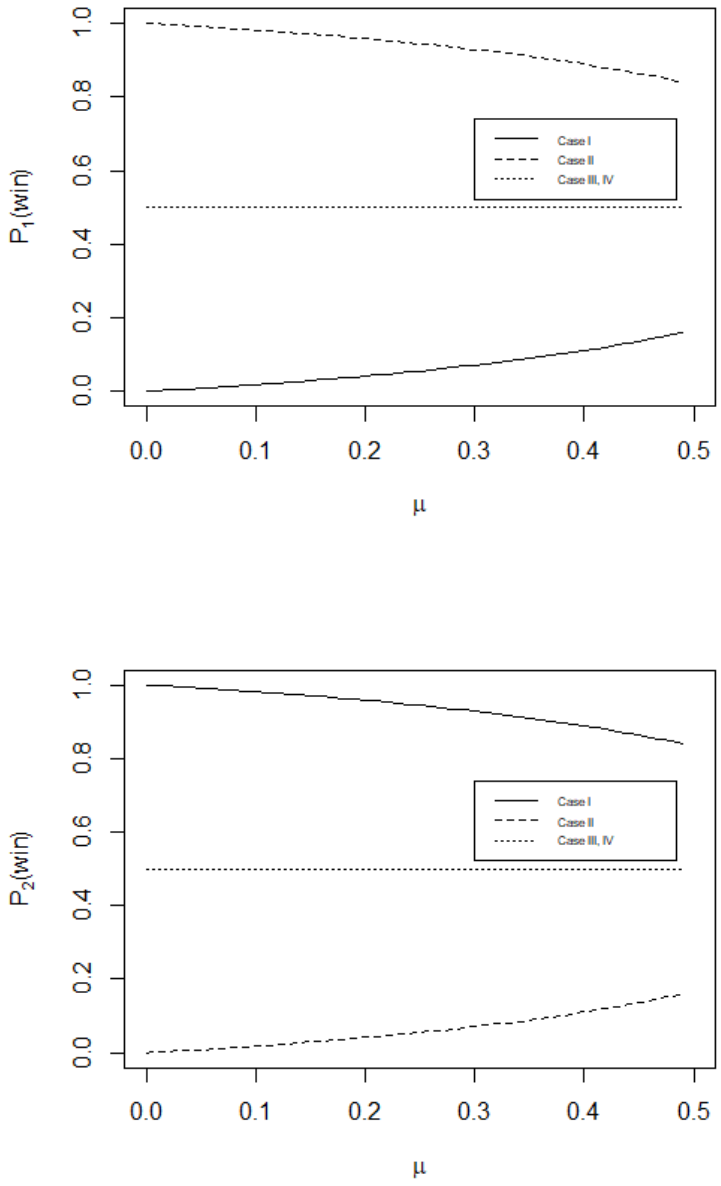


Figure 3: Probability of winning $P_1(\text{win})$, $P_2(\text{win})$ for poor and rich participants with varying μ

competitions - means that the poor participant is strictly better off (than facing probability $1/2$) if she participates while the rich participant doesn't participate (Case II). Also, the poor participant is better off participating for $\mu < \frac{1}{2}$ when the rich participant does participate as it raises the probability to $1/2$. That the competition becomes uncompetitive both when threshold $\overline{\nu}$ falls too much (letting both participants participate) or rises too much (letting both both participants withdraw) also follows from the properties of the winning probability.

Notice that the rich participant could withdraw and face either $1/2$ or lower probability (when the poor participant participates). By participating, therefore, the rich participant would either face probability $1/2$ (when poor participant also participates) or better (if the poor participant withdraws). Considering the probabilities alone, the rich participant would always participate. If the rich participant always participates, then so long as budget constraints permit, the poor participant is also better off participating since not participating brings her probability to win below $1/2$. If the utility was from probability alone, then both participants would participate and face $1/2$ probability.

However, it's not the probability to win that the participants are likely to optimise (maximise). The participants are more likely to face a finite life-time and utility from accumulated assets instead. We therefore explore the case where the participants optimise an intertemporally additive utility over their life-time. The life-time in the model is represented with two time-horizons - where the second-time horizon is one before the end of life when all wealth accumulated by either participant must disappear. With no accumulations at the end of life, both rich and poor participants must decide to spend all their incomes y_1 or y_2 (for poor and rich participants respectively so that $y_1 < y_2$) towards the promotion game. In the first-stage, however, the participants decide the amount to be spent towards promotion based upon the utility from wealth they obtain across the two time-horizons that span their lifetime. For the poor participant, the total utility to be optimised (assuming EUT) under the intertemporally additive utility $u(\cdot)$, a winning probability function $W(x)$ (for expenditure x) and a discount-factor δ is

$$u(y_1 - \nu_1) + \frac{1}{\delta}((1 - W(y_1))u(y_1) + W(y_1)u(y_2))$$

Similarly, for the rich participant, the total utility to be optimised (assuming EUT) is

$$u(y_2 - \nu_2) + \frac{1}{\delta}((1 - W(y_2))u(y_1) + W(y_2)u(y_2))$$

		Poor participant *withdraws*	Poor participant *participates*
Rich participant	*withdraws*	Case III	Case II
Participant	*participates*	Case I	Case IV

Table 1: Cases I,II, III, IV as participant strategies

		Poor participant *withdraws*	Poor participant *participates*
Rich participant	*withdraws*	$u(y_2) + \frac{(1-\frac{P}{N})u(y_2) + \frac{P}{N}u(y_1)}{\delta}$	$u(y_2) + \frac{u(y_2)p_2' + (1-p_2')u(y_1)}{\delta}$
		$u(y_1) + \frac{(1-\frac{P}{N})u(y_2) + \frac{P}{N}u(y_1)}{\delta}$	$u(y_1 - \overline{\nu}) + \frac{u(y_2)p_1' + (1-p_1')u(y_1)}{\delta}$
	participates	$u(y_2 - \overline{\nu}) + \frac{p_2 u(y_2) + (1-p_2)u(y_1)}{\delta}$	$u(y_2 - \overline{\nu}) + \frac{(1-\frac{P}{N})u(y_2) + \frac{P}{N}u(y_1)}{\delta}$
		$u(y_1) + \frac{p_1 u(y_2) + (1-p_1)u(y_1)}{\delta}$	$u(y_1 - \overline{\nu}) + \frac{(1-\frac{P}{N})u(y_2) + \frac{P}{N}u(y_1)}{\delta}$

Table 2: Utility from two periods for (rich,poor) participants - at the start of the first-period (with P poor participants among N total participants)

With above utilities, we discuss conditions for whether the Case I, II, III, IV can become Nash equilibria (see Table 1). We then proceed to discuss the implications for the effect of income differences on expenditure towards status. It is worth remarking that our discussion of the effect of the threshold $\overline{\nu}$ and the income differences $y_2 - y_1$ on participation (expenditure) assumes that the exogenous parameter μ is stable with respect to short-term changes in threshold investment $\overline{\nu}$ as well as disposable incomes y_1, y_2.

For ease of notation we define p_1, p_2, p_1' and p_2' - which correspond to the **two-player** model

discussed above.

$$
\begin{aligned}
p_1 &\equiv \frac{\mu}{6(1-\mu)} \\
p_2 &\equiv 1 - p_1 \\
p_1' &\equiv \frac{6-7\mu}{6(1-\mu)} \\
p_2' &\equiv 1 - p_1'
\end{aligned}
\tag{5}
$$

The consumer choices for the poor and rich participants using a general intertemporally additive function $u(A)$ (for assets A) are shown in Table 2. As we can verify from Figure 3 that we have $p_1 < 1/2 < p_2$ and $p_1' > 1/2 > p_2'$ for all values of $\mu < 1/2$.

$$
\begin{aligned}
&\text{Given } \mu < 1/2, \text{ we have} \\
&p_1 < 1/2 < p_2 \\
&p_1' > 1/2 > p_2'
\end{aligned}
\tag{6}
$$

Notice the Table 2 represents the condition when the budget is non-binding i.e. the poor participant is able to spend $\overline{\nu}$. i.e. $\overline{\nu} < y_1 < y_2$. When the poor participant faces a binding constraint i.e. $y_1 < \overline{\nu} < y_2$, then the rich participant knows that the best case probability to win for the poor participant is $\frac{1}{2}$. The rich participant's decision to participate would depend only on utility from savings and the relative advantage from gain in probability. This is elaborated further in Section 3.1.1.

Since the number of poor participants is likely to be higher than that of rich participants in a typical competition, the unequal number of poor and rich participants seems also worth considering in the model. Setting $N = 3, P = 2$ in the equations 1, 2 and 3, for example, we can interpret the **three-player** model with the following values of p_1, p_2, p_1' and p_2'

$$
\begin{aligned}
p_1 &= \frac{\mu}{8(1-\mu)} \\
p_2 &= 1 - 2p_1 \\
p_1' &= \frac{19\mu^2 - 40\mu + 20}{40(1-\mu)^2} \\
p_2' &= 1 - 2p_1'
\end{aligned}
\tag{7}
$$

Case	Description	Probability of poor participant to win	Response to μ
I	Only rich invest in skills	$\frac{\mu}{6(1-\mu)}$	With higher μ, the participants with raw-talent get rewarded. The probability to win increases with higher μ for the poorer participant.
II	Only poor invest in skills	$\frac{6-7\mu}{6(1-\mu)}$	With higher μ, the poor participants with access to skills get rewarded. Therefore, The probability to win decreases with higher μ for the poorer participant.
III,IV	Neither or Both invest in skills	$\frac{1}{2}$	Probabilty to win is independent of μ.

Table 3: The probability of winning for the poor participant in the 3-player model

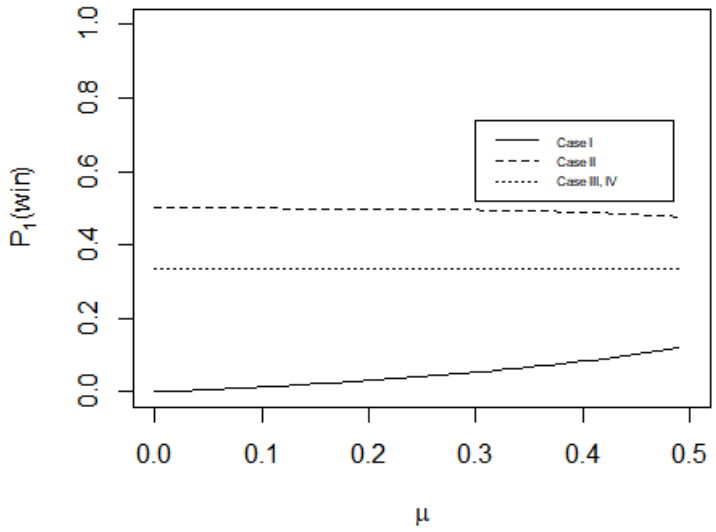
Case	Description	Probability of rich participant to win	Response to μ
I	Only rich invest in skills	$1 - \frac{\mu}{6(1-\mu)}$	With higher μ, the participants with raw-talent get rewarded. Therefore, the probability to win decreases with higher μ for the rich participant.
II	Only poor invest in skills	$1 - \frac{6-7\mu}{6(1-\mu)}$	With higher μ, the poor participants with access to skills get rewarded. Therefore, the probability to win increases with higher μ for the rich participant.
III,IV	Neither or Both invest in skills	$\frac{1}{2}$	Probabilty to win is independent of μ.

Table 4: The probability of winning for the rich participant in the 3-player model

Comparing these probabilities with those from the 2-player model in Equation 5 (also listed in Tables 3 and 4), we see that the probability to win for the poor participant in the 3-player model is lower with respect to the Case III probability (i.e. $1/N$). Likewise, the probability for the rich participant to win is higher in the 3-player model with respect to the Case III probability $(1/N)$. The implications for the unequal number of poor and rich participants are further discussed in Section 3.1.2.

3.1.1 Binding constraint

Assuming EUT participants we first consider the simpler case of binding constraint $y_1 < \overline{\nu} < y_2$. Since we assume a stability in long-term participant behaviour (which itself is a result of observed probability of success), mixed-strategies are not relevant in the current set up. Considering pure strategies for participation, the rich participant would thus participate only when the utility from participating would be higher than from not participating i.e. $u(y_2 - \overline{\nu}) + \frac{p_2 u(y_2)+(1-p_2)u(y_1)}{\delta} > u(y_2) + \frac{u(y_2)+u(y_1)}{2\delta}$. This results in the following condition



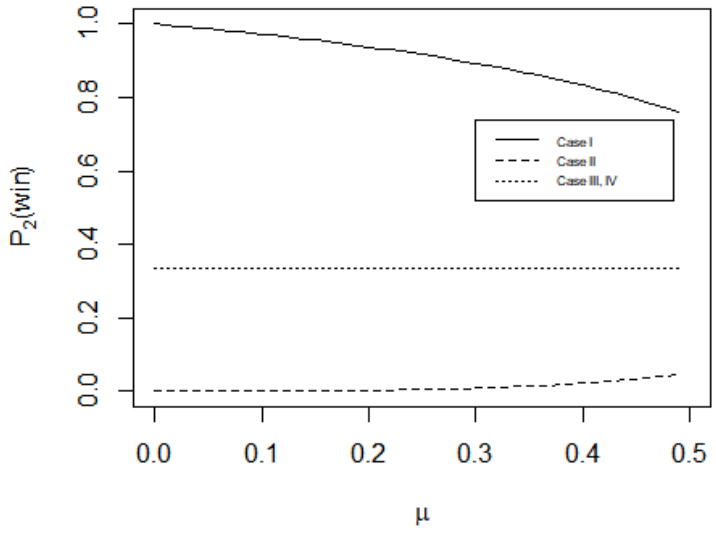


Figure 4: Probability of winning $P_1(\text{win})$, $P_2(\text{win})$ for poor and rich participants with varying μ

$$\frac{(p_2 - 1/2)(u(y_2) - u(y_1))}{\delta} > u(y_2) - u(y_2 - \overline{\nu})$$

With income differences being the same an increase in participation cost $\overline{\nu}$ discourages participation since RHS increases with $\overline{\nu}$ ($\because u(y_2 - \overline{\nu})$ is decreasing in $\overline{\nu}$) and may fail the above condition. Similarly, with participation cost $\overline{\nu}$ being the same, an increase in income-difference $y_2 - y_1$ is in the rich participant's interest whereas a narrowing of income differences reduces incentive for the participant to participate by possibly failing the above condition. All else being the same, higher impatience (lower time-discounted value of money or a higher δ) would also discourage participation. Therefore, when budget is binding for the poor participant i.e. if the poor participant cannot participate due to $y_1 < \overline{\nu}$, then narrower income differences, higher participation costs and higher patience for savings (common for both rich and poor participants) all encourage the rich participant to withdraw from the game. This also means that wider income differences ($y_2 - y_1$) combined with lower participation costs ($\overline{\nu}$) may together help maintain the equilibrium where only the rich participates. Similarly, wider income differences (higher $y_2 - y_1$) and more patience for savings (higher δ) would together maintain the incentive for the rich participant to participate. These conclusions very much fit with the sociological observations made in the literature about status (Hirsch, 1977; Galbraith, 2001).

3.1.2 Non-Binding constraint

When budget is non-binding, then participation of the rich participant is her dominant strategy if the condition is also $u(y_2) + \frac{u(y_2)p_2' + (1-p_2')u(y_1)}{\delta} < u(y_2 - \overline{\nu}) + \frac{u(y_2)+u(y_1)}{2\delta}$) i.e.

$$\frac{(1/2 - p_2')(u(y_2) - u(y_1))}{\delta} > u(y_2) - u(y_2 - \overline{\nu})$$

The inferences on how $y_2 - y_1$, $\overline{\nu}$ and δ influence the participation (see Section 3.1.1) remain the same for the above condition. To discuss the more general Nash equilibria, we now consider the conditions for all of the Cases I, II, III and IV to become a Nash equilibrium. For ease of notation we define the following for a given $\overline{\nu}, y_1$ and y_2

$$U_1 \equiv u(y_1) - u(y_1 - \overline{\nu})$$

$$U_2 \equiv u(y_2) - u(y_2 - \overline{\nu})$$

One could view U_1 and U_2 as the immediate gain in wealth utility from withdrawing. If $u(\cdot)$ represents a diminishing utility, then it is also necessary for $y_1 < y_2$ that

$$U_2 < U_1$$

Consider now that conditions for Case I to be a Nash Equilibrium. If Case I were to be a Nash equilibrium, then for the rich participant's decision to participate, the poor participant must get more utility from withdrawing than from participating and for poor participant's decision to withdraw, the rich must get more utility from participating than from withdrawing. The following conditions [2] must therefore hold simultaneously

$$\begin{aligned} \frac{(\frac{1}{N} - p_1)(u(y_2) - u(y_1))}{\delta} &< U_1 \\ U_2 &< \frac{(p_2 - 1/N)(u(y_2) - u(y_1))}{\delta} \end{aligned} \tag{8}$$

Since $p_2 - \frac{1}{N} \geq \frac{1}{N} - p_1$ (strict inequality follows when the number of poor participants P is higher than that for rich participants $N - P$ e.g., with $N = 3$ and $P = 2$), this is a fairly general condition for the equilibrium. An increase in U_1 due to higher $\overline{\nu}$ does not change the equilibrium conditions so long as the corresponding rise U_2 (due to higher $\overline{\nu}$) does not affect the second inequality. In other words, a higher $\overline{\nu}$ favours the equilibrium (Case I) so long as there is an incentive for the rich to participate. Likewise, a decrease in $\overline{\nu}$ would be favourable for the equilibrium (Case I) so long as the poor participants do not also have an incentive to participate.

Similar effects can be seen with income differences $y_2 - y_1$ (which increase $u(y_2) - u(y_1)$ as well as $\frac{U_2}{U_1}$) if we rewrite Equation 8 as

$$\begin{aligned} \frac{(\frac{1}{N} - p_1)}{\delta} &< \frac{U_1}{u(y_2) - u(y_1)} \\ \frac{U_2}{(u(y_2) - u(y_1))} &< \frac{(p_2 - 1/N)}{\delta} \end{aligned}$$

Limiting our attention to cases when $\overline{\nu} < y_1$, higher y_2 for fixed y_1 decreases $\frac{U_2}{u(y_2)-u(y_1)}$and makes it more likely for the rich to participate while it also decreases$\frac{U_1}{u(y_2)-u(y_1)}$ by raising the stakes enough for the poor to participate (making it less likely for first inequality to be fulfilled). It is easy to see that higher impatience δ encourages both poor and rich to withdraw. As we shall see, equilibrium conditions corresponding to Case II are far more stringent.

[2] $u(y_1) + \frac{p_1 u(y_2)+(1-p_1)u(y_1)}{\delta} > u(y_1 - \overline{\nu}) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} \Rightarrow u(y_1) - u(y_1 - \overline{\nu}) > \frac{(\frac{1}{N}-p_1)(u(y_2)-u(y_1))}{\delta}$ and $u(y_2 - \overline{\nu}) + \frac{p_2 u(y_2)+(1-p_2)u(y_1)}{\delta} > u(y_2) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} \Rightarrow u(y_2) - u(y_2 - \overline{\nu}) < \frac{(p_2-1/N)(u(y_2)-u(y_1))}{\delta}$

If Case II is a Nash Equilibrium, then for rich's decision to withdraw, the poor participant should get more utility from participating than from withdrawing and for poor participant's decision to participate, the rich participant should get more utility from withdrawing than from participating. This means that the following hold [3] simultaneously,

$$U_1 < \frac{(p_1' - 1/N)(u(y_2) - u(y_1))}{\delta}$$
$$U_2 > \frac{(1/N - p_2')(u(y_2) - u(y_1))}{\delta}$$

Since $\frac{(p_1'-1/N)(u(y_2)-u(y_1))}{\delta} < \frac{(1/N-p_2')(u(y_2)-u(y_1))}{\delta}$, the above cannot be satisfied so long as we have a diminishing utility which necessitates $U_2 < U_1$. Thus an equilibrium with Case II which incentivises the poor to participate but not the rich is possible only with a reference dependent utility with $U_1 < U_2$. With $U_1 < U_2$, we can write above as

$$\frac{U_1}{u(y_2) - u(y_1)} < \frac{(p_1' - 1/N)}{\delta} < \frac{(1/N - p_2')}{\delta} < \frac{U_2}{u(y_2) - u(y_1)}$$

Notice that despite a reference-dependent utility, a higher $\overline{\nu}$ or more impatience (δ) may disincentivise the poor from participation just the same (leading to Case III). Similarly, higher income differences (an increase in y_2 for fixed y_1) could encourage the rich to participate by decreasing $\frac{U_2}{u(y_2)-u(y_1)}$ (leading to Case IV). Even with the reference dependent utility, the equilibrium with Case II remains a fragile one.

If Case III were to be a Nash-Equilibrium, then for rich participant's decision to withdraw, poor should get more utility from withdrawing than from participating and for poor participant's decision to withdraw, the rich participant should get more from withdrawing than from participating. The following should thus hold[4],

$$U_1 > \frac{(p_1' - \frac{1}{N})(u(y_2) - u(y_1))}{\delta}$$
$$U_2 > \frac{(p_2 - \frac{1}{N})(u(y_2) - u(y_1))}{\delta}$$

Since $p_1' - \frac{1}{N} \leq p_2 - \frac{1}{N}$ (strict inequality when number of poor participants is higher), we have $\frac{(p_1'-\frac{1}{N})(u(y_2)-u(y_1))}{\delta} < \frac{(p_2-\frac{1}{N})(u(y_2)-u(y_1))}{\delta}$ and $U_2 < U_1$. A higher $\overline{\nu}$ pushes both U_2 and U_1 further away from participation but a fall in $\overline{\nu}$, incentivises the rich participant to participate

[3] $u(y_1 - \overline{\nu}) + \frac{p_1' u(y_2)+(1-p_1')u(y_1)}{\delta} > u(y_1) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} \Rightarrow u(y_1) - u(y_1 - \overline{\nu}) < \frac{(p_1'-1/N)(u(y_2)-u(y_1))}{\delta}$
and $u(y_2) + \frac{p_2' u(y_2)+(1-p_2')u(y_1)}{\delta} > u(y_2 - \overline{\nu}) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} \Rightarrow u(y_2) - u(y_2 - \overline{\nu}) > \frac{(1/N-p_2')(u(y_2)-u(y_1))}{\delta}$

[4] $u(y_1) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} > u(y_1 - \overline{\nu}) + \frac{p_1' u(y_2)+(1-p_1')u(y_1)}{\delta} \Rightarrow u(y_1) - u(y_1 - \overline{\nu}) > \frac{(p_1'-\frac{1}{N})(u(y_2)-u(y_1))}{\delta}$
and $u(y_2) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} > u(y_2 - \overline{\nu}) + \frac{p_2 u(y_2)+(1-p_2)u(y_1)}{\delta} \Rightarrow u(y_2) - u(y_2 - \overline{\nu}) > \frac{(p_2-\frac{1}{N})(u(y_2)-u(y_1))}{\delta}$

(leading to Case I).

$$\frac{U_1}{u(y_2)-u(y_1)} > \frac{(p_1' - \frac{1}{N})}{\delta}$$
$$\frac{U_2}{u(y_2)-u(y_1)} > \frac{(p_2 - \frac{1}{N})}{\delta}$$

Higher income differences i.e. higher y_2 for fixed y_1 also decrease $\frac{U_2}{u(y_2)-u(y_1)}$ and incentivise the rich participant to participate (leading to an equilibrium with Case I). Conversely, if the game is not lucrative for the rich it would automatically be infeasible for the poor i.e. if rich withdraws then poor automatically withdraws. If there is any net utility from playing the game, however, Case IV is the likely equilibrium.

If Case IV were to be a Nash Equilibrium, then for rich's decision to participate, the poor participant should get more from participating than from withdrawing and for poor's decision to participate, the rich participant should get more from participating than from withdrawing. Therefore, the following[5] must hold

$$U_1 < \frac{(\frac{1}{N} - p_1)(u(y_2)-u(y_1))}{\delta}$$
$$U_2 < \frac{(\frac{1}{N} - p_2')(u(y_2)-u(y_1))}{\delta}$$

With $\frac{(\frac{1}{N}-p_2')(u(y_2)-u(y_1))}{\delta} < \frac{(\frac{1}{N}-p_1)(u(y_2)-u(y_1))}{\delta}$ and $U_2 < U_1$, this equilibrium rests upon the incentive for the poor participant to participate. So long as the poorer participant finds the competition worthwhile, the rich would also participate.

The above results outline how participation costs influence the investments towards status. For very low $\overline{\nu}$, rich and poor would both participate and for very high $\overline{\nu}$, both would withdraw. For everything in between, the only equilibrium with a diminishing utility from wealth is one where rich participate whereas poor don't.

The above framework also allows us to discuss the implications of inequality for the same total wealth in the economy. As such, a segregation where the poor never participate in the status economy seems less favourable than one where both rich and poor participate in the status competitions. The current chapter explains that the likely outcome of a failure to maintain universal participation in status games (under diminishing utility from wealth) is that of a segregation where only the rich participate in skills leading to economic success. While a reference-based utility (which may be encouraged with segregation) does seem to permit an equilibrium where only the poor are concerned with status games (with respect to

[5] $u(y_1-\overline{\nu}) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} > u(y_1) + \frac{p_1u(y_2)+(1-p_1)u(y_1)}{\delta} \Rightarrow u(y_1) - u(y_1-\overline{\nu}) < \frac{(\frac{1}{N}-p_1)(u(y_2)-u(y_1))}{\delta}$
and $u(y_2-\overline{\nu}) + \frac{\frac{1}{N}u(y_2)+\frac{1}{N}u(y_1)}{\delta} > u(y_2) + \frac{p_2'u(y_2)+(1-p_2')u(y_1)}{\delta} \Rightarrow u(y_2) - u(y_2-\overline{\nu}) < \frac{(\frac{1}{N}-p_2')(u(y_2)-u(y_1))}{\delta}$

skills-enhancement), the result only suggests that a segregation of status games along with that of incomes is another possible long-term outcome of limited participation.

4 Conclusion

We have considered a model where status amounts to a position that only temporarily cements the participant's economic position. With a decision theory framework for uncertainty, the goal of the model has been to investigate role of income difference on expenditures that carry some status while also having some effect on the income mobility of the participants. The endogenisation of status in this manner allows us to use the framework to understand the role of income difference in aspirational consumption under urbanisation effects across economies that have varied status-determining cultural notions as well as development indicators.

The foremost result from the model is that under diminishing utility from wealth, the only two stable equilibria are those where both rich and poor participate in status competitions and those where only rich stand to benefit from status competitions. The model allows us to understand how a withdrawal from expenditure such as education in certain developing economies could be result of a combined effect of both participation costs and income-differences. More specifically, high participation costs push the equilibrium to one where where only rich participate in status-related competitions whereas high income differences may have the unlikely effect of incentivising the poor consumers to participate. In the context of developing economies, lowering of participation costs in education (as a status game) despite rise in income differences may thus also represent a likely escape from segregation of participation in the near-term for many developing economies. The role of a reference-based utility adds an important caveat to these results since contentment at lower levels of wealth could support segregation of competitions at different income levels in the economy.

5 Appendices

5.1 Convolution of Uniform Random Variables

The distribution of a convex combination $w = \mu r + (1-\mu)s$ of two random variables r, s - which are distributed according to densities $f_R(r)$ and $f_S(s)$ - can be derived using the convolution of two derived variables $x = \mu r$ and $y = (1-\mu)s$. If r and s are uniformly distributed variables ranging from 0 to 1, then x ranges from 0 to μ and y ranges from 0 to $(1-\mu)$. The CDF for $F_X(x)$ and $F_Y(y)$ are simply $F_X(x) = \frac{x}{\mu}$, $F_Y(y) = \frac{y}{1-\mu}$ (respectively). Therefore, the densities for x and y are $f_X(x) = \frac{1}{\mu}$ and $f_Y(y) = \frac{1}{1-\mu}$. The density of w is simple $f(w) = \int_{-\infty}^{\infty} f_Y(w-x) f_X(x) dx$. Since , $f_X(x)$ is $1/\mu$ only between 0 and μ. We have the following integral to evaluate for the convolution

$$f(w) = \frac{1}{\mu} \int_0^{\mu} f_Y(w-x) dx$$

Since $f_Y(y)$ is 1 only between 0 and $1-\mu$, we have the condition $0 < w - x < 1 - \mu \Rightarrow w - (1-\mu) < x < w$. Further, x is distributed only between 0 and μ. As shown in Figure 5, the integral would evaluate to the following[6] under the assumption $\mu < \frac{1}{2}$.

$$f(w) = \frac{1}{\mu} \int_0^{\mu} f_Y(w-x) dx = \begin{cases} \frac{1}{\mu} \int_0^{w} \frac{1}{1-\mu} dx & 0 < w < \mu \\ \frac{1}{\mu} \int_0^{\mu} \frac{1}{1-\mu} dx & \mu < w < 1-\mu \\ \frac{1}{\mu} \int_{w-(1-\mu)}^{\mu} \frac{1}{1-\mu} dx & 1-\mu < w < 1 \\ 0 & \text{otherwise} \end{cases}$$

This can be simplified as

$$f(w) = \begin{cases} \frac{w}{\mu(1-\mu)} & 0 < w < \mu \\ \frac{1}{1-\mu} & \mu < w < 1-\mu \\ \frac{1-w}{\mu(1-\mu)} & 1-\mu < w < 1 \\ 0 & \text{otherwise} \end{cases}$$

Figure 6 shows this trapezoidal density with μ set to $\frac{1}{5}$. The CDF $F_X(x)$ can be evaluated by integrating the above density as $P(X < x) = \int_{-\infty}^{x} f(t) dt$.

[6] Notice that as a convex combination, w is distributed only between 0 and 1.

$$F_X(x) = \begin{cases} \frac{x^2}{2\mu(1-\mu)} & 0 < x < \mu \\ \frac{\mu}{2(1-\mu)} + \frac{x-\mu}{1-\mu} & \mu < x < 1-\mu \\ \frac{\mu^2}{2\mu(1-\mu)} + \frac{2\mu(1-2\mu)}{2\mu(1-\mu)} + \frac{\mu^2-x^2+2x-1}{2\mu(1-\mu)} & 1-\mu < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Simplifying above, we have

$$F_X(x) = \begin{cases} \frac{x^2}{2\mu(1-\mu)} & 0 < x < \mu \\ \frac{x-\mu/2}{1-\mu} & \mu < x < 1-\mu \\ 1 - \frac{(x-1)^2}{2\mu(1-\mu)} & 1-\mu < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

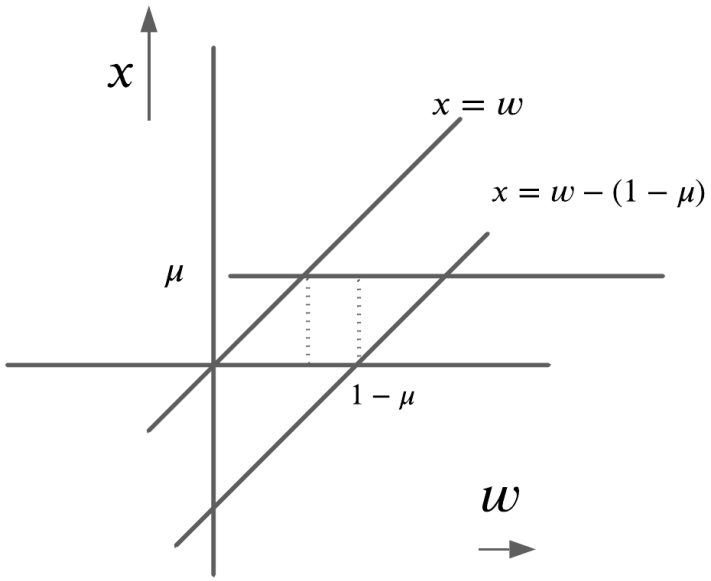


Figure 5: Ranges of w and $w - x$ for convolution of two scaled random variables

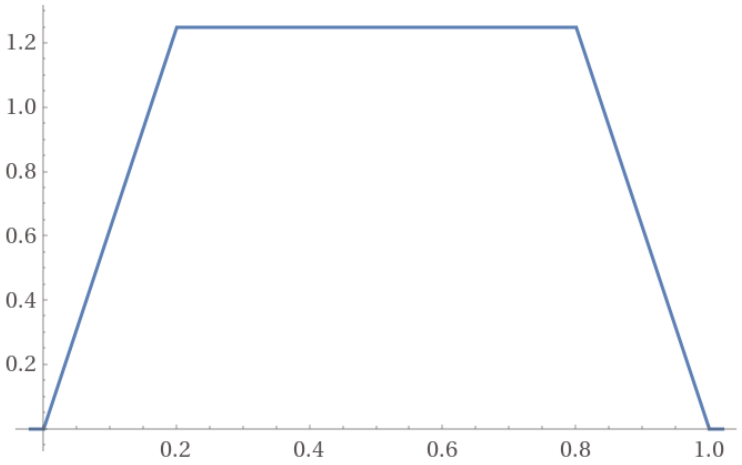


Figure 6: Combined density of w as the convex combination of two random variables with $\mu = \frac{1}{5}$

5.2 Double Integral of $P(\textbf{win}|r, s)$

To calculate $\int_{s=0}^{s=1} \int_{r=0}^{r=1} P(\text{win}|r, s)drds$ over the variable x (where $0 < r < 1$ and $0 < s < 1$), we evaluate the sum of three parts split by the boundaries of the piecewise function. The three parts are shown in Figure 7. If $x = \mu r + (1 - \mu)s$ and $F(x)$ is of the following piece-wise form

$$F(x) = \begin{cases} A & 0 < x < \mu \\ B & \mu < x < 1 - \mu \\ C & 1 - \mu < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Then the integral $\int_{s=0}^{s=1} \int_{r=0}^{r=1} F(x)drds$ can be written as

$$\int_{s=0}^{s=1} \int_{r=0}^{r=1} F(x)dsdr = \int_{r=0}^{r=1} \int_{s=0}^{s=\frac{\mu(1-r)}{1-\mu}} A\, dsdr + \int_{r=0}^{r=1} \int_{s=\frac{\mu(1-r)}{1-\mu}}^{s=1-\frac{\mu r}{1-\mu}} B\, dsdr + \int_{r=0}^{r=1} \int_{s=1-\frac{\mu r}{1-\mu}}^{s=1} C\, dsdr$$

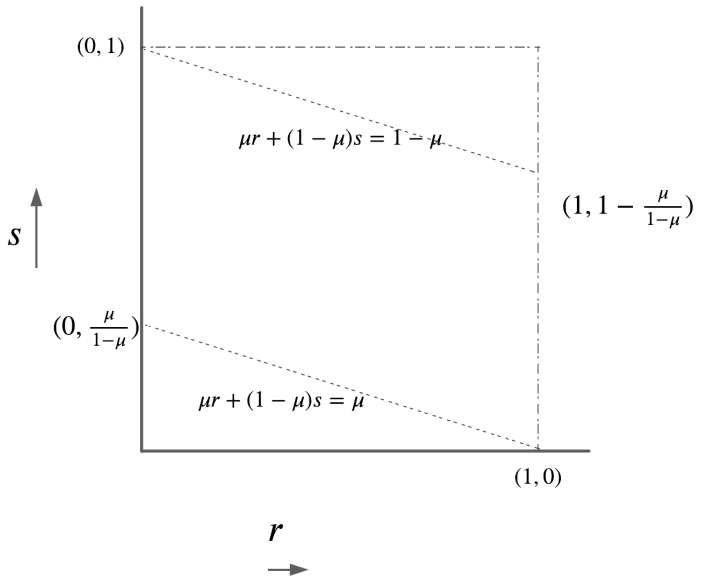


Figure 7: Integration over r, s split by boundaries $\mu r+(1-\mu)s = \mu$ and $\mu r+(1-\mu)s = 1-\mu$

Part III

Variation in Quality of Food Consumption in Tanzania

Abstract

While consumption of basic needs may have hardly any status value in societies where basic consumption needs have been fulfilled, food continues to carry some status value in societies undergoing economic transitions where it is either scarce or expensive. The presented chapter explores the role of permanent income in demand substitution for food quality in Tanzania where food continues to carry a status significance. Comparing effects of permanent income and that of prices of food commodities, we find a higher price-based quality associated with meats and fruits to be aligned with the income disparities as well as with the availability of electricity in Tanzania - suggesting that the access to urban amenities also significantly differentiates higher quality food consumption in the economy.

1 Introduction

While aspirational consumption is more often associated with extravagant items of luxury, there seems strong evidence for aspirational items becoming wider necessities in the population (Burke, 1996; Barnett, 2005; Silverstein, Fiske, & Butman, 2008; Bekir, El Harbi, & Grolleau, 2013). As the access to high quality items expands to poorer sections of the wider society, the notion of whether a certain item is aspirational or not also changes with time (Barnett, 2005; Silverstein et al., 2008; Bekir et al., 2013). In the context of sub-Saharan Africa (SSA), the current chapter examines what demand for higher quality may imply for trends in aspirational consumption. Using food as the basic need for this examination, the chapter asks two particular questions that may help evaluate the expected rise of high quality for basic needs in SSA (Srivastava et al., 2020). First, it asks whether certain high-quality items stand out in food consumption in a typical economy of sub-Saharan Africa (SSA) or not. Second, it asks whether permanent income in urban as well as non-urban areas significantly limits higher quality consumption for such items of high quality. If the effects of permanent income on higher quality are strong, then an overall rise in quality may largely depend on how newer food items are introduced in urban and rural markets. The chapter argues in favour of the conventional argument that a higher quality being limited to richer sections of society puts evident limits on the rise in consumption of such items (Hirsch, 1977). To explore permanent income effects on high price-based quality, the chapter employs a conventional quantity-based framework - while using a dietary diversity standard to test whether certain items of higher quality stand out in the survey of quality or not. It also tests whether the quality consumed may be higher in more urbanised regions or not.

The analysis of food as an unlikely category of aspirational consumption in the current chapter is motivated by two observations. First, there seem compelling results from sociological surveys for status associated with food in certain economies of SSA where a transformation from agrarian to industrial societies is under way (Weismantel et al., 1989; Pottier et al., 1999; Onuorah & Ayo, 2003; I. L. Ohna, 2007) . Second, a significant portion of household budget being spent on food implies that the variation in food quality could detail the effects of rising urbanisation on diets in a typical developing economy (Onuorah & Ayo, 2003). A survey of variation in quality is of significance for studying sub-Saharan African economies for other reasons as well. The strong preferences that we observe in the current chapter towards certain items - preferences which raise the overall food quality - are of interest from the perspective of food policy due to the prevalent link between SES and poor dietary diversity (DD) in the region (Hatløy, Hallund, Diarra, & Oshaug, 2000; Hoddinott & Yohannes, 2002; Ruel, 2003; Arimond & Ruel, 2004; M'Kaibi, Steyn, Ochola, & Du Plessis, 2017; Hailu & Woldemichael, 2019; Modjadji, Molokwane, & Ukegbu, 2020; Obayelu & Osho, 2020; Agada, Igbokwe, et al., 2015; Ogechi & Chilezie, 2017). A higher consumption of fruits and meats - which we find for urban and richer households - also has

implications for land use and environmental concerns through increases in meat production and consumption.

While extreme variations in food quality and diet diversity are observed many economies in SSA, the recent transition to urbanisation and compelling results from sociological surveys in support of relevance of food in status indication (I. L. Ohna, 2007)present Tanzania as an appropriate case-study for a study of variation in quality with respect to inequalities. Other than the sociological studies of such signaling value through food in both the pastoralist and agriculturalist communities (I. L. Ohna, 2007), the recent experience of higher consumption in Tanzania leaves wide disparities in asset ownership and urbanisation levels - a disparity that allows us to understand the role that variation in urban developments may play towards consumption of basic needs. The ongoing urbanisation also leaves a significant presence of semi-urban areas in Tanzania due in part to an incomplete transition away from subsistence farming in the economy (Pingali & Sunder, 2017; Wenban-Smith, Fasse, & Grote, 2016) - a scenario that allows us to understand the permanent income effects against urbanisation at a finer level of granularity.

A fundamental contention that we maintain for implications of permanent income effects on food demand and rise in consumption is that higher quality in any given food category may not be an instance of excessive consumption. This is foremost because the notion of food-related basic needs spans more than the minimal calorific intake required for survival (Doyal & Gough, 1991)- allowing for consumption of quality at the cost of its quantity for any particular food category to comprise of basic necessities. Further, even if consumers are compelled to increase quality in certain food categories rather than others (for reasons unobserved), one needs a specific normative criteria to argue that quality substitution of a particular food category with another is actually excessive. Instead of determining whether the quality consumed is excessive or not, our focus in the chapter remains on assessing whether consumption of higher quality is narrowed down to specific food categories relevant for food diversity or not and on explaining what that may imply for consumption trends. While we do use the FAO-specified nutrition categories (Kennedy, Ballard, & Dop, 2011) to help assess the extent to which higher quality in specific categories may be aligned with recommended food diversity, the FAO-categories are used only to provide a qualitative view of variation in quality while avoiding any particular judgment on what is considered necessary for the consumer. The empirical goal of the chapter is thus limited to testing whether higher permanent income results in a higher demand for quality across FAO-provided food categories in Tanzania or not. The chapter discusses the implications for aspirational consumption of factors other than permanent income - such as the role of availability of basic amenities and urbanisation. It argues that while high quality consumption may rise in future, a segregation of food demand based on high or low permanent income may depend on how newer food items are introduced in the market.

The analysis of variation in food quality relies on a robust framework for quality in demand

- where a measure for quality across commodities (such as fruits, fat, meats-proteins etc.) upholds the theoretical restrictions imposed by aggregation of demand across various goods (such as coconut, mango etc. that may be classified within a commodity:fruits) that are part of a commodity group (such as fruits). More specifically, the method follows the restrictions implied due to the Hick's commodity theorem (Hicks, 1946; Deaton & Muellbauer, 1980b; Nelson, 1991) which requires the physical quantities of goods consumed to be weighted by unchanging (relative) prices of the constituent goods before the quantities can be aggregated into the demand for group whose quality is measured. Thus, the quality of fruits as commodity - for example - can be evaluated only as long as the goods - coconut, mango etc. that may constitute the fruits commodity - exhibit a stable relative price-structure. To avoid confusion in the terminology associated goods and quality, we use the nomenclature provided by Cramer (1973) where "goods" or "items" are variants with different quality that all fall under a certain "commodity" (Cox & Wohlgenant, 1986).

The particular theoretical approach is important for our analysis of demand for quality due to two reasons. First, the approach is preferable over an arbitrary aggregation of quality from groups that does not follow the aggregation restrictions in a demand equation. Such restrictions may limit our ability to talk about substitution in quality across commodity groups - which is the core part of our assessment of how quality availed of certain food categories may be higher than others. Second, a robust aggregation based on market-prices also helps circumvent measurement issues arising out of use of unit-prices (Gibson & Kim, 2019).

The primary criterion of grouping the demand from goods consumed into a commodity group using the approach is a stable relative price-structure of market prices that are available in the consumption survey[7]. A classification of food items into commodities for which quality metrics are calculated is thus undertaken using past observed-market prices of all goods available in the survey alone - where a quality metric is associated with each food commodity such that the constituent goods of every commodity (which quality is calculated for) move together in terms of their annual prices[8] from 2008 to 2014. This is a departure from the unit-value approaches where average expenditure itself is considered as a price-control. It needs to be emphasised also that the groups of food-commodities over which quality is evaluated are determined based on a similar change in prices (rather than the prices themselves i.e. the classification does not place all high-price items in one group and low-price groups in another). The condition of a stable relative prices structure (due to Hicks Commodity Theorem) drives a clustering of relative price changes across all goods for food in the survey that we describe in 3.1. The FAO categories are then used as the nutritional basis upon which we assess how the food demand may have higher quality in certain commodity groups vs others. Thus, the commodity groups obtained from the

[7]The prices observed in local markets are those in the district where the household - whose consumption is recorded for budget shares and quality measures - is located.

[8]A locally observed price associated with an item is one that is observed for the item in the district where the household is located.

clustering are further divided based on the twelve FAO (Food and Agriculture Organization of the United Nations) categories - (i) cereals, (ii) vegetables, (iii) fruits, (iv) meat, (v) eggs (vi) fish and other sea foods (vii) legumes, nuts and seeds, (viii) milk and milk products (ix) oil and fats, (x) sweets (xi) spices, condiments and beverages and (xii) tubers and roots (Kennedy et al., 2011) in order to provide a perspective from dietary diversity in quality. The resultant commodities for which quality is evaluated are fat, meats-proteins, cereals, vegetables, milk, starches, complements[9], tubers, fruits, fish and chicken - steps that are further described in Section 3.1. The direct use of market-observed prices also has the advantage that it avoids using unit-values (price-deflator) as price controls in our analysis. More specifically, the market-prices - used as controls in the analyses - avoid the measurement errors that arise with the use of unit-values (which approximate the price of an item with the average of item expenditure divided by its quantity consumed).

It is worth highlighting inferring qualities for commodity-groups - based on an arbitrary or intuitive grouping criteria - may not follow separability in demand or other restrictions of demand analysis either (Deaton & Muellbauer, 1980b). As Nelson (1991) points out, an approach to consider a simple sum of quantities consumed from across items (or goods) within a commodity group (Deaton, 1988) may only make sense if one imposes severe restrictions on consumer preferences. The quality measure based on Hick's commodity theorem that relies on market-observed prices remedies such issues without imposing restricting assumptions on consumer demand (Hicks, 1946; Nelson, 1991). The measure also avoids intuitive grouping of goods into commodities - thus avoiding the use researcher's own judgment of substitution within commodity-groups.

Other than using the stable relative price-structure associated with a commodity (fat, cereals etc.), the quality metric for every household depends on the measurement of quantities of high vs low-price variants (goods) consumed within a food commodity (e.g., the quantities consumed of wheat, rice, maize for quality in cereals commodity). Considering a commodity-group of cereals comprising of rice, maize and wheat, for example, maize would contribute to lower quality of cereals-commodity than wheat and rice if the prices (per-kg) of wheat and rice are higher than that of maize. Other than the quality metric thus calculated for every food commodity (such as cereals), the budget shares per household for each of these commodities are used as dependent variables in the system of equations used for an AIDS estimation. In order to provide a comparison with a disaggregated view of goods consumed (Heien & Wessells, 1990) without any grouping into commodities, we also provide the results from estimation of price elasticities for a disaggregated set of goods consumed (see Appendix 7.2).

The econometric method uses an AIDS formulation with both quality metrics and budget shares as dependent variables (Deaton & Muellbauer, 1980a). This is the "Unrestricted

[9]Complements include miscellaneous items such as tea, sugar, coffee, spices and salt - that are not considered in other categories.

Method" AIDS formulation used by (McKelvey, 2011; Andalon & Gibson, 2017) - with the total household expenditure - interpreted as permanent income - as the main explanatory variable in the regression. The base-prices - used as controls - are market prices directly observed in the survey - rather than unit-prices which the empirical quality assessments are often compelled to rely upon (despite issues with measurement errors (Gibson & Kim, 2019)). The interpretation of household's total expenditure as the permanent income measure in the study - is common in the literature and relies on the standard consumption theory (see (Modigliani & Brumberg, 1954; Friedman, 1957)). To highlight the effect of wealth, we also present results from an alternative formulation where the logarithm of reported value of total assets owned by the households is used as the main explanatory variable. The use of owned assets (land, house, furniture, vehicles, electronics etc.) in wealth indices is also common in the health and consumption studies (Howe, Hargreaves, & Huttly, 2008; Howe, Hargreaves, Ploubidis, De Stavola, & Huttly, 2010; McKenzie, 2005; Booysen, Van Der Berg, Burger, Von Maltitz, & Du Rand, 2008).

The results indicate that higher permanent income (measured as total expenditure) is associated with higher demand for meat-related proteins and fruits in particular. Further, the effect is made stronger due to regional inequalities relating to variations in urbanisation levels in Tanzania. More specifically, electricity being sparsely available in the urban areas of Tanzania suggests strong effects of both electricity availability and urbanisation levels on price-based quality (meats and fruits in particular) [10]- despite controlling for permanent income (total expenditure). This seems aligned with other studies on effects of resource availability on food diversity in the region (see (Hailu & Woldemichael, 2019) who explore the role of water availability on food diversity).

The presented chapter contributes to the literature in a few different ways. First, it demonstrates that permanent income effects on quality in a typical developing economy may not be uniform across food categories nor across the income categories of consumers. This has implications for how such consumption may rise with more economic growth and income differences. Second, it investigates income effects on higher quality using a more thorough as well as robust quality measure than the more commonly used measures - such as the overall diversity measures or intuitive item classifications. Third, it provides and employs a feasible method to include food prices as controls in the discussion of regional variation in quality without relying on detailed market prices of every item.

In the sections that follow, the Section 2 surveys the literature on quality of consumption, food diversity and the relevance of often overlooked market prices in measurement of quality. The approach to view quality across various commodities are presented in Section 3. The details of the data used in the study are provided in Section 4 and a discussion of the distribution of quality for food follows in Section 5.

[10]Since the prices of all items in the commodity move together, only the base-price i.e. the minimum price in the commodity is used as a control. This is detailed further Section 6.

2 Literature Survey

While consumption of basic needs (such as food) may have hardly any status value in societies where the basic consumption needs have been fulfilled (Hirsch, 1977), food continues to carry a status value in societies where it is either scarce or signifies differences in life-styles (I. L. Ohna, 2007). The differences among agricultural, pastoral and urban life-styles in Tanzania that are reflected in food habits present us an interesting case-study where differences in food habits are indicative of one's social standing that are in turn influenced by occupational status or regional belonging.

As is often found in agrarian societies, the differences in food quality as well as in the social values associated with the several variants can be significant (Onuorah & Ayo, 2003; I. L. Ohna, 2007). It is common in Tanzania, for example, that the guests are served higher quality variants of commodities consumed by household. The variation in social value can be significant even for variants that are of similar nutritional value. In Tanzania, it is observed that while the staple diet (*ugali*) for many communities covers a wide range spanning maize, *kiwerege*, *dona*, rice, sorghum, cassava, millet and pumpkin, *kiwerege* and rice invariably carry more social value than millets or pumpkin (I. L. Ohna, 2007). It is worth remarking that the different types of flour - varied only by milling quality - tend to carry different social values. For example, maize can be of *kiwerege*, *sembe* and *dona* types - out of which *dona* is associated with poor communities and *kiwerege* carries the highest social value. This variation seems even wider for non-staple food (*mboga*) which can be used either as a side-dish or as a main meal (the latter being more often for the pastoralist communities). While *mboga* comprises of beef, mutton, *kuku* and other meats as well as beans, cabbage, fermented milk and leafy-greens, meat carries more social value than beans or cabbage. Further, meat and animal products are also used as special foods during ceremonies and to treat the sick (I. Ohna et al., 2012).The use of tomatoes, onion or red pepper invariably increase the social status associated with the food served (I. Ohna et al., 2012). The items of higher social value are invariably priced higher in the consumer markets.

The effects of recent economic changes on food habits also seem evident for communities in Tanzania - where an association with agriculture seems to valued generally higher. For example, *dona* is often valued more among the pastoralist community than among agriculturalists[11]. Urbanisation also seems to have brought particularly significant changes in food habits as more food items have been introduced to be consumed with between meals (with *chai* or as *kumbwa* (I. L. Ohna, 2007; I. Ohna et al., 2012)) - while a shift away consumption of wild foods and meats is also under way (I. L. Ohna, 2007). Given the status value of different varieties of food and the focus on food in the LSMS data, a study of variation in food quality is crucial to our exploration of high quality consumption for basic needs in the context of aspirational consumption in Tanzania.

[11] Some historical factors have also contributed to this variation such as the past government distribution of sorghum leading it to be regarded as of even lower social value than *dona* among the agriculturalists.

The study of variation in food quality has had a long tradition in empirical economics. One of the earliest approaches for measurement of quality had been to treat it as a separate good (Houthakker, 1952) - thus decomposing the quality price elasticity of demand (quantity purchased) into a quantity price elasticity of expenditure on quality and an income elasticity of quality measure. The treatment of quality as a separate good has been widely used ever since Deaton and Muellbauer (1980b). A common example of the approach is a simple repackaging method (Fisher & Shell, 1968) - which interprets demand for a high-quality good as multiples of lower quality goods. More specifically, the demand for goods with a certain price using the repackaging approach is viewed in terms of a base-price and a quality parameter - the implicit assumption being that the variants in every category are perfect substitutes.

For most empirical studies, however, the perfect substitution turns out to be a stringent requirement since the variation in quality among goods within a commodity may be often extreme. As an example, even though quality consumed of electricity under energy commodity may be higher than that of kerosene, the substitution between electricity and kerosene may in fact be little - thus making it difficult to justify a view of a demand for electricity in terms of the units of kerosene consumed[12]. Nevertheless, simple-repackaging does turn out to be appropriate in the context of cost-of-living index problems where a substitution is implied (Fisher & Shell, 1968).

Some constraints on either consumer preferences or prices are needed - when perfect substitution isn't possible - so as to measure the quality using both quantities consumed and prices of goods within a commodity. As such, the problem of measuring quality of a commodity comprising of multiple goods is that of assigning a quality metric to the entire commodity group using the a combination of quantities consumed of the goods within the commodity group at prices known for every good. As Nelson (1991) explains, the approach to measurement of food quality in such a manner must either assume weakly separable preferences with homothetic within-group preferences (Cox & Wohlgenant, 1986; Deaton, 1988) or a stable price-structure in a commodity group with a base-price associated with the commodity. The current chapter opts for the latter approach since it not only imposes fewer restrictions on the consumer preferences but also makes use of the market prices rather than the unit-prices[13]. It is worth emphasising that the homotheticity of within-group preferences presents an implausible constraint on the demand - since it implies either that the within-group income expansion path would be a straight line through the origin or that the group composition would remain independent of income - conditions. Such a restriction on preferences would mean - for example - that the ratio of electricity consumed to the kerosene consumed must be the same for both rich and poor consumers - a constraint that is unlikely to hold empirically.

[12]While one may address this issue with use of an energy-budget (calories), the unavailability of quantities consumed or prices in the consumption data limit wider applications of such an approach.

[13]The unit-prices are the total expenditure divided by quantities consumed that are used prices as controls in a regression with budget, quantity or quality as dependent variables.

More formally, consider that in a generic sense, a quality-metric associated with a certain commodity (e.g., cereals, fat etc.) depends on a grouping of items - which the commodity comprises of - subject to certain homogeneity restrictions. Given the quantity consumed $\{q_i\}_{i\in[1,n]} = \{q_1, q_2, ..., q_n\}$ and respective prices $\{p_i\}_{i\in[1,n]} = \{p_1, p_2, ..., p_n\}$ food $1 \leq i \leq n$ goods, the quality-metric solves a problem equivalent to the following disaggregated optimisation problem (Nelson, 1991) for a direct utility $U(q_1, q_2, ...)$

$$\begin{aligned} &\max U(q_1, q_2, ...q_n) \\ &\sum_{i=1}^{n} p_i q_i = x \end{aligned} \tag{9}$$

The quality metric problem is that of grouping the items $1 \leq i \leq n$ into groups $1 \leq G \leq M$ and assigning group-commodity price P_G so that the solution of the problem in Equation 9 is equivalent to that of the problem in Equation 10

$$\begin{aligned} &\max U(Q_1, Q_2, ...Q_M) \\ &\sum_{G=1}^{M} P_G Q_G = x \end{aligned} \tag{10}$$

Such an aggregation requires that the M groups (commodities) are separable and that the demand for items within each group is a homogenous function of the goods it contains (Nelson, 1991). The approach followed by Deaton (1988) - using the sum of physical quantities as measure of demand - requires that the homogeneity restriction is put on the household preferences while assuming that the within-group preferences are homothetic. However, the approach from Deaton (1988) (as well as that from Cox and Wohlgenant (1986)) does not use the homogenous commodities assumption and in fact permits heterogeneity in its use of sums of physical quantities as the measure of demand (Nelson, 1991). The Hick's commodity theorem-based approach - which imposes a restriction only on prices of goods - does not suffer with the issue as its use in the current chapter does not impose the stringent homotheticity assumption on preferences.

Simply put, the approach we adopt in the current chapter prevents considering chicken and eggs into one commodity unless the relative price structure for the two items (chicken and eggs) does not change. No other criteria - arbitrary or otherwise - is followed to group the goods into commodities. Subject to the restriction, the notion of quality is set to a price-weighted quantity (made possible due the availability of market prices) so that given the same expenditure, a lower quantity of the composite good (i.e. Q_G for the composite good G) corresponds to a higher quality and vice versa (see Section 3 for further details). The commodity groups thus selected are obtained by clustering goods into commodity

groups based on similar price-changes over the years 2008, 2010, 2012 and 2014. A further step sub-divides the clusters thus obtained based on the FAO provided categories to assist interpretations from a DD perspective[14].

The issue of unavailability of market prices - often sidelined in many empirical studies - has recently gathered some attention (McKelvey, 2011; Gibson & Kim, 2019). As such the measurement of quality is often performed without available market prices - by simply averaging over the expenditures (recorded costs) of a particular item in a given region (or over the entire population of households). This so-called unit-value (or price-deflator) method uses the average expenditure divided by quantity consumer for all the consumers as prices in the empirical analyses (see Lazaridis (2003); Tafere, Taffesse, Tamiru, Tefera, and Paulos (2010) for recent implementations). While using unit-values as prices may be appropriate for certain empirical concerns, the measurement errors introduced by their can be severe and difficult to ignore for finer conclusions relating to price and income elasticities. More specifically, the quantity consumed appearing both in the dependent variable (as budget share's numerator) and in the control variable (as the denominator for unit-value used as price) cannot be always mitigated (McKelvey, 2011; Gibson & Kim, 2019).

The approaches such as the RDMP ("real" deviations from regional/quarterly mean prices) proposed by Cox and Wohlgenant (1986) that are implemented without using locally observed market prices also suffer with the same issue[15]. The availability of detailed market-observed prices in the survey - therefore not only helps with the fulfillment of Hick's Commodity theorem through stable relative structure of prices - but also mitigates the measurement errors that would be implied with the use of unit-values.

The econometric for the price-based quality metric uses the "Unrestricted Method" for the AIDS model (Deaton & Muellbauer, 1980a; Andalon & Gibson, 2017; Gibson & Kim, 2019). Based on PIGL approaches originally popularised by Barten (1964), Gorman (1953) and Muellbauer (1975), the AIDS model readily conforms to the demand function restrictions and can be extended for other panel data scenarios. For instance, the structural changes in the parameters over time are addressed with a dynamic AIDS formulation (see G. Anderson and Blundell (1982, 1983) for a first-order dynamic model) and the demographic variables are accommodated using a "demographic translation" approach (Pollak & Wales, 1981) . With prices p_j for commodities $1 \leq j \leq n$, total expenditure x, coefficients γ_{ij} , β_i for commodities $1 \leq i \leq n$ and the price-index P, the demographic translation for the linear AIDS formulation uses the following α_i intercept in AIDS regression equation

[14]The FAO provided categories are (i) cereals, (ii) vegetables, (iii) fruits, (iv) meat, (v) eggs (vi) fish and other sea foods (vii) legumes, nuts and seeds, (viii) milk and milk products (ix) oil and fats, (x) sweets (xi) spices, condiments and beverages and (xii) tubers and roots.

[15]For situations where market-prices are not available to such detail, there have been a few techniques to address the measurement errors arising from the use of price-deflators. Deaton (1988) - for instance - uses cluster-average fitted budget shares and unit-values in a cross-sectional analysis - an approach that has been elaborated by McKelvey (2011) for panel data analyses.

$$\alpha_i = \rho_{io} + \sum_{k=1}^{s} \rho_{ik} d_k \tag{11}$$

Specifically, the above intercept α_i for $1 \leq i \leq n$ goods , d_k are demographic variables and constants ρ_{io} , ρ_{ik} are used in the following AIDS regression where w_i is the budget-share for the item i

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} ln(p_j) + \beta_i ln(\frac{x}{P}) \tag{12}$$

Compared with the measurement of elasticities for every individual good, the advantage with the aggregation over commodity groups is that the latter does not present the issue of non-consumption of high-quality goods. More specifically, if one were to calculate the price elasticities for individual goods, an obvious issue that arises is that certain items (particularly those with high-quality) might not be consumed at all for a significant part of the population. Heien and Wessells (1990) address this issue by correcting for non-consumption using the inverse Mills ratio as instrument into the AIDS regressions. The Heien-Wells (HW) approach effectively treats the consumption on high quality goods "truncated" for a significant part of the population[16]. While HW approach may be well-suited for comparison of elasticities for specific goods (see Appendix 7.2 for empirical results using the approach), the comparison of quality across multiple commodities is a more appropriate interpretation for the overview of quality we seek in the current chapter. Since it is much less likely for no goods to be consumed within a broader commodity group (fruits-veg, protein etc.) over which the quality measure is calculated, the approach with quality aggregated over broader commodity groups also circumvents the issue of non-consumption. No correction related to non-consumption - as used in HW method - is therefore used with our approach that aggregates quality over groups classified with homogeneity restrictions (see Section 3.1).

The variation in quality across multiple food sub-categories is examined using the estimated income elasticities for quality using total household expenditure as the main explanatory variable and urban-amenities as key controls. While primarily exploring the role of food quality in status, the current chapter also touches upon issues related with SES effects on food diversity that have been remarked by many empirical studies (Ruel, 2003; Hatløy et al., 2000). In particular, a lower within-country SES measure - whether it be in developed or developing economies - often corresponds to less healthy diets (Manyanga et al., 2017). In a broad survey across low and middle-income economies, Mayen, Marques-Vidal, Paccaud, Bovet, and Stringhini (2014) also find that with the exception of

[16]Heien and Wessells (1990) also use the unit-values instead of (unavailable) market prices. Further, the unit-values obtained from the households using a good are used as prices for those who don't consume the good.

fruits in diet, the residents of urban areas - particularly with higher SES - have more calorie intakes as well as consumption of protein, fats and vitamins relative to starchy items.

Some empirical studies also find a relevance of the availability of amenities such as clean water on dietary diversity (Mekuria, Wubneh, & Tewabe, 2017). A control for amenities such as electricity are also used - other than total expenditure as our main explanatory variable - in the current chapter to examine the variation in food quality across Tanzania. In its particular context, rapid urbanisation seems to have brought challenges to food security for the lower-income households in the towns and cities within Tanzania - even though the country has managed to maintain a broad self-sufficiency in basic food items (Wenban-Smith et al., 2016). The agricultural policy has also focused on a transition from subsistence farming to more efficient large-scale rural farming since urban farming seems common in the country (Wenban-Smith et al., 2016). As such, a significant number of households dependent on self-farming in reportedly rural settings may obfuscate the relationship between urban and rural food demand that may be apparent from the conventional models for urban-rural migration (Harris & Todaro, 1970). The control for urban-rural disparities is thus an important one for our survey of food quality in the Tanzanian economy.

3 Econometric Model

The measure of quality used in the current chapter for each food commodity group (such as fat, cereals etc.) comprising of multiple items (maize, rice, wheat etc. for cereals etc.) relies on higher weights to high-priced items (rice given higher weight than maize based solely on price etc.) within each commodity group. The commodity groups over which the quality is calculated are based on the price-structure restrictions implied by the Hicks Commodity Theorem (more details to follow in Section 3.1). A further sub-classification - which does not disrupt the commodities identified based on similarity of price structure - is used to assist interpretation from a DD perspective and is guided by FAO categories recommended for food diversity. It is worth highlighting that instead of measuring an overall food quality across all of the food items consumed, we inspect quality only for different commodity groups whose constituents satisfy the Hick's commodity theorem. The quality across food can thus be seen as a vector associated with the set of food commodities. As we detail in the Section 3.1, the quality vector $\{V_{fat},\ V_{meats-proteins},\ V_{cereals}\ ,\ V_{vegetables}$,$V_{milk}$,$V_{starches},\ V_{complements}$,$V_{tubers}$,$V_{fruits}$,$V_{fish}\}$ is based on price-correlation requirement from the Hicks Commodity Theorem - while the groups based on movements are further subdivided based on FAO categories to assist interpretations from a DD perspective.

The data used for the classification as well as the inputs towards estimation are described in Section 4.

3.1 Classification in commodity groups

The main criterion of grouping of food items into commodity groups is the co-movement of prices. The prices - shown in Figure 8 - are taken from the wave corresponding to years 2008, 2010, 2012 and 2014 in the survey. An unsupervised k-means clustering is then used to group the items into commodities based on price changes - so that the items within every commodity group identified have "similar" relative price-changes. The groups based strictly on price-movements thus obtained are then split into further categories aligned with the FAO recommended categories i.e. (i) cereals, (ii) vegetables, (iii) fruits, (iv) meat, (v) eggs (vi) fish and other sea foods (vii) legumes, nuts and seeds, (viii) milk and milk products (ix) oil and fats, (x) sweets (xi) spices, condiments and beverages and (xii) tubers and roots. More specifically, the k-means clustering is performed on the relative price-changes from available prices. Considering beef as an example good, the prices for beef in the years 2008, 2010, 2012, 2014 (see Figure 8) are $p_{beef} = \{3500.0, 4000.0, 5000.0, 6000.0\}$ - which correspond to a relative price-change vector $r_{beef} = \{\frac{4000}{3500}, \frac{5000}{4000}, \frac{6000}{5000}\} = \{1.14, 1.25, 1.2\}$. The relative price-change vectors for all the items (beef, goat, fresh_milk etc.) thus obtained are clustered using the k-means method - where the number of selected clusters is guided by the within-cluster sum of squared errors - an approach commonly used in k-means applications (Pollard, 1982). The classification we use as the basis for commodity groups are arrived at by setting the number of clusters to 5 which is when the within-cluster sum of squared errors drops to less than unity[17]. The classification of items based on these FAO-categories and the correlated-price-movement is presented in Table 5.

It is worth emphasising that the partition of commodity groups identified from the unsupervised k-means clustering into further sub-groups based on FAO categories does not pick goods from disparate clusters into a new sub-group. Instead, the further partition into sub-groups based on FAO categories is based on the observation that that the price-changes of goods within sub-groups formed by splitting a particular group identified by clustering would also be correlated with each other. The groups obtained by the classification based on correlation prices (shown in Table 6) are thus split into further sub-groups aligned with the FAO groups in Table 5. More specifically, the commodity cereals-milk-greens is split into milk, veg and cereals, fruits-fish is split into fruits and fish, startches-fats-complements is split into starches, fats and complements while the categories tubers and meatsproteins are retained. The resulting categories are therefore i) cereals ii) veg iii) fruits iv) meats-proteins v) fish vi) milk vii)fat viii) complements ix) tubers x) chicken and xi) starches whose constituents are shown in Table 5.

As explained before, the basis for the quality metric associated for every commodity G (corresponding to the left side of the Table 5) is the Hick's commodity theorem (Hicks, 1946) which states that a group of goods behave as if they were a single commodity so long as the prices of the group of goods changes in the same proportion. Simply put, a constant relative

[17] A higher number of clusters does not significantly reduce the within-cluster sum of squared error.

food group	food items
fat	cooking_oil, eggs
meats-proteins	beef, goat, pork, pulses, onion, salt, canned_milk, wild_birds, wild_meat
cereals	rice_husked, maize_green, maize_grain , wheat, maize_flour
vegetables	greens, othervegstarch
milk	fresh_milk, milk_products
starches	sugarcane, millet_grain, rice_paddy, potatoes, canned_drink, sweet_potato, bread
chicken	chicken
complements	sugar, tea, coffee, sweet, miscdrinkpowder, readymade_tea_coffee, honey, spices
tubers	yam, millet_flour, banana_green, cassava_flour
fruits	banana_ripe, cassava_fresh, peanuts, citrus, coconut, cashew_almonds,dried_canned_veg, nut_products, mangoes
fish	dried_canned_fish, fish_seafood, packaged_fish

Table 5: Food items classified into commodity groups after considering both price-movements and FAO categories

food group	food items
meats-proteins	beef, goat, pork, pulses, onion, salt canned_milk, wild_birds, wild_meat
cereals-milk-greens	rice_husked, greens , othervegstarch, maize_green, maize_grain , wheat, maize_flour , fresh_milk, milk_products
starches-fats-complements	cooking_oil, eggs, sugar, tea, sweet, sugarcane, coffee, chicken, bread, potatoes, sweet_potato, canned_drink, miscdrinkpowder, readymade_tea_coffee, honey, spices, millet_grain, rice_paddy
tubers	cassava_flour, yam, millet_flour, banana_green,
fruits-fish	banana_ripe, cassava_fresh, peanuts, citrus, coconut, mangoes, dried_canned_veg, dried_canned_fish, cashew_almonds, nut_products, fish_seafood

Table 6: Food items classified into commodity groups based on price-movements (without splitting into FAO-based categories)

price structure allows us to factor out the quantity of the composite commodity consumed from the demand function for the variants (goods) within the commodity - allowing 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^* from n goods with prices $p_G = \{p_{i,G}\}_{i\in[1,n]}$. With E_G as the total expenditure on group G, the constant relative price-structure permits us to write the commodity-aggregate specific demand function $g(E_G, p_G)$ in terms of the price-structure p_G^* as follows

$$g(E_G, p_G) \equiv g(\frac{E_G}{P_G}, p_G^*) = g(Q_G, p_G^*) \tag{13}$$

The formulation in Equation 13 depends both on the quantity $Q_G \equiv \frac{E_G}{P_G}$ and a quality variable V_G[18]. The essential appeal of aggregation is the simplified view of demand across all items. Consider for example, if we were to represent demands for items {wheat, rice, yam,coconut, eggs, beef, chicken, fish, goat, honey}. As such these items would each correspond to a demand function so that 10 functions describe the demand for all the 10 items. The aggregated view allows us to view demands for commodities {carb, protein and complements} - where the 10 goods are encompassed in just 3 respective functions for the 3 commodity groups. A key requirement for the validity of such a quality metric - however - is that the prices of goods within the commodity group have a constant relative price-structure i.e. a relative prices ratio that can be factored out of the observed prices (see (Deaton, 1988; Nelson, 1991; Gibson & Kim, 2019)). This condition - due to the Hicks commodity theorem - drives the k-means grouping of items described above.

In summary, the commodity aggregates used in the study are based on a *price-movement* sense (based on the Hick's commodity theorem) - while a further splitting of groups is guided by a *functional* sense implied by FAO categories. Every commodity aggregate on the left side of Table 5 has goods with a similar price-movement *and* the same functional significance while every commodity aggregate G on the left side of Table 6 has items that have only a similar price-movement[19] . To explain with particular example goods, consider two pairs of items with similar price-movements - with sugar and tea being the first pair that have similar price changes (see Table 6). Let the second pair be rice and cooking-oil - who also have price changes similar to each other (see Table 6). Notice however that not only do sugar and tea have similar price changes, they also have a similar functional sense implied by FAO categories. Therefore, they are classified in a group *complements* in the Table 5. Rice and cooking-oil on the other hand are split into two categories based on the different functional sense implied by FAO-categories (see Table 5). Other categories such as meatsproteins, fruits etc. are similarly based foremost on price-movement correlations before considering the functional

[18]The income elasticity of composite quantity Q_G is the sum of income elasticity of quality V_G and the income elasticity of physical quantity q_G. This is further explained in Section 3.2 (see (Nelson, 1991) for more details).

[19]A similar analysis of rice varieties is conducted by (McKelvey, 2011) .

sense implied by FAO categories for interpretation from a DD perspective.

3.2 Price-based quality

The vector of quality for every household comprises of the quality $\nu_{G\in\mathscr{G}}$ associated with the commodity group $G \in \mathscr{G}$ where $\mathscr{G}$ is the set of all commodities relevant for the household $\mathscr{G}$=$\{cereals, veg, fruits , meats\_proteins, fish , milk, fat, complements, tubers, starches\}$. Based on the local market-observed prices of the goods within the commodity group $G \in \mathscr{G}$ (such as cereals, fat etc.), the quality metric for the household ensures that a household who buys more of the high-priced good within a commodity group is viewed as one consuming higher quality of the commodity G. This is why no quality is considered for chicken, the only good within its own commodity group. As we've explained in Section 3.1, the k-means clustering interprets the restriction that there is a constant relative price structure on the constituent items of commodity group (the Hicks aggregation condition).

To calculate the quality associated with a group $G \in \mathscr{G}$ (such as the items : beet and goat in the commodity: meatsproteins) that have a constant within-group relative price-structure, a base-price P_G for the group G (meatsproteins) is selected as the minimum price of the items among the commodity aggregate constituents (e.g., among beef and goat). This is based on the approach suggested by (Nelson, 1991). The base-price P_G is then used to assign i) an aggregate-quantity of a composite good (for meatsproteins) comprising of quantities consumed of the items constituted in G (beef, goat) and ii) a quality metric corresponding to G (based on the proportion of goat and beef consumed). In other words, with the quantities consumed for beef and goat under meatsproteins commodity available to us, a quantity of "meatsproteins" is associated with the particular set of quantities consumed for beef and goat. The aggregate-quantity of "meatsproteins" is then calculated by relative-prices-weighed averages of beef and goat consumed (see details below). If beef per unit-weight is cheaper than goat-per-unit-weight, then this quality metric would evidently be lower when more beef is consumed for the same expenditure. This notion of higher-quality is thus based solely on price alone so that a more expensive item within the commodity group contribute to higher quality of the commodity.

In more formal terms, quantity-price tuples $(q_1, p_1), (q_2, p_2), (q_3, p_3)...$ within a commodity thus correspond to a quantity and quality pair (Q_G, P_G) as well as a quality metric V_G. We define the total physical sum of quantities consumed q_G as follows

$$q_G = \sum_{j\in G} q_{j,G}$$

Further, with $p_{i,G}$ as the prices for good i in commodity group G observed in the market, we represent the price-structure ratios for the good i in a group G as $p^*_{i,G}$. The ratio $p^*_{i,G}$ is

effectively a price-multiple with respect to the base-price P_G (the price of the base-commodity in the group G with respect to which the quality in G is being considered) so that the observed prices $p_{i,G}$ of the items within the group G bear the following relationship:

$$p_{i,G} = (P_G \times p^*_{i,G}) \forall i \in G \tag{14}$$

The quality metric V_G then corresponds to the a sum of quantities consumed $q_{i,G}$ weighted by the price-structure $(p^*_{i,G})$ for the commodity aggregate G as follows

$$V_G = \sum_{i \in G} (\frac{q_{i,G} p^*_{i,G}}{q_G}) \tag{15}$$

Following (Nelson, 1991), we reject the use of sum of physical quantities $q_G = \sum_{i \in G} q_i$ itself as Q_G (as used by (Deaton, 1988) - see Equations 10 and 13) in favour of the above relative-price-weighted sum of expenditure per physical unit (within the commodity G)[20].

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, 1991) for a more detailed theoretical discussion). Let's say we observe the consumption on goat and beef within a meatsproteins commodity with prices observed as 20£, 10£ (resp.) and the quantities consumed by the household 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 "meatsproteins". However, it is more difficult to argue that the households optimise the weight of the composite good - rather than comparing the disparate quantities of goods (see (Nelson, 1991) for details of this argument). Since quality is to be priced in the market, a quality-adjusted quantity is the more appropriate formulation where we consider a price-structured weighted quantity as the quality-adjusted quantity and express the pair i.e. $\{(p_{goat}, q_{goat}), (p_{beef}, q_{beef})\} = \{(20, 1.5), (10, 2)\}$ as $(10, 5 (= 2 \times 1.5 + 1 \times 2))$, $(0, 1.43 = \frac{5}{1.5+2})$ or $\{(P_{meats}, Q_{meats}), (0, V_{meats})\} = \{(10, 5), (0, 1.43)\}$. Here, P_G (or P_{meats}) is simply the minimum of $20, 10$ i.e. the prices of goat and beef respectively - corresponding to the price-structure p^*_G is (2,1) for {goat, beef}. In other words, consuming 1.5 kg of goat ($q_{goat} = 1.5$) is equivalent to consuming 3 units of commodity: meats-proteins $Q_{meats} = 2 \times 1.5$) with quality of 2 ($Q_{meats} = \frac{2 \times 1.5}{1.5} = 2$) since the price of goat is twice that of base good: beef.

[20](Nelson, 1991) argues that the composite good quantity Q_G can be equivalent to the sum of physical quantities q_G (i.e. $Q_G = q_G$) only if the quality effects are absent. With $U_G = \frac{E_G}{q_G}$ as the unit quantity consumed this is implied due to $Q_G = q_G \Rightarrow \frac{E_G}{P_G} = \frac{E_G}{U_G} \Rightarrow P_G = U_G$.

Good	*Prices*	*Quantity*		P_{meats}	Q_{meats} and V_{meats}
Goat	p_{goat} =20£	1.5	$\rightarrow$	10	$Q_{meats} = 5$ $(= 2 \cdot 1.5 + 1 \cdot 2)$
Beef	p_{beef} =10£	2.0			V_{meats} =1.43 $(= \frac{5}{1.5+2})$

Table 7: Quality Measurement in Illus. 1 (prices not representative of observed prices)

Comparing the $V_{meats} = 1.43$ above with another case where household consumes 0.5 kg of goat and 3 kg of beef so that the prices and quantities are $\{(20, 0.5), (10, 3)\}$ respectively for goat and beef, we have the same expenditure as before since $E_G = P_G Q_G = \sum_G p_i q_i = 40$. However, as shown in the table below, the quantity Q_G and quality V_G now corresponds to $(10, 4(= 2 \times 0.5 + 1 \times 3))$ and $(0, 1.14 = \frac{4}{3.5})$ respectively. The quality is lower for the latter household even though the sum of physical quantities - rather than quality-adjusted Q_G - is the same.

Good	*Prices*	*Quantity*		P_{meats}	Q_{meats} and V_{meats}
Goat	p_{goat} =20£	0.5	$\rightarrow$	10	$Q_{meats} = 4$ $(= 2 \cdot 0.5 + 1 \cdot 3)$
Beef	p_{beef}=10£	3.0			V_{meats} =1.14 $(= \frac{4}{3.5})$

Table 8: Quality Measurement in Illus. 2 (prices not representative of observed prices)

3.3 Econometric Method

The goal of the comparison of quality metrics across commodity groups defined in Section 3.2 is to measure how the demand for the diversity-based quality is distributed across the population and varies across prices, household needs, wealth and other household characteristics in the cross-section.

In the "Unrestricted method" of AIDS system of equations (see (McKelvey, 2011) and (Andalon & Gibson, 2017; Gibson & Kim, 2019)), each composite commodity group - unless it consists of just one good/item - corresponds to both a budget equation and a quality metric dependent variable. The income and price elasticities are calculated with the total-expenditure x and base-price P_G using the AIDS implementation for food categories defined in Table 5. As described earlier, the base-prices correspond to every category in the classification defined in Table 5 (based on prices shown in Figure 8). We use two formulations with the total household expenditure logarithm used as the main explanatory in the first formulation and the logarithm of the household assets value as the explanatory variable in the second

formulation. An indicator variable is also used as control for energy - specifying whether the household uses electricity or not.

With each of the two explanatory variables, we use both the quality-metric $ln\ V_G$ and the budget share w_G as dependent variables (for a food commodity G) in the AIDS system of equations. Thus with intercepts $\rho_{1,G}$, $\rho_{2,G}$ for budget-share and quality dependent variables equations (respectively) for given commodity group G among n commodity groups and coefficients $\rho_{1,Gk}, \rho_{2,Gk}$ for s demographic characteristics $\{d_k\}_{k\in[1,s]}$, $\beta_{G,1}, \beta_{G,2}$, $ln(x)$ as logarithm of total expenditure and coefficient $\gamma_{1,Gj}$, $\gamma_{2,Gj}$ for the base-prices corresponding to n commodity groups $\{P_j\}_{j\in[1,n]}$, we have the following system of equations (for each commodity group G)

$$w_G = \rho_{1,G} + \sum_{k=1}^{s} \rho_{1,Gk} d_k + \sum_{j=1}^{n} \gamma_{1,Gj} ln(P_j) + \beta_{G,1} ln(x) + \epsilon_{G,1} \tag{16}$$

$$ln\ V_G = \rho_{2,G} + \sum_{k=1}^{s} \rho_{2,Gk} d_k + \sum_{j=1}^{n} \gamma_{2,Gj} ln(P_j) + \beta_{G,2} ln(x) + \epsilon_{G,2} \tag{17}$$

Equivalently, with $ln(A)$ as the explanatory variable, we have the following system of equations

$$w_G = \rho'_{1,G} + \sum_{k=1}^{s} \rho'_{1,Gk} d_k + \sum_{j=1}^{n} \gamma'_{1,Gj} ln(P_j) + \beta'_{G,1} ln(A) + \epsilon'_{G,1} \tag{18}$$

$$ln\ V_G = \rho'_{2,G} + \sum_{k=1}^{s} \rho'_{2,Gk} d_k + \sum_{j=1}^{n} \gamma'_{2,Gj} ln(P_j) + \beta'_{G,2} ln(A) + \epsilon'_{G,2} \tag{19}$$

Once again, the intercepts $\rho'_{1,G}$, $\rho'_{2,G}$ corresponding to a commodity group G (of which there are n in number) are for budget-share w_G and quality metric $ln\ V_G$ dependent variable equations (respectively) while the coefficients $\rho'_{1,Gk}, \rho'_{2,Gk}$ are defined for s demographic characteristics :$\{d_k\}_{k\in[1,s]}$. Similarly, $\beta'_{G,1}, \beta'_{G,2}$ are coefficients of $ln(A)$ - the logarithm of total assets value and the coefficients $\gamma'_{1,Gj}$, $\gamma'_{2,Gj}$ are for the base-prices corresponding to n commodity groups $\{P_j\}_{j\in[1,n]}$. Notice that the s demographic variables in Equations 16 and 17 include the indicator variables related to access to electricity as well. Further, since the price for quality is not required, the price-coefficients in the $ln\ V$ equations are not subject to any restrictions - unlike the budget-share estimation equations where the constraints of symmetry are applied. More specifically, the symmetry conditions are imposed with $\gamma_{1,Gj} = \gamma_{1,jG}$ for all j, G in the equations for budget-share equations (and equivalently

$\gamma'_{1,Gj} = \gamma'_{1,jG}$ for all j, G).

The homogeneity and adding-up restrictions mean that the sum of all expenditure shares in the AIDS model is unity - thus leading to the residuals variance-covariance matrix becoming singular (Barten, 1969; G. Anderson & Blundell, 1982; Fujii, Khaled, & Mak, 1985; Asche, 1996; Li, Song, & Witt, 2004). As a solution to this well-known estimation problem, it is common to delete one of the equations to estimate and recover its coefficients through the adding-up restrictions (Paris & Caracciolo, 2014; Li et al., 2004). Notice that the adding-up restrictions require that $\sum_G^n \rho_{1,G} = 1$, $\sum_G^n \beta_{G,1} = 0$ and $\sum_G^n \gamma_{1,Gj} = 1$ (for all j) while the homogeneity restrictions require that and $\sum_j^n \gamma_{1,Gj} = 0$ (for all G). Thus one of the commodity group (chicken) - which does not correspond to a quality is ignored from the estimation. The coefficients in the budget-share equation for chicken can be recovered as $\rho_{chicken} = 1 - \sum_{G \neq chicken}^n \rho_{1,G}$ and $\gamma_{1,chicken,j} = -\sum_{G \neq chicken}^n \gamma_{1,Gj}$ respectively.

In order to describe an exact equation set to be estimated, we consider the commodity $G = meatsproteins$ along with the demographic variables d_h (number of household members), d_τ (age of the household-head), d_χ (urban residence) and d_ξ (electricity-access). There are two equations for budget-share and quality of the meats-proteins commodity (which we name `qmeatsproteins` and `qVmeatsproteins` respectively - for reference purposes) :

$qmeatsproteins$:

$$
\begin{aligned}
w_{meatsproteins} = \rho_{1,meatsproteins} + \rho_{1,meatsproteins,h} d_h + \rho_{1,meatsproteins,\tau} d_\tau \\
+ \rho_{1,meatsproteins,\chi} d_\chi + \rho_{1,meatsproteins,\xi} d_\xi + \gamma_{1,meatsproteins,fat} ln(P_{fat}) \\
+ \gamma_{1,meatsproteins,meatsproteins} ln(P_{meatsproteins}) + \gamma_{1,meatsproteins,cereals} ln(P_{cereals}) \\
+ \gamma_{1,meatsproteins,veg} ln(P_{veg}) + \gamma_{1,meatsproteins,milk} ln(P_{milk}) \\
+ \gamma_{1,meatsproteins,starches} ln(P_{starches}) + \gamma_{1,meatsproteins,complements} ln(P_{complements}) \\
+ \gamma_{1,meatsproteins,tubers} ln(P_{tubers}) + \gamma_{1,meatsproteins,fruits} ln(P_{fruits}) \\
+ \gamma_{1,meatsproteins,fish} ln(P_{fish}) + \beta_{meatsproteins,1} ln(x) + \epsilon_{meatsproteins,1}
\end{aligned}
$$

$qVmeatsproteins$:

$$
\begin{aligned}
ln\, V_{meatsproteins} = \rho_{2,meatsproteins} + \rho_{2,meatsproteins,h} d_h + \rho_{2,meatsproteins,\tau} d_\tau \\
+ \rho_{2,meatsproteins,\chi} d_\chi + \rho_{2,meatsproteins,\xi} d_\xi + \gamma_{2,meatsproteins,fat} ln(P_{fat}) \\
+ \gamma_{2,meatsproteins,meatsproteins} ln(P_{meatsproteins}) + \gamma_{2,meatsproteins,cereals} ln(P_{cereals}) \\
+ \gamma_{2,meatsproteins,veg} ln(P_{veg}) + \gamma_{2,meatsproteins,milk} ln(P_{milk}) \\
+ \gamma_{2,meatsproteins,starches} ln(P_{starches}) + \gamma_{2,meatsproteins,complements} ln(P_{complements}) \\
+ \gamma_{2,meatsproteins,tubers} ln(P_{tubers}) + \gamma_{2,meatsproteins,fruits} ln(P_{fruits}) \\
+ \gamma_{2,meatsproteins,fish} ln(P_{fish}) + \beta_{meatsproteins,2} ln(x) + \epsilon_{meatsproteins,2}
\end{aligned}
$$

For symmetry, the following restrictions are applied. Notice that $\gamma_{1,fat,meatsproteins}$ corresponds to the equation for fat (`qfat`).

$$
\begin{aligned}
\gamma_{1,meatsproteins,fat} &= \gamma_{1,fat,meatsproteins} \\
\gamma_{1,meatsproteins,cereals} &= \gamma_{1,cereals,meatsproteins} \\
\gamma_{1,meatsproteins,veg} &= \gamma_{1,veg,meatsproteins} \\
\gamma_{1,meatsproteins,milk} &= \gamma_{1,milk,meatsproteins} \\
\gamma_{1,meatsproteins,starches} &= \gamma_{1,starches,meatsproteins} \\
\gamma_{1,meatsproteins,complements} &= \gamma_{1,tubers,complements} \\
\gamma_{1,meatsproteins,tubers} &= \gamma_{1,tubers,meatsproteins} \\
\gamma_{1,meatsproteins,fruits} &= \gamma_{1,fruits,meatsproteins} \\
\gamma_{1,meatsproteins,fish} &= \gamma_{1,fish,meatsproteins}
\end{aligned}
$$

These conditions are repeated in the estimation for every other $G \in \{fat, cereals, veg, milk, meatsproteins, starches, complements, tubers, fruits, fish\}$ i.e. 10 commodities in total. The system of equation is thus used to observe the variation of measures of quality across the total expenditure and household characteristics in Tanzania. The entire set of equations consists of two equations for each commodity (10) - where one equation (per commodity) has budget-share as the dependent variable and another equation has quality-metric as the dependent variable. These $10 \times 2 = 20$ equations with respective constraints are listed in Appendix 7.4. The AIDS estimation for the system of equations that accommodates the constraints uses a 3SLS method (instead of SUR - which requires at least one regressor not used in other equations).

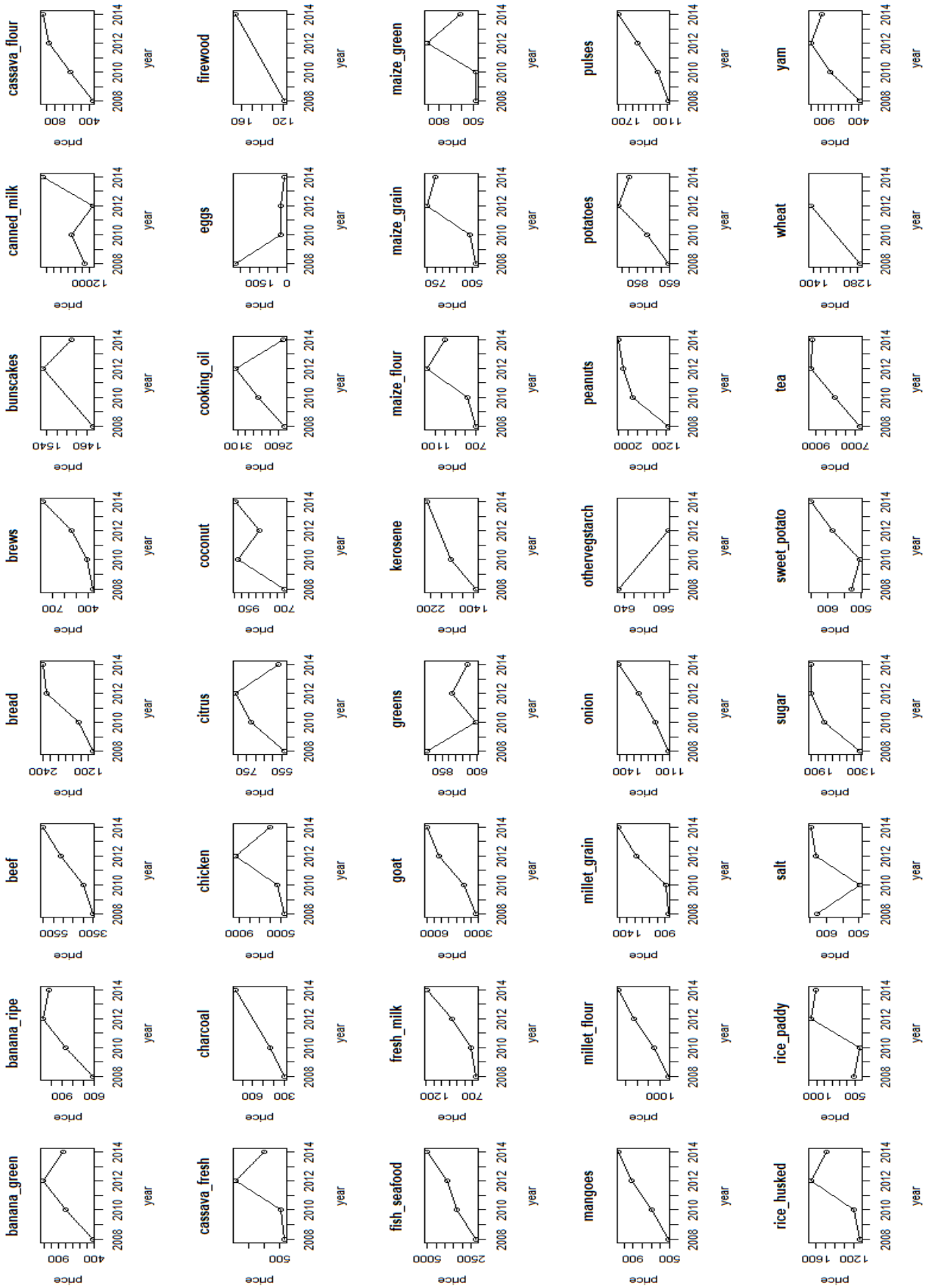


Figure 8: Market Prices recorded in 2008, 2010, 2012 and 2014 surveys

4 Data

The data used for our analysis has been obtained from the Living Standard Measurement Study (LSMS) - conducted by the World Bank - for the year 2014 - which is the last wave in the survey. The survey has been designed to assist in the understanding of links between asset ownership, community access (to market, schools), food prices and agriculture activities. The data in the surveys from multiple countries is largely uniform in its data-model (fields) and is often geo-referenced with a sufficiently wide-coverage in the country to allow a study of spatial variation in the observed variables. The years used for the analyses are the most recent years in the survey for Tanzania (2014) i.e. the last wave available in the survey at the time of writing this chapter. The price-data which we used in the classification described in Section 3.1 is based on prices for waves (years) 2008, 2010, 2012 and 2014.

The LSMS survey records past expenditure with diary and recall methods. The diary method is used for food items while other purchases are recorded as weekly, monthly and yearly recalls. The expenditures on food are recorded with a greater level of detail than other expenditures. The demographic variables from survey that we use in our a cross-sectional analysis are based on the current locality and the number of members in the family.

The consumption microdata for the year (2014) - used for budget shares and quality measures - is combined from across the recall and weekly diaries - merging the weekly data with yearly data by multiplying the past week's consumption into the number of weeks in the year. The details of this process are described in Appendix 7.1. Once combined, the market prices for the goods consumed by a household are matched with the reported consumed quantities of the food items [21].

The prices relevant for a household are marked as those that are available as market prices in the district where the household is located. Since the market-observed prices in the survey are recorded at every district level or at finer levels, we match these prices with every household in the survey based on the district the household is located in. This is necessitated also because not all wards and enumerated-areas(EAs) within every district in the survey have prices recorded for all relevant food items. Further, as price observations at the district level are recorded in different quantities (grams, kilograms for weight and litres, mls for fluids etc.), we infer the price at the district level using a regression over price records for the good in the district. This is only a means to standardise the prices in a standard unit such as a kilogram or a litre from the prices that are recorded in several units. Since the price per unit of any item declines with the increase in quantity purchase, the price data from across different recorded quantities are assimilated using a regression to obtain the price corresponding to a standard unit-quantity of an good (kg, litres etc.). A simple quadratic expression $\frac{m}{q} = \beta\frac{1}{q^2} + c$ to used to approximate the inverse proportionality of price-per-unit

[21] The absence of a good in the recall record for a household as the good is not to be interpreted as evidence of non-usage of the good. With the case of electricity, for example, the field for "main lighting fuel" is more reliable for the test of electricity usage than the last expenditure on electricity.

with the quantity purchase. Thus, observed marked-prices m are regressed against the quantities q they're recorded in where β is the coefficient and c is a constant . The price of any arbitrary quantity $\underline{q}$ of a good is thus simply $\frac{\beta}{\underline{q}^2} + c$ and for the price of unit-quantity that we use for base-prices and relative price structures, the price is simply $\beta + c$ (for $\underline{q} = 1$).

asset	*type*
bike	transport
motorbike	transport
car	transport
sewingmachine	household
bed	household
watch	household
chair	household
table	household
cupboard	household
sofa	household
sports_hobby	household
mobile	electric
waterheater	electric
camera	electric
phone	electric
musicplayer	electric
videoplayer	electric
musicsystem	electric
ac_fan	electric
waterpump	electric
tv	electric
dishtv	electric
computer	electric
refrigerator	electric
land	household
housing	household

Table 9: Long-term assets in the LSMS

The values of assets owned by the household that we use as a measure for wealth are obtained from the record of durable goods in the survey. The durable goods in the LSMS data range from short-term durable items such as mobile phones to the expensive and long-term assets such as land or housing. Of these, we are interested only in the transferable long-term owned assets (see Table 9). The exclusion of short-term assets or the items that are not transferred over a generation is based on consideration of wealth whose relevance is sought as an SES measure. The empirical studies often prefer the data on ownership of assets (Filmer & Pritchett, 1999, 2001) for SES effects since the inter-generational transfers of wealth are what make a break in the inequality gap (i.e. mobility through reduction of wealth differences) difficult (Obasuyi & Rasiah, 2019). Another reason why the income data has not been used directly in the current study is that it is available only for around 30% of the households that

we have the diary data for.

The assets transferred over a generation that we use for SES measures can be readily verified by inspecting households in the data that are split when a young member of the unsplit family starts a new household - treating the assets except land and house that are more expensive than the bed (furniture) as those that contribute to the long-term asset-values. The details of the records of asset-ownership in the LSMS data include the number n_t of durable goods owned by the household, the reported cost C_t at which the durable good was purchased and the reported price Π_t which the household expects by selling the durable good in the current market at the time of survey. We use the reported price Π_t of the durable goods - rather than reported cost C_t - to infer the asset-values in the area surrounding the household for two reasons. First, Π_t is not susceptible to the errors associated with the recall of the purchase value which varies over years in the panel data for the household and item. Second, Π_t also encapsulates the perceived depreciation of the durable good over time - making it more appropriate for a comparison against non-durable consumption at a given time t. The total cost of the assets of a household is thus obtained simply as a product of the number of assets n_t reported by the household and the price Π_t.

	All	Rural	Urban
Mean (SD) Household size	4.97 (2.91)	5.39 (3.06)	4.12 (2.39)
Mean (SD) age of household head	44.87 (16.05)	46.46 (16.56)	42.01 (14.69)
Mean (SD) number of rooms per household	3.28 (2.06)	3.56 (2.14)	2.79 (1.81)
Mean (SD) Total Expenditure (Tanzanian Shillings)	2004450 (2004450)	1495100 (1236870)	3017537 (2102633)
Percentage with household head educated secondary or higher	21.31	14.64	33.30
Percentage of household heads employed in Agriculture	55.63	66.37	10.74
Total Number of Households	4068	2707	1361

Table 10: Descriptive statistics for LSMS Tanzania 2014

5 Empirical Analysis

5.1 Descriptive Statistics

The summary statistics from the LSMS microdata for 2014 are presented in the Table 10. Tanzania appears to be a relatively young economy with a relatively large family size - given

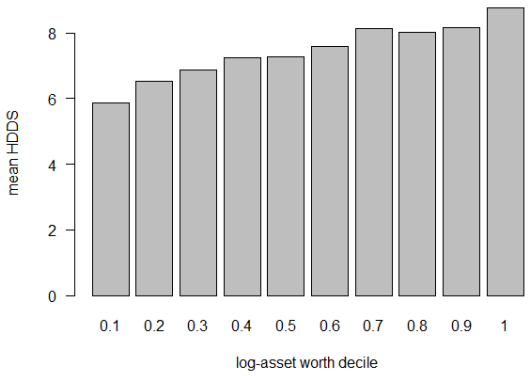


Figure 9: Diversity Score changes by log asset deciles

the average household member being nearly 5 and the average age of the family-head being 45 years. Also remarkable is that a significantly higher number of household heads in the rural areas are employed in agriculture.

A quick survey of food diversity scores shows that the dietary diversity varies significantly across Tanzania (see Figure 9). However, they do not indicate an extreme segmentation based on assets as is reported for many economies in the region (Nti, 2011; Hailu & Woldemichael, 2019). It is worth emphasising that FAO standards do not prescribe what's considered a low dietary diversity status and it is common for studies to therefore consider empirically validated thresholds for the analysis of diversity. (McDonald et al., 2015) , for example, set this threshold to 3. Considering the threshold diversity score of 3, the diversity scores in Tanzania do not appear too low (see Figure 9).

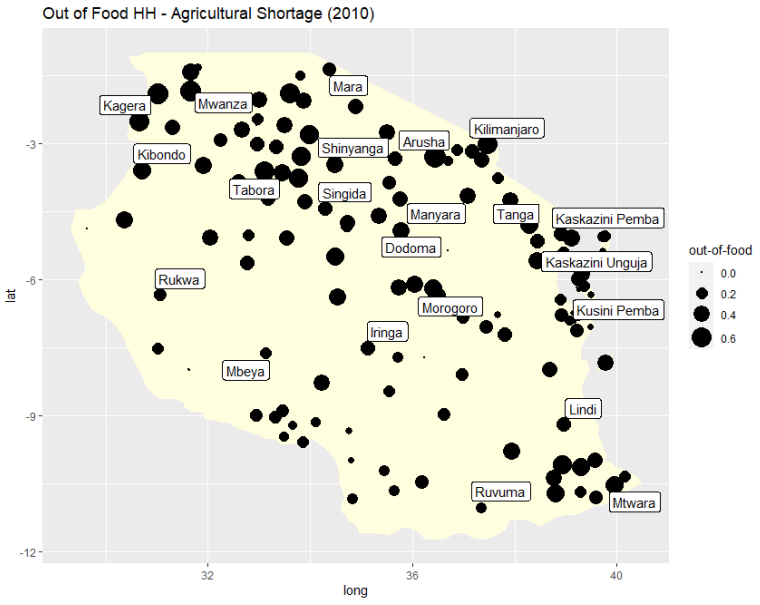


Figure 10: Fraction of households in Tanzania having run of food due to agricultural inputs in the past year (2010)

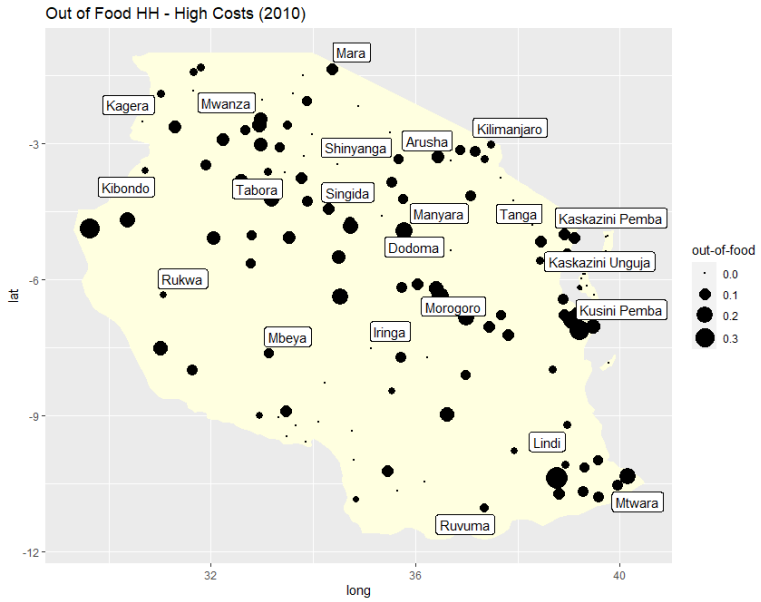


Figure 11: Fraction of households having run of food due to costs in the past year (2010)

The spatial distribution of food insecurity provides us an overview of regional disparities in the country. To survey food insecurity, we use the directly observed food deprivation in the survey to compare the fraction of households having run out of food (either due to lack of agricultural inputs or high prices) in the past year across urban and rural areas (see Figures 10 and 11). While the figures do suggest a significant proportion of households running out of food in urban areas (due to high costs), the overall hunger instances remain proportionately lower in the urban areas (when compared to the households having run out of food due to agricultural reasons- see Figure 11). While the food insecurity does not seem significantly stronger in either urban or rural areas, the extremity of urban-rural or regional differences seems evident from a comparison of assets and occupations. As a non-parametric view of occupations[22] and value of assets-owned show (Figure 12), the higher-paid occupations seems severely to urban - particularly eastern coastal - areas in Tanzania.

5.2 Results

The key goal of our empirical analysis is to test whether the effects of permanent income (as well as asset-ownership) result in higher quality overall or if the effects are limited to specific food-categories in Tanzania. The disaggregated view of quality into commodities based on FAO categories allows us to relate the choice for quality in high-level commodity groups with the effects of permanent income - while also controlling with local prices in an AIDS framework. The effect of permanent income measures seems significant on food quality with respect to the meats-proteins commodity. We also find that access to higher food quality

[22]The variable `occupation_rank` shown in the figure is defined as an ordered variable indicating an income-based rank inferred from the occupation of the head-household using the available income data (see Appendix 7.3 for mapping of occupations to occupation-ranks).

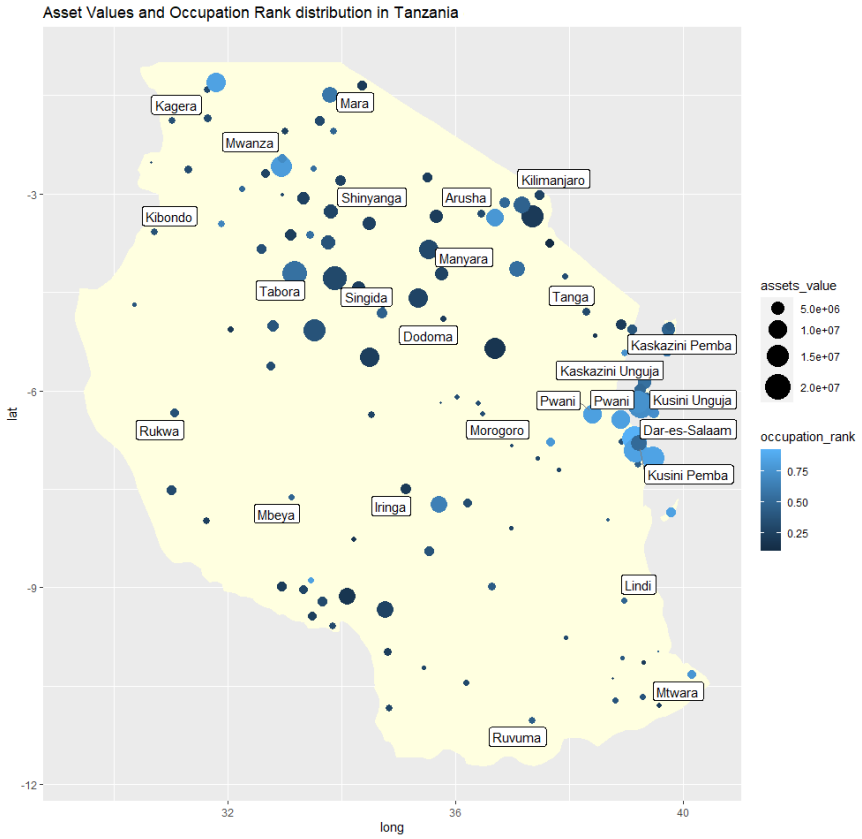


Figure 12: Occupations and consumer-owned assets in Tanzania

is higher in localities with access to electricity.

Variable	Description
`ln_tot_exp`	logarithm of per-head total expenditure x_t per household
`lp-`	base-prices for various commodities
`hsize`	number of members in a household
`age`	age of the household-head
`logA`	logarithm of sum of reported market-price of assets owned by the household
`isurban`	indicator variable specifying if the region is Dar-es-Salaam, Mbeya, Mwanza or Magharibi Unguja(Zanzibar)
`has_electric`	indicator variable specifying if the household uses electricity or not

Table 11: Control Variables

The estimation results from the AIDS formulation - which we have specified in Equations 16 and 17 - are presented in Tables 13 and 14. The logarithm of the total expenditure `ln_tot_exp` is used as the main explanatory variable in the results presented in

Table 13 . In an alternative formulation for which the results presented in Table 14, the logarithm of total household owned (long-term) assets `logA` is used as the main explanatory variable. An important set of controls are the prices corresponding to the food commodities - which are listed in Table 5. As described in Section 3.3, the AIDS formulation in Equations 16 and 17 use the logarithm of base-prices of food categories (commodities) as price-controls. These are represented as the variables starting with the prefix `lp-` (e.g., `lpdensefoods, lpnonfresh, lpprotein` etc.) in the results presented in Tables 13 and 14. Recall that the base-price P_G for every commodity G in the Equation 14 is simply the minimum price of the goods within the commodity group that is derived from the locally observed market prices (at the district level) rather than the unit-values (estimated by averaging expenditure over a locality) for goods within a commodity. Every composite commodity in the AIDS regression (see Equations 16 and 17) thus corresponds to an `lp-` variable. It is worth remarking that the prices of other items in the commodity are incorporated in the price-structure that is considered in the measurement of quality. Other important controls used for results in Tables 13 and 14 are listed in Table 21.

Of the controls in Table 21, the indicator variable `has_electric` - specifies whether the household uses electricity (used in both Tables 13 and 14) and is inferred from the field for source of energy for lighting in the survey. Two further household characteristics which we explore the effects of are `hsize` (number of household members) and `age` (age of the household-head). Whether the households reside in urban areas or not are considered with a indicator variable `isurban` (indicator variable for densely populated urban areas).

Variable	Description
qcereals	budget share for the entire commodity: cereals
qVcereals	quality metric for the entire commodity: cereals
qcomplements	budget share for the entire commodity: complements
qVcomplements	quality metric for the entire commodity: complements
qfat	budget share for the entire commodity: fat
qVfat	quality metric for the entire commodity: fat
qfish	budget share for the entire commodity: fish
qVfish	quality metric for the entire commodity: fish
qfruits	budget share for the entire commodity: fruits
qVfruits	quality metric for the entire commodity: fruits
qmeatsproteins	budget share for the entire commodity: meats-proteins
qVmeatsproteins	quality metric for the entire commodity: meats-proteins
qmilk	budget share for the entire commodity: milk
qVmilk	quality metric for the entire commodity: milk
qstarches	budget share for the entire commodity: starches
qVstarches	quality metric for the entire commodity: starches
qtubers	budget share for the entire commodity: tubers
qVtubers	quality metric for the entire commodity: tubers
qveg	budget share for the entire commodity: vegetables
qVveg	quality metric for the entire commodity: vegetables

Table 12: Dependent Variables

The dependent variables of the AIDS regressions in the columns of Tables 13 and 14- starting either with q- or with qV - indicate the budget-share and quality metric respectively. The columns starting with q- indicate the budget shares that are used as the dependent variables in AIDS regressions (see Equation 16) while the columns starting with qV- indicate the logarithm of the quality metric as the dependent variable (i.e. $ln\ V$ in the Equation 17). The list of all dependent variables are presented in Table 12. Recall that the classification of items - which is based on prices-correlation - results in the commodities and constituents good listed in Table 5. The total number of equations (the Equations 16 and 17 specified in Section 3.3) for the estimation is therefore $10 \times 2 = 20$ (see Appendix 7.4 for the entire list of estimation equations). The estimation uses a three-stage least-squares method for systems of equations (see (Zellner & Theil, 1992)) for the Equations 16 and 17 and with symmetry constraints specified in Section 3.3.

Discussion

The results obtained from the AIDS regression shown in Tables 13 and 14 suggest a relevance of permanent income (measured with total expenditure) and ownership of assets -

along with access to electricity - on the demand for quality in food. While there is a some heteroskedasticity in the results (tested for with Breush-Pagan statistic), the effects of household expenditure and total assets owned appear significant for quality.

In Table 13, the effect of permanent income (`ln_tot_exp`) is positive on quality across all food commodities i.e. the income-elasticities are positive for quality within food categories (commodities). This does not seem surprising since the income-elasticities for quality are expected to be positive so long as a higher total expenditure `ln_tot_exp` is spent on the more expensive variants (goods) within a food category (commodity). The negative income elasticities for budget-share of some food commodities are also predictable - if we consider that certain goods are preferred for higher (price-based) quality and therefore lower quantity (budget share). Of particular attention is how the budget shares show a decline with total expenditure logarithm for commodities `fat` and `meatsproteins` (see coefficients of `ln_tot_exp` under `qfat` and `qmeatsproteins` in Table 13) . More specifically, the increase in food expenditure is associated with less quantity (at the cost of higher quality) of `fats` and `meatsproteins` - while quality consumed for both fats and meats-proteins rises with higher total expenditure logarithm (see coefficients of `ln_tot_exp` under `qVfat` and `qVmeatsproteins` in Table 13). With cereals, on the other hand, we see that the budget shares rise with higher total expenditure logarithm (see coefficient of `ln_tot_exp` under `qcereals` in Table 13) while the cereals-quality also increases with higher total expenditure logarithm (see coefficient of `ln_tot_exp` under `qVcereals` in Table 13). The demand for quality (price-based) for meat is unlike that for cereals in the sense that higher-income households increase both price-based quality and quantity (budget-shares) for cereals whereas they seem increase quality without increasing quantity (budget shares) for meats.

Looking at the effect of access to electricity (`has_electric`) in Table 13, we see a higher quality across food categories - an effect similar to that of total expenditure logarithm `ln_tot_exp` (except for fats where quality is not significantly higher for consumers with access to electricity). It could be argued that having access to electricity is endogenous with higher prosperity. Thus it may simply be that those with higher income are the one with access to electricity and have a better access to quality through ownership of electric appliances. However, the correlation of having access to electricity with prosperity is more likely through being in the urban area rather through income. This is because the indicator variable `has_electric` used for access to electricity is based only on the usage of electricity as the primary source of lighting at home. The basic usage of electricity hardly reflects the income of the household - and is more strongly correlated with being in an area (more often urban) where electricity is available (particularly given the sparse distribution of electricity in Tanzania). Since availability of refrigeration facilities in a market may improve the availability of fruits and meats in a wider area, the positive externality through being in richer areas - one that includes access to electricity - is more likely to be the cause for higher quality in food categories.

The effects of hsize on quality are also worth noting - as larger families (higher hsize) corresponds to lower quality (particularly in protein and fruits-veg) but lower their budget-share for consumption. This may struggle in maintaining high quality in larger families with children.

A key advantage of using the proposed framework of quality is that it allows us to examine the effect of prices of commodities on quality. Looking at the effect of (base-)prices (variables lp-) in both Tables 13 and 14, we see with lpmeatsproteins - for example - that the high base-price of meats-proteins (the cheapest meat-protein available in the district) corresponds to higher quality in fruits and veg. Thus, the households in districts where base-price of meatsproteins (price of cheapest meat-protein) is high also avail higher quality of fruits and veg - an indication of how meat consumption may be associated with higher income in Tanzania.

The results in Table 13 also suggest certain substitutions in quality across some food commodities. For example, with a rise in prices of cereals (lpcereals), the quality in meats-proteins (see coefficient of lpcereals under qVmeatsproteins in Table 13) increases but the quality in starches (see coefficient of lpcereals under qVstarches in Table 13) and vegetables (see coefficient of lpcereals under qVveg in Table 13) declines - indicating that the consumers tend to increase quality in meat-consumption while reducing the quality for starches and veg. In other words, with higher prices of (cheapest) meat, the consumer is more likely to increase quality elsewhere (e.g., in fruits and vegetables) - indicating once again how meat consumption rises with higher income.

On similar lines, we see that the quality in starches seems to drop with rise in prices of fruits (see coefficient of lpfruits under qVstarches in Table 13) whereas the quality in fruits does not rise with the base-price of starches (see coefficient of lpstarches under qVfruits in Table 13). In other words, when (cheapest) fruit is more expensive, starch quality is also lower for consumers whereas (cheapest) starch being more expensive does not raise the quality of fruits for consumers. It is more often thus that the consumers spend higher on quality in starches when the fruits become expensive (rather than spending higher on quality in fruits when starch becomes more expensive). A similar relation is not seen between tubers and starches whose consumption of quality seem to move together.

The observation from the budget shares that the expenditure on veg and meatsproteins is higher - particularly for meatsproteins - in urban areas can also be verified with a simple calculation of total expenditure divided by the commodity base-price. Estimating the effective quantity of meatsproteins consumed as the meatsproteins expenditure divided by its base-price, the urban areas seem to consume about 1.61 times more meatsproteins and spend 1.63 times more (comparing just expenditure) - not as much for the expenditure on chicken or fish. With effective quantity of veg, the urban areas consume about 1.1 times more meatsproteins and spend also about 1.1 times more than the rural areas. Looking at the expenditures in urban and rural areas for cereals, we

see that unlike with meatsproteins, the differences in expenditure (or in effective quantity) for cereals across urban and rural areas and are not wide. The expenditures also suggest that while the consumption of veg rises with total expenditure logarithm ln_tot_exp , it does not rise as much as meatsproteins in urban areas and does not differ a lot from the expenditures in rural areas.

In Table 14 - where total assets logarithm logA is used as the main explanatory variable instead of total expenditure ln_tot_exp - the coefficient of total assets logarithm is positive for meatsproteins (dependent variable: qmeatsproteins) - suggesting that those with lower assets do consume less meatsproteins overall. In other words, the lower quantity of meatsproteins for those with low-assets-ownership does not mean a higher quality in meatsproteins (compare signs under qmeatsproteins of the coefficients for logA and ln_tot_exp in Tables 13 and 14). The quality consumed for meats-proteins does rise with higher total assets logarithm as well. Comparing with cereals, on the other hand, we see that the budget shares are lower for those with higher total assets - even though the quality consumed is higher.

The access to electricity (has_electric) indicates a higher quality across food categories in Table 14 as well. The budget shares rise for fruits and milk in Table 14 whereas they fall for fat (qfat) and meatsproteins (qmeatsproteins) - even though the quality metrics rise (both qVfat and qVmeatsproteins) for the consumers having electricity. This may be due to refrigeration storage facilities available in certain localities. The consumption of fish (budget shares qfish) is not affected by the has_electric variable as much as meatsproteins - suggesting a higher local consumption for fish (likely less dependent on refrigeration facilities overall). For fruits and milk on the other hand, both the quality consumed and budget-share are higher for consumers with access to electricity (see Table 14).

In both Tables 13 and 14, we see that the effect of isurban (apart from the base-price P_G) on quality is significant on quality across all food categories. This could be an indication of higher income occupations concentrated in the eastern coast and centre of Tanzania. More specifically, quality for meatsproteins (qVmeatsproteins) is slightly lower in urban areas (compare coefficients of isurban under qVmeatsproteins and qVfruits in both Tables 13 and 14) and is better with access to electricity. In other words, those with no access to electricity in the urban areas are less likely to spend higher on quality of meats-proteins compared to the consumers in rural areas. The effects of both total expenditure and assets-ownership are significant in the quality of consumption for food. The access to electricity in urban environment creates clear differences in lifestyle choices as well as dependence on higher budgets - an effect that remains significant for access to quality in nutrition. In summary, even though there may be some positive externality with being in the urban areas, the differences in quality of consumption associated with high permanent income as well as assets possession remain significant.

While the question of whether higher quality for food can be considered aspirational or not could be answered only after surveys of subjective well-being of consumers, there are two observations pertaining to determinants of aspirational consumption that we can make from the above results. First, the demand for quality is strong for specific food items (rather than across wider nutritional categories). Second, the role of regional identity is likely to be strong towards aspirational consumption due to the extreme life-style differences between urban and rural areas. The observations seem to be supported by the presence of a strong positive externality with rich surroundings in the data on subjective-welfare in LSMS survey in Tanzania as well (Atsebi & i Carbonell, 2019).

Robustness Checks

As a test of the robustness of results, we calculate the food-quality using the market-provided prices and instead of the expenditure reported by the household. More specifically, in the results shown in Tables 13 and 14, we have used diary-costs on goods as reported by the household - costs that are recorded by the LSMS for the goods that were consumed from purchases in the last week. If a household didn't purchase anything in the past week, there would only be quantities consumed (not costs) that would be reported for the household 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 households 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 food categories (commodities).

Table 13: reg3 results with electricity usage indicator control (2014)

	(1) qcereals	(2) qcomplements	(3) qfat	(4) qfish	(5) qfruits	(6) qmeatsproteins	(7) qmilk	(8) qstarches	(9) qtubers	(10) qVcereals	(11) qVcomplements	(12) qveg	(13) qVfat	(14) qVfish	(15) qVfruits	(16) qVmeatsproteins	(17) qVmilk	(18) qVstarches	(19) qVtubers	(20) qVveg
isurban	-0.0317***	-0.0170***	0.0002	0.0197***	0.0094*	0.0118	-0.0041*	0.0074**	-0.0039	-0.1542**	0.0360	0.0055*	-0.8281	0.0484*	0.1402**	-0.4567***	-0.0581***	0.0190	0.0577	0.1297***
	(0.000)	(0.000)	(0.964)	(0.000)	(0.020)	(0.120)	(0.043)	(0.007)	(0.260)	(0.004)	(0.104)	(0.021)	(0.283)	(0.011)	(0.003)	(0.000)	(0.000)	(0.764)	(0.121)	(0.000)
has_electric	-0.0641***	-0.0095*	0.0019	-0.0050	0.0253***	0.0296***	0.0066**	0.0175***	-0.0041	0.0828	0.0294	-0.0008	-1.4587	0.0367	0.2107***	0.4544***	0.0939***	0.5391***	0.1254**	0.1341***
	(0.000)	(0.029)	(0.653)	(0.398)	(0.000)	(0.000)	(0.002)	(0.000)	(0.278)	(0.136)	(0.206)	(0.767)	(0.071)	(0.069)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
hh_age	0.0002	0.0002	-0.0003**	0.0004*	-0.0002*	0.0002	-0.0000	-0.0001	-0.0002*	0.0023	0.0008	0.0000	-0.0081	-0.0005	-0.0027*	0.0044**	-0.0003	-0.0019	-0.0024*	-0.0013**
	(0.407)	(0.057)	(0.009)	(0.015)	(0.020)	(0.274)	(0.560)	(0.301)	(0.044)	(0.100)	(0.180)	(0.883)	(0.688)	(0.321)	(0.028)	(0.003)	(0.403)	(0.253)	(0.015)	(0.004)
hsize	-0.0035**	0.0042***	0.0026***	0.0038***	-0.0018**	-0.0003	-0.0014***	-0.0009*	-0.0004	-0.0887***	-0.0310***	-0.0009*	0.2308*	0.0000	-0.0484***	-0.0574***	-0.0141***	-0.0294**	-0.0228***	-0.0208***
	(0.004)	(0.000)	(0.000)	(0.000)	(0.003)	(0.808)	(0.000)	(0.029)	(0.478)	(0.000)	(0.000)	(0.011)	(0.049)	(0.994)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
ln_tot_exp	0.0641***	-0.0091***	-0.0326***	-0.0052*	0.0095***	-0.0532***	0.0042***	0.0112***	0.0049**	0.6292***	0.2089***	-0.0002	0.6893*	0.1220***	0.3857***	0.5941***	0.0799***	0.4104***	0.1870***	0.1455***
	(0.000)	(0.000)	(0.000)	(0.025)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.855)	(0.032)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
lpcereals	-0.0060	0.0001	-0.0014	0.0031	0.0166***	-0.0166***	-0.0011	0.0023	0.0049**	-1.5037***	0.0516**	-0.0018	0.6404	0.0194	0.0237	0.3641***	-0.0228*	-0.1978***	-0.0431	-0.0503***
	(0.253)	(0.970)	(0.397)	(0.187)	(0.000)	(0.000)	(0.252)	(0.133)	(0.002)	(0.000)	(0.001)	(0.217)	(0.247)	(0.100)	(0.477)	(0.000)	(0.012)	(0.000)	(0.088)	(0.000)
lpcomplements	0.0001	0.0019	0.0003	-0.0012	0.0016	0.0060	-0.0005	0.0015	-0.0066***	1.5261***	-0.0461	-0.0031	4.4315**	-0.0882**	0.0245	1.7650***	0.0163	0.4861***	0.0484	-0.1422***
	(0.970)	(0.805)	(0.768)	(0.429)	(0.615)	(0.273)	(0.522)	(0.263)	(0.000)	(0.000)	(0.260)	(0.058)	(0.002)	(0.001)	(0.767)	(0.000)	(0.419)	(0.000)	(0.440)	(0.000)
lpfat	-0.0014	0.0003	0.0002	0.0013	-0.0020*	0.0032	-0.0008	-0.0019**	0.0013	-0.0110	-0.0236***	-0.0002	-8.8267***	0.0064	0.0098	-0.0016	-0.0028	-0.1016***	0.0115	-0.0025
	(0.397)	(0.768)	(0.850)	(0.202)	(0.044)	(0.075)	(0.077)	(0.003)	(0.065)	(0.419)	(0.000)	(0.714)	(0.000)	(0.156)	(0.426)	(0.916)	(0.423)	(0.000)	(0.224)	(0.590)
lpfish	0.0031	-0.0012	0.0013	0.0052*	-0.0027	-0.0041	-0.0023***	-0.0015	0.0030**	0.0100	-0.0214*	-0.0007	-0.4615	0.0413***	0.0326	-0.0128	-0.0078	0.0209	0.0306*	0.0004
	(0.187)	(0.429)	(0.202)	(0.013)	(0.059)	(0.110)	(0.000)	(0.104)	(0.003)	(0.606)	(0.011)	(0.382)	(0.111)	(0.000)	(0.063)	(0.542)	(0.116)	(0.374)	(0.023)	(0.946)
lpfruits	0.0166***	0.0016	-0.0020*	-0.0027	0.0079*	-0.0243***	0.0015*	0.0023*	0.0005	0.0458	-0.0046	-0.0013	-0.9906	-0.0084	-0.3593***	-0.0227	-0.0027	-0.1496**	-0.0263	-0.0169
	(0.000)	(0.615)	(0.044)	(0.059)	(0.010)	(0.000)	(0.028)	(0.043)	(0.632)	(0.259)	(0.795)	(0.301)	(0.112)	(0.498)	(0.000)	(0.601)	(0.773)	(0.002)	(0.343)	(0.195)
lpmeatsproteins	-0.0166***	0.0060	0.0032	-0.0041	-0.0243***	0.0382***	0.0017	-0.0054**	-0.0044*	-0.0368	0.0432	0.0058**	8.0136***	0.0343	0.1864**	-2.2568***	0.0109	-0.0440	-0.0364	0.1751***
	(0.000)	(0.273)	(0.075)	(0.110)	(0.000)	(0.000)	(0.153)	(0.003)	(0.012)	(0.593)	(0.153)	(0.002)	(0.000)	(0.104)	(0.003)	(0.000)	(0.486)	(0.597)	(0.438)	(0.000)
lpmilk	-0.0011	-0.0005	-0.0008	-0.0023***	0.0015*	0.0017	0.0019***	0.0001	-0.0014**	0.0094	-0.0048	0.0009*	-0.2357	-0.0076*	0.0283**	-0.0062	0.0176***	0.0107	0.0017	0.0032
	(0.252)	(0.522)	(0.077)	(0.000)	(0.028)	(0.153)	(0.000)	(0.752)	(0.001)	(0.333)	(0.257)	(0.023)	(0.109)	(0.016)	(0.001)	(0.554)	(0.000)	(0.369)	(0.806)	(0.330)
lpstarches	0.0023	0.0015	-0.0019**	-0.0015	0.0023*	-0.0054**	0.0001	0.0032**	-0.0002	0.0228	0.0239**	-0.0003	-0.2865	0.0002	0.0124	0.0015	-0.0043	-0.0307	-0.0022	-0.0020
	(0.133)	(0.263)	(0.003)	(0.104)	(0.043)	(0.003)	(0.752)	(0.001)	(0.699)	(0.173)	(0.001)	(0.598)	(0.260)	(0.962)	(0.415)	(0.936)	(0.293)	(0.145)	(0.852)	(0.715)
lptubers	0.0049**	-0.0066***	0.0013	0.0030**	0.0005	-0.0044*	-0.0014**	-0.0002	0.0037***	-0.0097	0.0020	-0.0007	-0.6671**	0.0180***	-0.0171	0.0468**	-0.0108**	-0.0172	0.0179	0.0015
	(0.002)	(0.000)	(0.065)	(0.003)	(0.632)	(0.012)	(0.001)	(0.699)	(0.000)	(0.479)	(0.730)	(0.234)	(0.001)	(0.000)	(0.169)	(0.002)	(0.002)	(0.303)	(0.073)	(0.753)
lpveg	-0.0018***	-0.0031***	-0.0002***	-0.0007***	-0.0013	0.0058***	0.0009*	-0.0003***	-0.0007	-0.0530***	-0.0202***	0.0014**	-1.6173	-0.0154***	0.0586***	0.1228***	0.0062***	0.0231***	-0.0021***	0.0337***
	(0.217)	(0.058)	(0.714)	(0.382)	(0.301)	(0.002)	(0.023)	(0.598)	(0.234)	(0.011)	(0.028)	(0.157)	(0.000)	(0.017)	(0.002)	(0.000)	(0.205)	(0.365)	(0.886)	(0.000)
_cons	-0.6042***	0.1824***	0.5310***	0.1178***	-0.0430	0.9945***	-0.0289*	-0.1060***	-0.0190	-8.3633***	-1.7428***	0.0398**	-2.3817	-1.0892***	-4.0316***	-6.9911***	-0.8506***	-5.2349***	-2.0281***	-1.2673***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.068)	(0.000)	(0.012)	(0.000)	(0.347)	(0.000)	(0.000)	(0.004)	(0.623)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

p-values in parentheses

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

Table 14: reg3 results with total assets (logarithm) and electricity usage indicator control (2014)

	(1) qcereals	(2) qcomplements	(3) qfat	(4) qfish	(5) qfruits	(6) qmeatsproteins	(7) qmilk	(8) qstarches	(9) qtubers	(10) qVcereals	(11) qVcomplements	(12) qveg	(13) qVfat	(14) qVfish	(15) qVfruits	(16) qVmeatsproteins	(17) qVmilk	(18) qVstarches	(19) qVtubers	(20) qVveg
isurban	-0.0134	-0.0197***	-0.0062	0.0194***	0.0104**	-0.0028	-0.0040*	0.0079**	-0.0024	-0.0367	0.0658**	0.0058*	-0.7816	0.0705***	0.2010***	-0.3775***	-0.0484**	0.0675	0.0893*	0.1544***
	(0.111)	(0.000)	(0.133)	(0.000)	(0.009)	(0.724)	(0.044)	(0.004)	(0.484)	(0.526)	(0.005)	(0.015)	(0.310)	(0.000)	(0.000)	(0.000)	(0.001)	(0.295)	(0.018)	(0.000)
has_electric	-0.0047	-0.0210***	-0.0211***	-0.0043	0.0298***	-0.0172*	0.0078***	0.0217***	0.0018	0.4745***	0.1259***	0.0003	-1.3344	0.1183***	0.4277***	0.7374***	0.1332***	0.7203***	0.2375***	0.2203***
	(0.594)	(0.000)	(0.000)	(0.457)	(0.000)	(0.034)	(0.000)	(0.000)	(0.635)	(0.000)	(0.000)	(0.911)	(0.094)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
logA	-0.0077***	0.0042***	-0.0016	-0.0043***	0.0021*	0.0045**	0.0014**	0.0031***	-0.0019*	0.0633***	0.0464***	-0.0010*	0.2649	0.0042	0.0521***	0.1156***	0.0146***	0.0958***	0.0191*	0.0156***
	(0.000)	(0.000)	(0.084)	(0.000)	(0.013)	(0.008)	(0.001)	(0.000)	(0.015)	(0.000)	(0.000)	(0.049)	(0.105)	(0.316)	(0.000)	(0.000)	(0.000)	(0.000)	(0.018)	(0.000)
hh_age	-0.0001	0.0002	-0.0000	0.0005**	-0.0004**	0.0005*	-0.0001	-0.0002**	-0.0002	-0.0036*	-0.0018**	0.0000	-0.0196	-0.0015**	-0.0067***	-0.0027	-0.0012**	-0.0071***	-0.0042***	-0.0027***
	(0.775)	(0.123)	(0.910)	(0.001)	(0.001)	(0.026)	(0.077)	(0.002)	(0.062)	(0.019)	(0.003)	(0.584)	(0.339)	(0.006)	(0.000)	(0.096)	(0.002)	(0.000)	(0.000)	(0.000)
hsize	0.0036**	0.0024***	0.0003	0.0045***	-0.0016*	-0.0058***	-0.0014***	-0.0008	0.0005	-0.0519***	-0.0249***	-0.0007	0.2275	0.0094**	-0.0289***	-0.0355***	-0.0109***	-0.0187	-0.0117	-0.0124***
	(0.008)	(0.000)	(0.673)	(0.000)	(0.015)	(0.000)	(0.000)	(0.091)	(0.387)	(0.000)	(0.000)	(0.071)	(0.061)	(0.003)	(0.000)	(0.000)	(0.000)	(0.067)	(0.052)	(0.000)
lpcereals	0.0015	-0.0017	-0.0082***	0.0024	0.0189***	-0.0221***	0.0001	0.0056***	0.0057***	-1.4272***	0.0747***	-0.0021	0.6073	0.0339**	0.0758*	0.4352***	-0.0096	-0.1343**	-0.0203	-0.0341**
	(0.788)	(0.560)	(0.000)	(0.324)	(0.000)	(0.000)	(0.913)	(0.000)	(0.000)	(0.000)	(0.000)	(0.163)	(0.272)	(0.006)	(0.028)	(0.000)	(0.299)	(0.003)	(0.430)	(0.008)
lpcomplements	-0.0017	0.0030	0.0017	-0.0008	0.0001	0.0076	-0.0009	0.0004	-0.0066***	1.3209***	-0.1476**	-0.0028	3.5007*	-0.1468***	-0.1352	1.4214***	-0.0148	0.3250**	-0.0329	-0.2106***
	(0.560)	(0.692)	(0.124)	(0.587)	(0.962)	(0.165)	(0.262)	(0.756)	(0.000)	(0.000)	(0.001)	(0.079)	(0.013)	(0.000)	(0.112)	(0.000)	(0.472)	(0.004)	(0.603)	(0.000)
lpfat	-0.0082***	0.0017	0.0038***	0.0013	-0.0030**	0.0088***	-0.0013**	-0.0033***	0.0005	-0.0690***	-0.0407***	-0.0003	-8.8596***	-0.0059	-0.0261*	-0.0518**	-0.0109**	-0.1391***	-0.0064	-0.0160**
	(0.000)	(0.124)	(0.000)	(0.197)	(0.002)	(0.000)	(0.002)	(0.000)	(0.484)	(0.000)	(0.000)	(0.589)	(0.000)	(0.204)	(0.040)	(0.001)	(0.002)	(0.000)	(0.500)	(0.001)
lpfish	0.0024	-0.0008	0.0013	0.0047*	-0.0024	-0.0034	-0.0023***	-0.0013	0.0027**	0.0141	-0.0179*	-0.0009	-0.4367	0.0412***	0.0364*	-0.0039	-0.0072	0.0279	0.0314*	0.0011
	(0.324)	(0.587)	(0.197)	(0.023)	(0.086)	(0.191)	(0.000)	(0.158)	(0.006)	(0.497)	(0.042)	(0.296)	(0.132)	(0.000)	(0.045)	(0.864)	(0.157)	(0.245)	(0.022)	(0.868)
lpfruits	0.0189***	0.0001	-0.0030**	-0.0024	0.0082**	-0.0260***	0.0017*	0.0028*	0.0008	0.0937*	0.0149	-0.0012	-0.8222	0.0065	-0.3217***	0.0516	0.0053	-0.1163*	-0.0070	-0.0003
	(0.000)	(0.962)	(0.002)	(0.086)	(0.008)	(0.000)	(0.012)	(0.013)	(0.409)	(0.027)	(0.423)	(0.357)	(0.188)	(0.615)	(0.000)	(0.281)	(0.578)	(0.020)	(0.805)	(0.983)
lpmeatsproteins	-0.0221***	0.0076	0.0088***	-0.0034	-0.0260***	0.0423***	0.0005	-0.0086***	-0.0051**	0.0170	0.0843**	0.0060**	8.6268***	0.0567*	0.2325***	-2.1152***	0.0163	-0.0114	-0.0076	0.2018***
	(0.000)	(0.165)	(0.000)	(0.191)	(0.000)	(0.000)	(0.689)	(0.000)	(0.004)	(0.813)	(0.008)	(0.002)	(0.000)	(0.010)	(0.000)	(0.000)	(0.310)	(0.893)	(0.872)	(0.000)
lpmilk	0.0001	-0.0009	-0.0013**	-0.0023***	0.0017*	0.0005	0.0020***	0.0005	-0.0012**	0.0279**	0.0021	0.0009*	-0.1894	-0.0029	0.0411***	0.0164	0.0204***	0.0241*	0.0082	0.0084*
	(0.913)	(0.262)	(0.002)	(0.000)	(0.012)	(0.689)	(0.000)	(0.256)	(0.003)	(0.007)	(0.629)	(0.021)	(0.198)	(0.379)	(0.000)	(0.152)	(0.000)	(0.045)	(0.231)	(0.015)
lpstarches	0.0056***	0.0004	-0.0033***	-0.0013	0.0028*	-0.0086***	0.0005	0.0040***	0.0002	0.0770***	0.0452***	-0.0002	-0.1288	0.0148**	0.0509**	0.0721***	0.0040	0.0087	0.0177	0.0142*
	(0.000)	(0.756)	(0.000)	(0.158)	(0.013)	(0.000)	(0.256)	(0.000)	(0.788)	(0.000)	(0.000)	(0.704)	(0.610)	(0.007)	(0.001)	(0.000)	(0.325)	(0.681)	(0.126)	(0.015)
lptubers	0.0057***	-0.0066***	0.0005	0.0027**	0.0008	-0.0051**	-0.0012**	0.0002	0.0037***	0.0020	0.0065	-0.0008	-0.6590**	0.0197***	-0.0094	0.0583***	-0.0090*	-0.0069	0.0209*	0.0037
	(0.000)	(0.000)	(0.484)	(0.006)	(0.409)	(0.004)	(0.003)	(0.788)	(0.000)	(0.892)	(0.292)	(0.184)	(0.001)	(0.000)	(0.463)	(0.000)	(0.011)	(0.684)	(0.040)	(0.452)
lpveg	-0.0021***	-0.0028	-0.0003***	-0.0009***	-0.0012***	0.0060***	0.0009	-0.0002	-0.0008***	-0.0564**	-0.0215***	0.0013***	-1.6391	-0.0172***	0.0558***	0.1160	0.0056	0.0223**	-0.0040*	0.0317***
	(0.163)	(0.079)	(0.589)	(0.296)	(0.357)	(0.002)	(0.021)	(0.704)	(0.184)	(0.010)	(0.026)	(0.188)	(0.000)	(0.011)	(0.005)	(0.000)	(0.264)	(0.391)	(0.784)	(0.000)
_cons	0.3323***	0.0112	0.1167***	0.0995***	0.0596***	0.2383***	0.0104	0.0083	0.0677***	-0.5637**	0.5892***	0.0489***	4.6900	0.5407***	0.6369***	-0.1000	0.0720	-0.7448**	0.3107*	0.5505***
	(0.000)	(0.441)	(0.000)	(0.000)	(0.000)	(0.000)	(0.080)	(0.323)	(0.000)	(0.004)	(0.000)	(0.000)	(0.084)	(0.000)	(0.000)	(0.636)	(0.136)	(0.001)	(0.017)	(0.000)

p-values in parentheses

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

6 Conclusions

The availability of market-prices and the details on household characteristics as well as owned assets allow us to understand the extent to which quality may be segregated by differences in income and urbanisation levels. Given the disparate consumer baskets for the urban and rural households as well as the disparity in household incomes, the differences in expenditure on quality is hardly surprising - but differences in quality for meat and fresh goods esp for urban households with access to electricity indicates a peculiar environmental difference in areas where urban amenities are available. Whether we use the income measure based on the total expenditure or the measure of assets, the relatively affluent households appear to avail significantly higher quality and diversity in food - particular in the meats-proteins category. Further, the apparent demand for higher quality of meats-proteins - seems a consequence of increasing urbanisation and availability of appliances (as well as storage facilities in urban areas) - a trend that may not be aligned with improvements in nutritional diversity. Under these conditions, a consideration of life-style changes arising out of urbanisation seems essential for the treatment of a specific consumption category as aspirational. One cannot deny that the rise in consumption - when considering food quality which has an evidential role for status in a developing economy such as Tanzania - is influenced by increasing urbanisation in the context of SSA.

This interpretation of empirical results also fits in with the observation on the role of positive externality identified by studies on subjective welfare in Tanzania. More specifically, while poor relative incomes are meant to be strongly related with the general happiness and future outlook of a consumer, the evidence from the LSMS data suggests that the positive externalities from being in a richer area significantly influence the consumer's happiness levels recorded in the survey as well (Atsebi & i Carbonell, 2019). The improvements in food quality indicate a particular channel of how positive externalities from being in richer areas might influence the quality of life and the subjective welfare of the resident consumers.

In summary, while it is difficult to treat the rise in food quality with income as generally aspirational, the effects of both urbanisation and income on quality suggest that such consumption should continue to rise with improvements in household incomes as well as urban development. However, with food as the most basic of needs and one where higher price-based quality seems easy to achieve for the markets, the positive effects of income on quality seem also likely to remain - thus also limiting the future rise in high-quality consumption.

7 Appendices

7.1 Data Preparation Steps

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

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

    (a) All goods 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 goods were also multiplied by 52 (to estimate annual consumption)

    (d) Monthly recall goods 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 goods 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 and extreme 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 goods that are consumed by 10 or less households (goods - 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 goods

(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 goods for consumption. The LSMS survey records the quantities for food goods - even if they were not purchased in the past week. But for other non-food goods (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 goods 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 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.

7.2 Cross-sectional AIDS with adjustment for non-consumption

The use of Hick's commodity theorem for classification of items implies that a substitution between two goods is possible only if they belong to the same commodity group. The functional criteria of separability combined with the Hick's commodity theorem restriction is however a view adopted by the researcher rather than a claim on household's behaviour (see (Nelson, 1991) for substantiation of this argument). To provide a comparison with a disaggregated approach, we provide the provide the results from item-wised view of elasticities suggested by (Heien & Wessells, 1990).

Model

The first-stage of the Heien-Wells (HW) method is the following probit where the dependent variable Y_{ih} is an indicator variable denoting whether the household h consumes the item i from n goods or not. Given s demographic variables d_{jh} for $j \in [1, s]$, prices p_{kh} for $k \in [1, n]$ and household-income m_h, the inverse mills is calculated for the following:

$$Y_{ih} = f(p_{1h}, p_{2h}, ..., p_{nh}, m_h, d_{1h}, d_{2h}, ...d_{sh}) \tag{20}$$

Notice that the inverse mills ratio R_{ih} is calculated for every household using the probit regression (Equation 20)

$$R_{ih} = \begin{cases} \frac{\phi(p_h, m_h, d_h)}{\Phi(p_h, m_h, d_h)} & q_{ih} > 0 \\ \frac{\phi(p_h, m_h, d_h)}{1-\Phi(p_h, m_h, d_h)} & q_{ih} = 0 \end{cases} \tag{21}$$

R_{ih} is then used as an instrument in the following estimation equation

$$w_i = \rho_{io} + \sum_{k=1}^{s} \rho_{ik} d_{kh} + \sum_{j=1}^{n} \gamma_{ij} ln(p_{jh}) + \beta_i ln(\frac{x_h}{P}) + \delta_i R_{ih} \tag{22}$$

Typically one uses the following formulation for $ln(P)$

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

But a simpler linear-AIDS formulation $P = \sum_{i=1}^{n} w_i ln(p_{ih})$ is also often used (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, 1990) 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.

Implementation and Results

The first step with the selection approach is a probit where the dependent variable is an indicator variable that 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 20) . The inverse-mills ratio calculated from the probit (as explained with the Equation 21 - see (Heien & Wessells, 1990)) is then used as an instrument in the AIDS regressions(see Equation 22). Notice that the instrument is calculated for every household and every item - and is part of the additivity constraint that is used for estimation.

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 15, 16 and 17. The own-price and cross-price elasticities presented in the results provide us an overview of the demand across rare and commonly consumed goods.

The main control variables 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), `lntotexp` (log of the total expenditure) and `occupationrank` (occupation rank - an ordered variable derived from the occupation of the head of the household). The other control variables starting with lp- are logarithm of price corresponding to every commodity in the second-stage AIDS regression (see Equation 22). 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 22).

The above item-wise analysis suggests a split between the goods 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 show that the goods available in the industrialised regions of the Tanzania seem to have positive elasticities for income and education rank whereas the goods of agrarian production seem to matter less to the highly educated, middle-age populations huddled in the urban areas. Similarly, how households respond to changes in kerosene and electricity prices indicates the limited availability of industrial goods. A contrast between the goods that are influenced by electricity prices and those that are influenced by kerosene prices demonstrates this urban-rural split as well (demand for goods such as rice, beef is less sensitive to electricity than it is to kerosene). It isn't simply that the households with electricity are more likely to consume the more expensive goods (due to their higher incomes) but that a certain goods become less relevant in the urban settings or are simply unavailable in the rural settings.

Going further, the signs of the coefficients suggest that mangoes (less consumed) and rice (commonly consumed) to be substitutes while fish appears as a complement to cassava-flour ([qfish_seafood]lpcassava_flour <0 etc.) rice. Similarly, fresh milk is a complement to beef, beer, cassava_flour ([qcassava_flour]lpfresh_milk = -0.001) and greens - while it is a substitute to bread and mangoes respectively. The substitution and complementarity thus observed are hardly an indicator of the overall quality since the reported cross-price elasticities do not consider the functional separation. More specifically, one cannot make much out of the cross-elasticities of kerosene with respect to buns or cassava unless without imposing strict separability assumptions (Deaton, 1993) or model restrictions (such as more buns would means less cooking in the household etc.). An aggregated view of the basket is thus preferred given our interest in the choice of quality exercised by the household when faced with multiple varieties of several qualities and price.

Table 15: Estimation for all goods 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 16: Estimation for all goods 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 17: Estimation for all goods 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

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

7.3 Occupation-ranks

It is common for the empirical studies in sub-Saharan Africa to use the income data only indirectly. There are two mains reasons for this. First, the income-data is much more sparse than for occupation or education (only about a quarter of households, for example, have income data available in the LSMS survey for Tanzania). Second, the predominance of informal sectors further subjects the income data to significant measurement errors.

While the income data may not be of direct utility to its sparsity, it is useful in providing an overview of occupational differences in Tanzania. As the household occupation-data is more widely available than the income data, we use a mapping of incomes on to the list of occupational (ordered) ranks by sorting the occupations based on income. More specifically, the occupations are sorted based on the average income received and categorised in broad ranks ranging form 0 to 3 (non-paying occupations are unemployed status is set to 0 and a higher-rank corresponds to a higher-income occupation). The occupation ranks thus assigned are shown in Table 31. The variation in education levels is also given a similar treatment and used as education ranks for comparison. In other words, an ordered variable corresponding to the varied education levels is inferred as occupation-rank based on income received by the highest-education level recorded for individuals in the data. The mappings from education-levels to education ranks are listed in Table 29.

Table 18: Occupations as ranks in Tanzania

code	occupation area	occupation rank	median income	number of entries
14	Student	0	NA	43
13	Job Seeker	0	NA	7
12	Paid Family Work	0	NA	7
16	Unemployed	0	NA	40
11	Unpaid Family Work	0	NA	23
17	Unemployed (too young)	0	NA	0
1	Livestock/Agriculture	1	400000.00	1097
2	Fishing	1	1660800.00	19
3	Mining	1	666666.67	8
4	Tourism	1	46666.67	1
7	Private Sector	2	3000000.00	873
9	Non-Agricultural (w Employer)	2	2662400.00	2
10	Non-Agricultural (w/o Employer)	2	2116800.00	2
8	Non-Government/Religious Org	3	2440000.00	36
5	Government	3	7168000.00	329
6	Parastatal	3	10736000.00	18

Table 19: Education levels as ranks in Tanzania

	education code	education rank
PP	1	1
ADULT	2	2
D1	11	2
D2	12	2
D3	13	2
D4	14	2
D5	15	2
D6	16	2
D7	17	2
D8	18	3
OSC	19	3
MS COURSE	20	3
F1	21	3
F2	22	3
F3	23	3
F4	24	3
O COURSE	25	4
F5	31	4
F6	32	4
A COURSE	33	4
DIPLOMA	34	4
U1	41	4
U2	42	4
U3	43	4
U4	44	4
U5&	45	4

7.4 Estimation Equations

For each of the 10 relevant commodities $\{meatsproteins,\ fat,\ \ cereals\ ,\ \ veg,\ milk,\ starches,\ \ complements,\ tubers,\ fruits,\ \ fish\}$, we have one equation each with budget share as the dependent variable (indicated with equation starting with q such as $qmeatsproteins, qfat$ etc.) and another (for each commodity) with the quality-metric as dependent variable (indicated with equation starting with qV such as $qVmeatsproteins, qVfat$ etc.). The entire list is as follows.

$qmeatsproteins$:

$$\begin{aligned} w_{meatsproteins} = {} & \rho_{1,meatsproteins,\chi} d_{\chi} + \rho_{1,meatsproteins,\xi} d_{\xi} + \gamma_{1,meatsproteins,fat} ln(P_{fat}) \\ & + \gamma_{1,meatsproteins,meatsproteins} ln(P_{meatsproteins}) + \gamma_{1,meatsproteins,cereals} ln(P_{cereals}) \\ & + \gamma_{1,meatsproteins,veg} ln(P_{veg}) + \gamma_{1,meatsproteins,milk} ln(P_{milk}) \\ & + \gamma_{1,meatsproteins,starches} ln(P_{starches}) + \gamma_{1,meatsproteins,complements} ln(P_{complements}) \\ & + \gamma_{1,meatsproteins,tubers} ln(P_{tubers}) + \gamma_{1,meatsproteins,fruits} ln(P_{fruits}) \\ & + \gamma_{1,meatsproteins,fish} ln(P_{fish}) + \beta_{meatsproteins,1} ln(x) + \epsilon_{meatsproteins,1} \end{aligned}$$

$qVmeatsproteins$:

$$\begin{aligned} ln\ V_{meatsproteins} = {} & \rho_{1,meatsproteins,\chi} d_{\chi} + \rho_{1,meatsproteins,\xi} d_{\xi} + \gamma_{1,meatsproteins,fat} ln(P_{fat}) \\ & + \gamma_{1,meatsproteins,meatsproteins} ln(P_{meatsproteins}) + \gamma_{1,meatsproteins,cereals} ln(P_{cereals}) \\ & + \gamma_{1,meatsproteins,veg} ln(P_{veg}) + \gamma_{1,meatsproteins,milk} ln(P_{milk}) \\ & + \gamma_{1,meatsproteins,starches} ln(P_{starches}) + \gamma_{1,meatsproteins,complements} ln(P_{complements}) \\ & + \gamma_{1,meatsproteins,tubers} ln(P_{tubers}) + \gamma_{1,meatsproteins,fruits} ln(P_{fruits}) \\ & + \gamma_{1,meatsproteins,fish} ln(P_{fish}) + \beta_{meatsproteins,1} ln(x) + \epsilon_{meatsproteins,1} \end{aligned}$$

$qfat$:

$$\begin{aligned} w_{fat} = {} & \rho_{1,fat,\chi} d_{\chi} + \rho_{1,fat,\xi} d_{\xi} + \gamma_{1,fat,fat} ln(P_{fat}) \\ & + \gamma_{1,fat,fat} ln(P_{fat}) + \gamma_{1,fat,cereals} ln(P_{cereals}) \\ & + \gamma_{1,fat,veg} ln(P_{veg}) + \gamma_{1,fat,milk} ln(P_{milk}) \\ & + \gamma_{1,fat,starches} ln(P_{starches}) + \gamma_{1,fat,complements} ln(P_{complements}) \\ & + \gamma_{1,fat,tubers} ln(P_{tubers}) + \gamma_{1,fat,fruits} ln(P_{fruits}) \\ & + \gamma_{1,fat,fish} ln(P_{fish}) + \beta_{fat,1} ln(x) + \epsilon_{fat,1} \end{aligned}$$

$qVfat$:

$$ln\, V_{fat} = \rho_{1,fat,\chi} d_\chi + \rho_{1,fat,\xi} d_\xi + \gamma_{1,fat,fat} ln(P_{fat}) + \gamma_{1,fat,fat} ln(P_{fat}) + \gamma_{1,fat,cereals} ln(P_{cereals}) + \gamma_{1,fat,veg} ln(P_{veg}) + \gamma_{1,fat,milk} ln(P_{milk}) + \gamma_{1,fat,starches} ln(P_{starches}) + \gamma_{1,fat,complements} ln(P_{complements}) + \gamma_{1,fat,tubers} ln(P_{tubers}) + \gamma_{1,fat,fruits} ln(P_{fruits}) + \gamma_{1,fat,fish} ln(P_{fish}) + \beta_{fat,1} ln(x) + \epsilon_{fat,1}$$

$qcereals$:

$$w_{cereals} = \rho_{1,cereals,\chi} d_\chi + \rho_{1,cereals,\xi} d_\xi + \gamma_{1,cereals,fat} ln(P_{fat}) + \gamma_{1,cereals,cereals} ln(P_{cereals}) + \gamma_{1,cereals,cereals} ln(P_{cereals}) + \gamma_{1,cereals,veg} ln(P_{veg}) + \gamma_{1,cereals,milk} ln(P_{milk}) + \gamma_{1,cereals,starches} ln(P_{starches}) + \gamma_{1,cereals,complements} ln(P_{complements}) + \gamma_{1,cereals,tubers} ln(P_{tubers}) + \gamma_{1,cereals,fruits} ln(P_{fruits}) + \gamma_{1,cereals,fish} ln(P_{fish}) + \beta_{cereals,1} ln(x) + \epsilon_{cereals,1}$$

$qVcereals$:

$$ln\, V_{cereals} = \rho_{1,cereals,\chi} d_\chi + \rho_{1,cereals,\xi} d_\xi + \gamma_{1,cereals,fat} ln(P_{fat}) + \gamma_{1,cereals,cereals} ln(P_{cereals}) + \gamma_{1,cereals,cereals} ln(P_{cereals}) + \gamma_{1,cereals,veg} ln(P_{veg}) + \gamma_{1,cereals,milk} ln(P_{milk}) + \gamma_{1,cereals,starches} ln(P_{starches}) + \gamma_{1,cereals,complements} ln(P_{complements}) + \gamma_{1,cereals,tubers} ln(P_{tubers}) + \gamma_{1,cereals,fruits} ln(P_{fruits}) + \gamma_{1,cereals,fish} ln(P_{fish}) + \beta_{cereals,1} ln(x) + \epsilon_{cereals,1}$$

$qveg$:

$$w_{veg} = \rho_{1,veg,\chi} d_\chi + \rho_{1,veg,\xi} d_\xi + \gamma_{1,veg,fat} ln(P_{fat}) + \gamma_{1,veg,veg} ln(P_{veg}) + \gamma_{1,veg,cereals} ln(P_{cereals}) + \gamma_{1,veg,veg} ln(P_{veg}) + \gamma_{1,veg,milk} ln(P_{milk}) + \gamma_{1,veg,starches} ln(P_{starches}) + \gamma_{1,veg,complements} ln(P_{complements}) + \gamma_{1,veg,tubers} ln(P_{tubers}) + \gamma_{1,veg,fruits} ln(P_{fruits}) + \gamma_{1,veg,fish} ln(P_{fish}) + \beta_{veg,1} ln(x) + \epsilon_{veg,1}$$

$qVveg$:

$$\begin{aligned} ln\, V_{veg} = \rho_{1,veg,\chi} d_\chi + \rho_{1,veg,\xi} d_\xi + \gamma_{1,veg,fat} ln(P_{fat}) \\ + \gamma_{1,veg,veg} ln(P_{veg}) + \gamma_{1,veg,cereals} ln(P_{cereals}) \\ + \gamma_{1,veg,veg} ln(P_{veg}) + \gamma_{1,veg,milk} ln(P_{milk}) \\ + \gamma_{1,veg,starches} ln(P_{starches}) + \gamma_{1,veg,complements} ln(P_{complements}) \\ + \gamma_{1,veg,tubers} ln(P_{tubers}) + \gamma_{1,veg,fruits} ln(P_{fruits}) \\ + \gamma_{1,veg,fish} ln(P_{fish}) + \beta_{veg,1} ln(x) + \epsilon_{veg,1} \end{aligned}$$

$qmilk$:

$$\begin{aligned} w_{milk} = \rho_{1,milk,\chi} d_\chi + \rho_{1,milk,\xi} d_\xi + \gamma_{1,milk,fat} ln(P_{fat}) \\ + \gamma_{1,milk,milk} ln(P_{milk}) + \gamma_{1,milk,cereals} ln(P_{cereals}) \\ + \gamma_{1,milk,veg} ln(P_{veg}) + \gamma_{1,milk,milk} ln(P_{milk}) \\ + \gamma_{1,milk,starches} ln(P_{starches}) + \gamma_{1,milk,complements} ln(P_{complements}) \\ + \gamma_{1,milk,tubers} ln(P_{tubers}) + \gamma_{1,milk,fruits} ln(P_{fruits}) \\ + \gamma_{1,milk,fish} ln(P_{fish}) + \beta_{milk,1} ln(x) + \epsilon_{milk,1} \end{aligned}$$

$qVmilk$:

$$\begin{aligned} ln\, V_{milk} = \rho_{1,milk,\chi} d_\chi + \rho_{1,milk,\xi} d_\xi + \gamma_{1,milk,fat} ln(P_{fat}) \\ + \gamma_{1,milk,milk} ln(P_{milk}) + \gamma_{1,milk,cereals} ln(P_{cereals}) \\ + \gamma_{1,milk,veg} ln(P_{veg}) + \gamma_{1,milk,milk} ln(P_{milk}) \\ + \gamma_{1,milk,starches} ln(P_{starches}) + \gamma_{1,milk,complements} ln(P_{complements}) \\ + \gamma_{1,milk,tubers} ln(P_{tubers}) + \gamma_{1,milk,fruits} ln(P_{fruits}) \\ + \gamma_{1,milk,fish} ln(P_{fish}) + \beta_{milk,1} ln(x) + \epsilon_{milk,1} \end{aligned}$$

$qstarches$:

$$\begin{aligned} w_{starches} = \rho_{1,starches,\chi} d_\chi + \rho_{1,starches,\xi} d_\xi + \gamma_{1,starches,fat} ln(P_{fat}) \\ + \gamma_{1,starches,starches} ln(P_{starches}) + \gamma_{1,starches,cereals} ln(P_{cereals}) \\ + \gamma_{1,starches,veg} ln(P_{veg}) + \gamma_{1,starches,milk} ln(P_{milk}) \\ + \gamma_{1,starches,starches} ln(P_{starches}) + \gamma_{1,starches,complements} ln(P_{complements}) \\ + \gamma_{1,starches,tubers} ln(P_{tubers}) + \gamma_{1,starches,fruits} ln(P_{fruits}) \\ + \gamma_{1,starches,fish} ln(P_{fish}) + \beta_{starches,1} ln(x) + \epsilon_{starches,1} \end{aligned}$$

$qV starches$:

$$ln\ V_{starches} = \rho_{1,starches,\chi} d_\chi + \rho_{1,starches,\xi} d_\xi + \gamma_{1,starches,fat} ln(P_{fat}) + \gamma_{1,starches,starches} ln(P_{starches}) + \gamma_{1,starches,cereals} ln(P_{cereals}) + \gamma_{1,starches,veg} ln(P_{veg}) + \gamma_{1,starches,milk} ln(P_{milk}) + \gamma_{1,starches,starches} ln(P_{starches}) + \gamma_{1,starches,complements} ln(P_{complements}) + \gamma_{1,starches,tubers} ln(P_{tubers}) + \gamma_{1,starches,fruits} ln(P_{fruits}) + \gamma_{1,starches,fish} ln(P_{fish}) + \beta_{starches,1} ln(x) + \epsilon_{starches,1}$$

$qcomplements$:

$$w_{complements} = \rho_{1,complements,\chi} d_\chi + \rho_{1,complements,\xi} d_\xi + \gamma_{1,complements,fat} ln(P_{fat}) + \gamma_{1,complements,complements} ln(P_{complements}) + \gamma_{1,complements,cereals} ln(P_{cereals}) + \gamma_{1,complements,veg} ln(P_{veg}) + \gamma_{1,complements,milk} ln(P_{milk}) + \gamma_{1,complements,starches} ln(P_{starches}) + \gamma_{1,complements,complements} ln(P_{complements}) + \gamma_{1,complements,tubers} ln(P_{tubers}) + \gamma_{1,complements,fruits} ln(P_{fruits}) + \gamma_{1,complements,fish} ln(P_{fish}) + \beta_{complements,1} ln(x) + \epsilon_{complements,1}$$

$qV complements$:

$$ln\ V_{complements} = \rho_{1,complements,\chi} d_\chi + \rho_{1,complements,\xi} d_\xi + \gamma_{1,complements,fat} ln(P_{fat}) + \gamma_{1,complements,complements} ln(P_{complements}) + \gamma_{1,complements,cereals} ln(P_{cereals}) + \gamma_{1,complements,veg} ln(P_{veg}) + \gamma_{1,complements,milk} ln(P_{milk}) + \gamma_{1,complements,starches} ln(P_{starches}) + \gamma_{1,complements,complements} ln(P_{complements}) + \gamma_{1,complements,tubers} ln(P_{tubers}) + \gamma_{1,complements,fruits} ln(P_{fruits}) + \gamma_{1,complements,fish} ln(P_{fish}) + \beta_{complements,1} ln(x) + \epsilon_{complements,1}$$

$qtubers$:

$$w_{tubers} = \rho_{1,tubers,\chi} d_\chi + \rho_{1,tubers,\xi} d_\xi + \gamma_{1,tubers,fat} ln(P_{fat}) + \gamma_{1,tubers,tubers} ln(P_{tubers}) + \gamma_{1,tubers,cereals} ln(P_{cereals}) + \gamma_{1,tubers,veg} ln(P_{veg}) + \gamma_{1,tubers,milk} ln(P_{milk}) + \gamma_{1,tubers,starches} ln(P_{starches}) + \gamma_{1,tubers,complements} ln(P_{complements}) + \gamma_{1,tubers,tubers} ln(P_{tubers}) + \gamma_{1,tubers,fruits} ln(P_{fruits}) + \gamma_{1,tubers,fish} ln(P_{fish}) + \beta_{tubers,1} ln(x) + \epsilon_{tubers,1}$$

$qVtubers$:

$$\begin{aligned} ln\, V_{tubers} = \rho_{1,tubers,\chi} d_\chi + \rho_{1,tubers,\xi} d_\xi + \gamma_{1,tubers,fat} ln(P_{fat}) \\ + \gamma_{1,tubers,tubers} ln(P_{tubers}) + \gamma_{1,tubers,cereals} ln(P_{cereals}) \\ + \gamma_{1,tubers,veg} ln(P_{veg}) + \gamma_{1,tubers,milk} ln(P_{milk}) \\ + \gamma_{1,tubers,starches} ln(P_{starches}) + \gamma_{1,tubers,complements} ln(P_{complements}) \\ + \gamma_{1,tubers,tubers} ln(P_{tubers}) + \gamma_{1,tubers,fruits} ln(P_{fruits}) \\ + \gamma_{1,tubers,fish} ln(P_{fish}) + \beta_{tubers,1} ln(x) + \epsilon_{tubers,1} \end{aligned}$$

$qfruits$:

$$\begin{aligned} w_{fruits} = \rho_{1,fruits,\chi} d_\chi + \rho_{1,fruits,\xi} d_\xi + \gamma_{1,fruits,fat} ln(P_{fat}) \\ + \gamma_{1,fruits,fruits} ln(P_{fruits}) + \gamma_{1,fruits,cereals} ln(P_{cereals}) \\ + \gamma_{1,fruits,veg} ln(P_{veg}) + \gamma_{1,fruits,milk} ln(P_{milk}) \\ + \gamma_{1,fruits,starches} ln(P_{starches}) + \gamma_{1,fruits,complements} ln(P_{complements}) \\ + \gamma_{1,fruits,tubers} ln(P_{tubers}) + \gamma_{1,fruits,fruits} ln(P_{fruits}) \\ + \gamma_{1,fruits,fish} ln(P_{fish}) + \beta_{fruits,1} ln(x) + \epsilon_{fruits,1} \end{aligned}$$

$qVfruits$:

$$\begin{aligned} ln\, V_{fruits} = \rho_{1,fruits,\chi} d_\chi + \rho_{1,fruits,\xi} d_\xi + \gamma_{1,fruits,fat} ln(P_{fat}) \\ + \gamma_{1,fruits,fruits} ln(P_{fruits}) + \gamma_{1,fruits,cereals} ln(P_{cereals}) \\ + \gamma_{1,fruits,veg} ln(P_{veg}) + \gamma_{1,fruits,milk} ln(P_{milk}) \\ + \gamma_{1,fruits,starches} ln(P_{starches}) + \gamma_{1,fruits,complements} ln(P_{complements}) \\ + \gamma_{1,fruits,tubers} ln(P_{tubers}) + \gamma_{1,fruits,fruits} ln(P_{fruits}) \\ + \gamma_{1,fruits,fish} ln(P_{fish}) + \beta_{fruits,1} ln(x) + \epsilon_{fruits,1} \end{aligned}$$

$qfish$:

$$\begin{aligned} w_{fish} = \rho_{1,fish,\chi} d_\chi + \rho_{1,fish,\xi} d_\xi + \gamma_{1,fish,fat} ln(P_{fat}) \\ + \gamma_{1,fish,fish} ln(P_{fish}) + \gamma_{1,fish,cereals} ln(P_{cereals}) \\ + \gamma_{1,fish,veg} ln(P_{veg}) + \gamma_{1,fish,milk} ln(P_{milk}) \\ + \gamma_{1,fish,starches} ln(P_{starches}) + \gamma_{1,fish,complements} ln(P_{complements}) \\ + \gamma_{1,fish,tubers} ln(P_{tubers}) + \gamma_{1,fish,fruits} ln(P_{fruits}) \\ + \gamma_{1,fish,fish} ln(P_{fish}) + \beta_{fish,1} ln(x) + \epsilon_{fish,1} \end{aligned}$$

$qVfish$:

$$\begin{aligned} ln\ V_{fish} = \rho_{1,fish,\chi} d_\chi + \rho_{1,fish,\xi} d_\xi + \gamma_{1,fish,fat} ln(P_{fat}) \\ + \gamma_{1,fish,fish} ln(P_{fish}) + \gamma_{1,fish,cereals} ln(P_{cereals}) \\ + \gamma_{1,fish,veg} ln(P_{veg}) + \gamma_{1,fish,milk} ln(P_{milk}) \\ + \gamma_{1,fish,starches} ln(P_{starches}) + \gamma_{1,fish,complements} ln(P_{complements}) \\ + \gamma_{1,fish,tubers} ln(P_{tubers}) + \gamma_{1,fish,fruits} ln(P_{fruits}) \\ + \gamma_{1,fish,fish} ln(P_{fish}) + \beta_{fish,1} ln(x) + \epsilon_{fish,1} \end{aligned}$$

The symmetry restrictions corresponding the ten commodities are as follows.

$$\begin{aligned} \gamma_{1,meatsproteins,fat} &= \gamma_{1,fat,meatsproteins} \\ \gamma_{1,meatsproteins,cereals} &= \gamma_{1,cereals,meatsproteins} \\ \gamma_{1,meatsproteins,veg} &= \gamma_{1,veg,meatsproteins} \\ \gamma_{1,meatsproteins,milk} &= \gamma_{1,milk,meatsproteins} \\ \gamma_{1,meatsproteins,starches} &= \gamma_{1,starches,meatsproteins} \\ \gamma_{1,meatsproteins,complements} &= \gamma_{1,complements,meatsproteins} \\ \gamma_{1,meatsproteins,tubers} &= \gamma_{1,tubers,meatsproteins} \\ \gamma_{1,meatsproteins,fruits} &= \gamma_{1,fruits,meatsproteins} \\ \gamma_{1,meatsproteins,fish} &= \gamma_{1,fish,meatsproteins} \end{aligned}$$

$$\begin{aligned} \gamma_{1,fat,cereals} &= \gamma_{1,cereals,fat} \\ \gamma_{1,fat,veg} &= \gamma_{1,veg,fat} \\ \gamma_{1,fat,milk} &= \gamma_{1,milk,fat} \\ \gamma_{1,fat,starches} &= \gamma_{1,starches,fat} \\ \gamma_{1,fat,complements} &= \gamma_{1,complements,fat} \\ \gamma_{1,fat,tubers} &= \gamma_{1,tubers,fat} \\ \gamma_{1,fat,fruits} &= \gamma_{1,fruits,fat} \\ \gamma_{1,fat,fish} &= \gamma_{1,fish,fat} \end{aligned}$$

$$\gamma_{1,cereals,veg} = \gamma_{1,veg,cereals}$$
$$\gamma_{1,cereals,milk} = \gamma_{1,milk,cereals}$$
$$\gamma_{1,cereals,starches} = \gamma_{1,starches,cereals}$$
$$\gamma_{1,cereals,complements} = \gamma_{1,complements,cereals}$$
$$\gamma_{1,cereals,tubers} = \gamma_{1,tubers,cereals}$$
$$\gamma_{1,cereals,fruits} = \gamma_{1,fruits,cereals}$$
$$\gamma_{1,cereals,fish} = \gamma_{1,fish,cereals}$$

$$\gamma_{1,veg,milk} = \gamma_{1,milk,veg}$$
$$\gamma_{1,veg,starches} = \gamma_{1,starches,veg}$$
$$\gamma_{1,veg,complements} = \gamma_{1,complements,veg}$$
$$\gamma_{1,veg,tubers} = \gamma_{1,tubers,veg}$$
$$\gamma_{1,veg,fruits} = \gamma_{1,fruits,veg}$$
$$\gamma_{1,veg,fish} = \gamma_{1,fish,veg}$$
$$\gamma_{1,milk,starches} = \gamma_{1,starches,milk}$$
$$\gamma_{1,milk,complements} = \gamma_{1,complements,milk}$$
$$\gamma_{1,milk,tubers} = \gamma_{1,tubers,milk}$$
$$\gamma_{1,milk,fruits} = \gamma_{1,fruits,milk}$$
$$\gamma_{1,milk,fish} = \gamma_{1,fish,milk}$$

$$\gamma_{1,starches,complements} = \gamma_{1,complements,starches}$$
$$\gamma_{1,starches,tubers} = \gamma_{1,tubers,starches}$$
$$\gamma_{1,starches,fruits} = \gamma_{1,fruits,starches}$$
$$\gamma_{1,starches,fish} = \gamma_{1,fish,starches}$$
$$\gamma_{1,complements,tubers} = \gamma_{1,tubers,complements}$$
$$\gamma_{1,complements,fruits} = \gamma_{1,fruits,complements}$$
$$\gamma_{1,complements,fish} = \gamma_{1,fish,complements}$$
$$\gamma_{1,tubers,fruits} = \gamma_{1,fruits,tubers}$$
$$\gamma_{1,tubers,fish} = \gamma_{1,fish,tubers}$$
$$\gamma_{1,fruits,fish} = \gamma_{1,fish,fruits}$$

Part IV

Permanent Income Effects on Education Expenses in Taznania and Nigeria

Abstract

The economies in sub-Saharan Africa show significant permanent income effects which limit wider access to higher education in the region. Given the status value of education in the region, we examine the extent of limitations on education expenses in the context of aspirational consumption and compare effects of permanent income against regional wealth on education expenses in Tanzania and Nigeria across differing levels of availability of education facilities. The results indicate that the availability of secondary-education facilities and local wealth levels can have a stronger effect than that of permanent income in a typical developing economy and may contribute to a widespread rise in such consumption across the region.

1 Introduction

The participation rates for secondary and higher education remain one of the lowest for the economies of sub-Saharan Africa (SSA). Despite the role of education in social mobility (Galbraith, 1994), the aspirational value of education in SSA may appear to be little - since it remains inaccessible to the vast majority and inadequate in providing mobility to those with lower incomes(Dabalen, Oni, & Adekola, 2001; Oketch, 2016; Ndyali, 2016; Minnis, 2006). Due to the persistent long-term effects of insufficient household wealth (Glewwe & Jacoby, 2004; Ilie & Rose, 2018) and deficient non-formal education initiatives (Minnis, 2006), it may appear that secondary or higher education is likely to remain a luxury for the richer sections of the society in SSA. The presented chapter argues - however - that an extreme variation in urban developments and education attainment levels suggest that wealth effects may be ultimately less important for education expenses than the uneven access to education facilities and disparate urbanisation in the region. So far as education can be considered of aspirational value (often in economies where education attainment has been historically low), the observation also has a significance in the context of aspirational consumption since education expenses would be likely to rise broadly across varied socio-economic levels in response to urban development - becoming a wider common need for the population.

To elaborate this view, we compare education expenses in two economies in the region with respect to the effects of household permanent income against regional wealth levels. Comparing Tanzania and Nigeria - countries that bear similar geographical and demographic metrics while exhibiting significant differences in economic development and their degree of dependence on agrarian production - we find a strong role of availability of education facilities on education expenses and that of local wealth (accumulation of assets). The results of a comparison between slightly more urbanised economy with a relatively less urbanised economy also suggest that a rise in local asset ownership should coincide with reduction of permanent income effects on education over time.

The effects of household permanent income on education - i.e. the extent to which the household permanent income contributes to education expenses - are commonly observed through various mechanisms. Other than through costs (such as transportation) imposed on households when public education access is poorer, the household income also influences access to education due to information asymmetries that may help in future job-searches, better representations in the nascent political institutions of developing economies or certain social characteristics that align with higher household income (Behrman & Knowles, 1999)[23]. Educational inequalities tend to create a poverty trap for the lower-income households (Glewwe & Jacoby, 2004; Behrman & Knowles, 1999) - a concern that has been

[23]The core issue evaluated by Behrman and Knowles (1999) is that family dynasties are reinforced when children from higher-income households are significantly more likely to receive better schooling. Taking a detailed view of investments in education and the variation in quality, Behrman and Knowles (1999) note systematic associations between children's school progress and household income through genetic endowments, additional support etc..

important for policy in SSA (K'akumu & Olima, 2007; Owusu & Agyei-Mensah, 2011; Minnis, 2006).

At the same time, the uneven urbanisation also requires us to consider the extreme variation in local wealth (accumulated assets in a locality) for the context of aspirational consumption in the region. While household wealth is often reported to be more important than physical access for education expenses (Glewwe & Jacoby, 2004; Behrman & Knowles, 1999), the extreme regional variation in wealth and the peer-effects that they may imply must also be factored in when considering rises in per-capita expenditure. Despite the limitations posed by wealth and permanent income, the overall effect of urban development could be that of a broad overall increase in education expenses. A few idiosyncratic reasons may contribute to such effects as well - for example, many countries in SSA have heeded the policy recommendation in the last decades to prefer wider access to primary education over selective provision of secondary-level education (Heyneman, 1980b; Jee-Peng, Jimenez, & Psacharopoulos, 1986; Heyneman, 2003; Psacharopoulos & Patrinos, 2004; Heyneman & Lee, 2014) - thus letting rapid urban developments create a high demand for education.

Setting aside the physical access to education - i.e. the higher availability of education opportunities in urban areas which we consider as a separate control - the two fundamental mechanisms how local wealth may influence education expenses are the relative abundance of opportunities for higher income employment and the social environment of the urban areas. The better opportunities of employment and wealth overall make education more affordable so that the households in rural areas - where incomes are invariably lower and education somewhat of a luxury- end up underinvesting in education. The social environment of urban areas is also often different - leading to a stronger need for finding education-dependent employment in more economically developed or high-asset areas.

The current study measures effects of permanent income - the first explanatory variable - as the annual total household expenditure. It uses the average assets in the immediate area surrounding the household as a measure of local wealth i.e. the second explanatory variable . The main goal of using local wealth - rather than urban-rural indicators - as an explanatory variable is to take into consideration the spatial variation of owned assets in the SSA without relying too much on the urban-rural dichotomy that have been recently criticised in the literature (Abay, Asnake, Ayalew, Chamberlin, & Sumberg, 2021; Wiggins & Proctor, 2001). Despite several empirical studies using urban-rural effects to explain disparities in education expenditures in SSA (Ebaidalla, 2018; Donkoh & Amikuzuno, 2011) as well as in other developing economies (Bayar & İlhan, 2016; Knight & Shi, 1996; Acar, Günalp, & Cilasun, 2016), we thus attempt to consider a spatial distribution of resources and employment opportunities in the SSA rather than the broad or implicit urban-rural classifications used in surveys. The measurement for the local wealth uses the average assets-accumulation[24] in the

[24]The term "assets" signifies the goods that are owned by the household. These constitute items that are more often referred to as durable goods in the consumption literature from the developed economies. There seem at least two reasons why the developing economics literature more often refers to the durable household possessions

district where the household is located[25]. With permanent income effects and local wealth effects as the main explanatory variables whose effects are compared after controlling for physical access (availability of education opportunities) and social factors (religious identity etc.), the study finds that the household permanent income effects for education to be more severe in less urbanised areas. To compare this assessment with that using the more conventional approach using urban-rural indicators, we also provide estimation results from an alternative formulation where a ruralness-score of the district based on urban-rural distinction is used instead of the local wealth measure and the permanent income measure is replaced with the direct measure of household assets value (see Section 5.3 for details).

It is worth highlighting the importance of physical access as a control - a factor that directly influences education expenditure in the above analysis. Given the focus on secondary education in development policy as well as literature in the last decades (Fuller, 1976; Tilak, 2002, 2007; Heyneman, 2003, 1980b), we focus on the role of accessibility of secondary education as a control in the two economies. The results we obtain with local wealth levels are also particularly significant since the effects of income are more often reported to be stronger than the effects of physical access in empirical studies from the region (Obasuyi & Rasiah, 2019; Filmer & Pritchett, 1999, 2001; Minnis, 2006; Behrman & Knowles, 1999; Glewwe & Jacoby, 2004; Kambon & Busby, 2000). As such, if access to facilities and local wealth both limit education expenses more significantly, then more urban development could have a strong impact on how education expenses rise - a result that is relevant for how aspirational needs could rise in the economy.

The social factors associated with a household are another important set of controls that one cannot ignore as factors influencing participation in education. While the effects of discernible factors such as segregation or exclusion based on social identities education have been often highlighted in SSA (K'akumu & Olima, 2007; Owusu & Agyei-Mensah, 2011), the specific role of social factors that may encourage or support education present a more challenging task to the empirical studies. The early attempts to explain role of social factors through the notion of social capital (Bourdieu, 2018) - encompassing habits, preferences and behaviour cultivated in richer environments - have been rejected by the empirical studies in favour of certain environmental factors - such as the feedback for improvement of students, the subjective evaluations of students or other measures of social interactions instead (Manski, 2000; Sobel, 2002). The current study thus focuses on specific social factors at the household level - encompassing the acquisition of skills from parents or from the social environment (networks) valued in the employment markets and higher levels of education. The parental educational levels - which are widely observed to have a significant effect on

as assets. First, the financial assets are far less relevant in the developing economies. Second, the durable goods such as bicycles or stoves - however insignificant - are far more important for consumption behaviour in the developing world.

[25]The long-term asset ownership in a 6-km area is considered for the local wealth measure. The 6-km scale in the district where the household residence is located - is based on the geographical distances between districts and the average distance to the education facility recorded in the consumption survey.

education attainment in the developing economies (Tansel, 1997) - are the first important social factor we consider in the chapter[26]. The other relevant social factor that may favour participation in education is the main occupation in the household. More particularly, the agrarian or non-agrarian nature of the household occupation in a largely agrarian economy may indicate a non-transactional advantage for individuals through the social networks they acquire. Other household characteristics of social significance that influence participation in education vary across the economies. In Tanzania, for example, the speakers of English language seem to be generally more educated and have higher incomes. The ability to speak English could indicate an exposure to education in ways that the poorer or rural households may not be exposed. The religious identities also have some implications for participation in education in countries such as Nigeria (Cooray & Potrafke, 2011; Alesina, Hohmann, Michalopoulos, & Papaioannou, 2019).

An essential claim of the human capital theory (Tilak, 2002) - which the empirical literature on determinants of education expenditure often relies on - is that the educational investments must change over the life cycle of consumers. The current chapter takes advantage of the reporting of household expenditures on education at individual level in the surveys for both the countries and considers how the education expenditure may vary with the age of the children in the household. The variables for childrens' ages are important for our comparison also because the education expenditures for a population where the majority joins the workforce after attaining primary education would be very different from one where the employment opportunities are higher for educated individuals. Since large households have higher basic needs than smaller households, we also consider the number of children as another relevant control that is closely related with lifecycle concerns.

As mentioned earlier, the two economies in SSA that we have selected for our comparison of education expenditure have different levels of economic development and a different degree of dependence on agrarian output. More particularly, Nigeria relies more heavily on energy exports while Tanzania's economy is more reliant on its agricultural output - a difference that is critical for the varied level of urbanisation in the two economies. While the educational attainments in the populations of the two countries are evidently shaped by the idiosyncrasies of the policy and distinct social factors, the varying levels of urban developments in the two economies have general implications for how demand towards education may rise in the region.

To summarise the methodology, household permanent income and local wealth levels are the main explanatory variables in our comparison of educational expenditures while physical access to education, the social or demographic factors and the household characteristics serve as important controls. The first explanatory variable i.e. the permanent income of the

[26]It is possible to argue from a human capital perspective (Wu, Zhang, & Zhang, 2008; Behrman & Knowles, 1999; Glewwe & Jacoby, 2004; Spaull, 2013) that as accumulations transferred to further generations (Becker & Tomes, 1979), the parental education levels have an effect similar to that of wealth (which has a strong effect on educational expenditures).

household uses the logarithm of household's total annual expenditure - a measure from standard consumption theory (Modigliani & Brumberg, 1954; Friedman, 1957) that is commonly used in studies on consumption. The local wealth levels - used as the second explanatory variable - is based on average asset-ownership levels in the households' vicinity (district). The physical access to education - is measured with the number of secondary schools in every district so that a higher score of `secondary_schools` for a district implies more secondary schools in the district. Further robustness checks are conducted by redefining the local wealth measures which vary the boundaries of households vicinity and examine the significance of local wealth in the household's neighborhood of similar occupations (agrarian/non-agrarian) and education ranks (primary, secondary).

In the alternative formulation - which we provide only in order to compare results from the main formulation with a more conventional approach using urban-rural demarcations from the survey - the total value of long-term assets owned by the household i.e. the measure of household wealth is used as the first explanatory variable and a "ruralness" score based on urban-rural indicators from the survey is used as the second explanatory variable. The use of wealth-indices based on the stock of durable goods (excluding short-term durable goods that are not transferred over a generation) is also common in the studies from health and household economics (Howe et al., 2008, 2010; McKenzie, 2005; Booysen et al., 2008). The results from the alternative formulation also support our inference that rising local wealth (measured with conventional measures using urban-rural indicators in the survey) may coincide with weakened severity of household permanent income effects.

Our main conclusion in the study is that so long as education can be considered of aspirational significance, the lower severity of permanent income effects with higher local wealth means that expenditure on education is likely to experience a significant rise as a common need across all sections of a society rather than being limited to a smaller richer section of the society. The conclusion is drawn, however, with the caveat that weakening of household permanent income effects must also overlap with improvements in employment opportunities and social equity. The current study finds that the segregation of education opportunities and the disparities in future employment opportunities across the country are more evident in Tanzania where urban-differences seem more extreme.

The current chapter contributes to the literature on aspirational consumption by highlighting the role of mobility perceptions amidst urban development in the developing economies. While education expenses are not commonly treated as aspirational, their status value as well as their value in future economic gains makes education an important case-study for how mobility-related expenses change under income inequality and economic growth. The chapter also demonstrates that despite the challenges of urban overpopulation and lifestyle changes which entail rise in aspirational consumption, the overall effect of urban and economic development may be a desirable one for the education levels of the population. The weakening in severity of permanent income effects with the rise in

economic development could in fact indicate pathways to future improvements in income equality - changes that encompass a reduction in education inequalities through development of the industrial sector and rise in demand for education. In the context of aspirational consumption, all such expenditures that grant mobility opportunities to the consumer seem likely to continue to increase as a mass common need in the economies of SSA.

In the sections that follow, the Section 2 surveys the literature on education expenditures, Section 3 details the method used in empirical analysis, Section 4 describes the data used for Nigeria and Tanzania in the study and Sections 5, 6 discuss the results from the empirical analyses.

2 Literature Survey

> There is no well educated literate population that is poor, [and] there is no illiterate population that is other than poor.
>
> *Beyond contentment*, John Kenneth Galbraith, (1994)

While education may not be considered a basic necessity in all poverty-measurement surveys, it is seldom disputed that a lack of education contributes to poverty. Even in the developing economies, the poorest individuals tend to have a strong demand for education (Tilak, 2002; Gertler & Glewwe, 1990). The literature goes as far to suggest that when education is not state-funded in the developing economies, there might be a compulsion - rather than a mere willingness - to pay for education among the poor (Tilak, 2002; Tilak et al., 2002). With a high demand for educated workforce in an industrialising economy, there seems little disagreement in policy therefore - whether from a human capital perspective (Barro, 1991, 1999) or the capability perspective (Sen, 1992, 1978; Nussbaum, 2003) - that economic growth bears a strong relationship with educational attainment.

An important consideration for education attainment in the population is that while higher education provides needed skills for higher incomes, the expenditure on education is often limited by the wealth one possesses. The differences in education elasticities based on income have in fact been emphasised in the literature early on since the mid-last century (Benson, 1961; Hashimoto & Heath, 1995). The empirical studies from the SSA also report significant wealth effects on education (Glewwe & Jacoby, 2004). Inspecting the enrollments data, for example, Chernichovsky (1985) note that the children from households that have large cattle stocks face less reductions in enrollment, because higher wealth may allow the households to take care of the cattle and prevent the childrens' withdrawals from schools[27]. Using a benefit incidence analysis to test if the education is pro-rich, Ilie and Rose (2018) show that the poor

[27]One caveat with the use of education enrollment data is that they are often not considered rigorous for empirical assessments - as they ignore the difference in educational levels and measure "flow" rather than "stock" with respect to educational inequalities (Obasuyi & Rasiah, 2019; Thomas, Wang, & Fan, 1999).

young people receive only 5% of funds allocated to tertiary education compared with 54% for the rich. The wealth effects on education have had somewhat of a historic significance and have informed views on policy as well as notions of necessity.

While the effects of permanent income remain significant in SSA, much of the discussion on the high graduate unemployment and low higher-education participation has focused on the poor returns to education in the region. On one hand, the literature attributes these problems to the prevalent mismanagement in public education (Languille, 2019) and on the other it highlights a misplaced focus on formal education in the informal economy settings (Minnis, 2006). Instead of focusing on the returns-to-education, we argue that the poor returns may not offer a complete view of lower participation in education. One reason is that the education needs or the perceived value of education is not be adequately captured with a returns-to-education approach. As Heyneman (2003) point out, the returns-to-education approaches tend to ignore the positive externalities from education[28]. The other reason is simply that the role of physical access and wealth distribution may be far more dominant than the issue of poor returns-to-education in an economy. While there should be little doubt on that the concerns for return-to-education are important, they do need to be viewed in the context of effects of permanent income and availability of education facilities. More specifically, it is possible that the demand for education is suppressed due to inadequate income (relatively high prices of education) in the economy, structural regional differences or a sparse distribution of education opportunities rather than due to lower pay from post-education employment opportunities alone. If urbanisation trends improve such regional inequalities, then a rise in education expenditure as a common need for the masses seems evident[29]. A more holistic view towards education expenses must therefore consider wider changes in the economic environment of developing economies.

Although empirical studies have generally found physical access to education to be less significant than wealth effects in SSA (Filmer & Pritchett, 1999; Kambon & Busby, 2000; Obasuyi & Rasiah, 2019), there is also evidence that the effects of physical access become stronger at higher (secondary and tertiary) levels of education (Ngware, Onsomu, Muthaka, & Manda, 2006). The effect of urban-rural differences - which may be correlated with difference in both physical access and local wealth - has in fact been widely reported to be significant in SSA (Ilie & Rose, 2018; Castro-Leal, Dayton, Demery, & Mehra, 1999;

[28]In a debate on whether the policy should favour free basic education to all (at the point of use) over the investments in higher education or not, (Heyneman, 2003, 1980a) argue that the positive externalities from education are simply not factored in the productivity measures used in the returns-to-education approach proposed by Jee-Peng et al. (1986). Heyneman (1980b) elaborate that the individuals benefit culturally and physically from education and that such changes are ignored when using a returns-of-education approach. It would seem also that not only does formal education help workers pick up more skills on their own (Fuller, 1976), the skills from secondary or higher education are often necessary in the labour market (Tilak, 2007).

[29]Such trends have implications for education policy as well - if education expenditures are better explained with physical access, then the policy may need to focus on building more accessible schools whereas if permanent income effects are more significant, then there is a need to make education more affordable for the lower-income households.

Glewwe & Jacoby, 2004; Bossuroy & Cogneau, 2013). The findings from Porzio et al. (2021) also suggest that the changes in social environment of the urbanised areas may create a more competitive setting over time and across the wider populations. Both local wealth and household permanent income (aside from the physical access to education that we use as a control) are thus important concerns in the current study that motivate our elaboration of the decline in effects of permanent income with rise in economic development.

The effects of social factors on education - particularly the link between cultural capital and education - are also of significance in a comparison of education expenses. Social factors for education have been of interest in the sociological literature at least since the beginning of the post-war era (Bourdieu, 2018). While this early research had viewed education as a certain high-culture participation that enhances the social capital of a household, the later refinements have focused on the specific role of social networks and parental acquisitions or skills (Byun, 2007; Yamamoto & Brinton, 2010; De Graaf, De Graaf, & Kraaykamp, 2000; Yan, Peng, Hao, Irfan, & Wu, 2021; Katsillis & Rubinson, 1990; Alesina et al., 2019; Bossuroy & Cogneau, 2013). De Graaf et al. (2000) find - for example - that it is the reading habits among the richer households rather than high-culture-participation that affects school performance in Netherlands. More importantly, as De Graaf et al. (2000) argue, it is possible to explain a significant part of parental social background with explicit measures of resources accessible to the students. In another study, Gould, Simhon, and Weinberg (2020) point out that higher educated parents invariably spend more time with children - a social factor that contributes to how parental education levels influence the next generation. In fact, the high-culture participation does not always influence educational performance positively either. As Byun (2007) find, the high-culture participation affects educational performance negatively in South Korea as it takes time off from study. In most cases, the cultural factors thus seem to matter in education only through economic factors that are relevant for participation in education. Recent literature is thus more often concerned with a "network social capital" - focusing on the role of social networks in consumer decisions instead (Mouw, 2006; Manski, 2000). The scope of social effects in the current study is therefore limited to that of urbanisation and other household (religious identity, language spoken) or paternal characteristics (paternal education-level, occupation etc.).

The general trends of rise in aspirational consumption across base-of-pyramid (BoP) countries that have been recently highlighted (Srivastava et al., 2020) do require us to consider the specific structural issues faced by the respective economies. The hopes of a rise in household expenditure on education - or any other category of consumption that may be considered aspirational - can hardly be detached from the macroeconomic issues faced by economies in SSA. In the specific context of SSA, the high-land labour ratios in the region(E. Anderson & McKay, 2004)) - which characterise strains on long-term economic growth while contributing to inequality, migration-pressures (Kosec, Ghebru, Holtemeyer,

Mueller, & Schmidt, 2018) and employment opportunities in the region - offer unique challenges to rise in household budgets. Similarly, the high susceptibility of the region to volatilities in commodity prices affects household budgets significantly (Deaton, 1999) and the presence of a large informal sector constrains funding for "non-necessities" such as education. Finally, the political powers availed by the lower-income workforce - which an underdeveloped and commodity reliant industrial sector (Kaplinsky & Morris, 2016) has struggled to absorb (Geda, 2019; Arbache & Page, 2009; Clementi, Fabiani, & Molini, 2019; Minnis, 2006) - also remain relatively little and have implications for income inequalities as well as the education sector. A trend of rise in consumption or expenditure in the region is thus subject to an interplay of such factors in the long-term.

As many such factors are shared by both Nigeria and Tanzania, the current study focuses only on the contexts that differentiate the two economies in the current chapter. In particular, we note that Nigeria has had more investments in commodities sector (oil and petroleum) than Tanzania - which remains a relatively agrarian economy. Additionally, the policy of education has also been different in the two economies - with Tanzania having a history of restrictions of secondary education at a time when other countries in the region were expanding their education sector (Söderbom, Teal, Wambugu, & Kahyarara, 2006; Appleton, Bigsten, & Manda, 1999; Buchert, 1992) . These differences as well as the disparities in permanent income, urbanisation and access to secondary education are reflected in the levels of educational attainment that we detail in the subsequent sections.

3 Econometric Method

The goal of our empirical analysis is to find out if with rise in urban and economic development, the permanent income effects better explain the disparity in education expenses rather than local wealth in the two economies. As mentioned earlier, our main formulation is based on the local wealth measure - whereas results are also provided with an alternative formulation based on urban-rural indicators for comparison. In the main formulation, the logarithm of total expenditure is used as the main explanatory variable while the local wealth measure - the second explanatory variable - is calculated as the average assets in the consumer's district (for smaller districts, the average over multiple districts within a 6-km boundary is taken). In the alternative formulation using urban-rural indicators, the two explanatory variables are the household wealth based on durable goods from the survey and a ruralness score based on urban-rural indicators. The measures for local-wealth (as well as the first explanatory variable in the alternative formulation) rely on the market-value of durable goods (including land, housing) that are reported in the survey (see Section 4 for details). It is worth highlighting that the durable goods are considered as part of the consumer wealth only if they are transferable over generations. The ownership of goods that won't last more than a lifetime enter as consumption for the household rather than

as assets contributing to household wealth.

In a simple econometric model, the representative household balances i) her non-durable consumption (excluding education) with ii) increasing expenditure on education for the children. The non-durable consumption that the household may spend on excludes the expenses on maintenance of owned assets. The educational expenses and non-durable consumption excluding assets-related-costs thus form the consumer budget in the model. In this model for cross-sectional comparison of education expenses, we do not consider the substitution with savings towards durable goods. Other than our focus on cross-sectional effects, another reason why we avoid panel methods in the current study is a severe measurement error with time-varying household asset values in the survey (see Section 4 for details of asset values recorded in the consumption microdata). More specifically, the variation in time of ownership of assets is not accurate enough for us to measure the changing value of household assets possessions - due to the inconsistent recalls of the number of assets over time for the same household.

It is worth highlighting that the econometric analysis is limited for households having one or more children. In other words, only households with at least one member with their age less than 18 are considered in the analysis. Since the households with lower working age are likely to have less accumulations as well as needs, the average paternal age in the household is also used as a control in the comparison. It is worth pointing out that the young households with no children in school or those where adult members are engaged in continuing education while not having any children are excluded from the analysis. The exclusion would be a more serious issue if one were to investigate factors determining whether children enter education or not but as our focus remains on the *ceteris paribus* effects of local wealth on household education expenses, the selection bias thus introduced remains of less concern for our analyses. The number of children (household-size), the educational level of the parents, the accessibility of schools in the district and the social characteristics of the household such as religion and language spoken are other important controls used in the analyses. Notice that for the paternal educational levels, we use an educational rank that is standardised by differentiating between primary, secondary and higher educational levels (see Section 4 for details).

Given the recent industrialisation in the sub-Saharan Africa, the dichotomy of agrarian and non-agrarian occupations is of particular importance to our analyses and is used as a control in the analysis. However, we are also aware of not mixing up the peri-urban with the rest of the rural settings in the economy (Wiggins & Proctor, 2001) and thus focus on the comparison of the educational expenses across local wealth levels in our analysis.

The analysis uses two formulations with the budget-share of education expenses and logarithm of education expenditure as dependent variables. Such an approach - using both budget-share and logarithm of education expenditure as dependent variables - is similar to the empirical method used by Acar et al. (2016) - where the budget shares offer a better view of the household-level demand while the logarithm of expenditure facilitate an appropriate

comparison of wide-differences in educational expenses across the economy.

The first regression uses the budget share of education expenses w_{educ} as the dependent variable

$$w_{educ} = \rho + \alpha_{educ} \cdot r + \beta_{educ} \cdot log(x_i) + \gamma_{educ} \cdot log(p_{ne}) + \delta_{educ} \cdot s + \epsilon_{ne} \tag{23}$$

The main explanatory variables used in the empirical analysis are the logarithm of the total annual expenditure $log(x_i)$ (for a household i in a given year) and the local wealth measure r. The demographic and social identity variables ρ (e.g., age of the children, the family size and other variables listed in Table 21) , the access to education s, the indicator variable specifying if there are household members with an age appropriate for primary, secondary or tertiary education and the prices p_{ne} associated with non-durable expenditure in the region (calculated as unit-value i.e. the average non-durable expenditure excluding asset-related costs in the district) are all used as controls in the analysis. In above Equation 23, ρ is the intercept, α_{educ} is the coefficient for local wealth, β_{educ} is the coefficient for $log(x_i)$, γ_{educ} is the coefficient for logarithm of prices p_{ne} and δ_{educ} is the set of coefficients for physical access s. The variables used in the results are listed in Table 21. Notice that the total expenditure x_i appearing both as an explanatory variable and as the denominator of the education budget-share dependent variable may lead to a measurement error in an elasticities-interpretation of education expenses. The issue however is of less significance as we only wish to examine the relative importance of local wealth level with respect to household permanent income.

In the alternative formulation provided for a comparison with the analysis based on urban-rural indicators, we use the logarithm of total assets owned ($log(A)$) and a ruralness score based on urban-rural indicators as the explanatory variables. The second regression thus uses the logarithm of educational expenses $log(x_{educ})$ as the dependent variable against the same set of control variables (instead of w_{educ} in Equation 23)

$$log(x_{educ}) = \rho + \alpha_{educ} \cdot r + \beta_{educ} \cdot log(x_i) + \gamma_{educ} \cdot log(p_{ne}) + \delta_{educ} \cdot s + \epsilon_{ne} \tag{24}$$

While the appearance of educational expenses both in the dependent variable and in the logarithm of the total expenditure $log(x)$ (used as the explanatory variable) may cause some endogeneity, it does not interfere with the comparison of relative effects of r and $log(x)$ in the model. The estimation results from the comparison of education expenditure for Tanzania and Nigeria - which uses a tobit (due to zero expenditure on education for a significant number of households) - are presented in Section 5.3 after a discussion of the descriptive statistics in Section 5.2.

4 Data

The educational expenses and the non-asset related expenditures are obtained from the Living Standard Measurement Study (LSMS) which is conducted by the World Bank. The survey is unique in its and aims to improves the understanding of links between asset ownership, community access (to market, schools), food prices and agriculture activities. The data in the surveys from multiple countries is largely uniform in its data-model (fields) and is often geo-referenced with a sufficiently wide-coverage in the country to allow a study of spatial variation in the observed variables. The years used for the cross-section analyses are the most recent years in the survey for Tanzania (2014) and Nigeria (2015) - i.e. the last waves available in the survey at the time of writing this chapter.

As described earlier, the local wealth measure uses the log-averages of the values of durable goods accumulated in a district (calculated as per the specifications in Table 26). The durable goods owned by every consumer available in the surveys range from the short-term durable goods such as mobile phones to the expensive and long-term assets such as land or houses. Given our interest in the wealth that excludes consumption that is not inherited, we focus on durable goods that can be transferred over a generation. For both Tanzania and Nigeria, we therefore include items as long-term assets only if they are costlier than furniture - which is the cheapest asset to be transferred between the generations in households. That the assets costlier than furniture are transferred over a generation is readily verified in the data by inspecting households that are split when a young member of the unsplit family starts a new household.

The details of the records of durable goods in the LSMS data include the number n_t of durable goods owned by the household, the reported cost C_t at which the durable good was purchased and the reported price P_t which the household expects by selling the durable good in the current market at the time of survey. We use the reported price P_t of the durable goods to infer the goods-values in the area surrounding the consumer - partly because P_t is not susceptible to errors associated with recall of the purchase value (which often varies for the same item over years in the panel data for a household). Further, P_t encapsulates the perceived depreciation of the durable good over time - a feature that makes it more appropriate for a comparison against non-durable consumption at a given year t. The total cost of the durable goods owned by the household is thus obtained simply as a product of the number of assets n_t reported by the consumer with price P_t. These costs are used both for the local-wealth measures and the assets-measure. As explained earlier, the permanent-income measure uses the logarithm of total expenditure at the household level.

At the time of writing this chapter, the precise geographic-location of every household (or the ward she belongs) to is not available to us in the LSMS data for Tanzania. As a result, we have relied only on the geo-coding of districts for the wider level consumer vicinity based on Euclidean distance. Partly to avoid the absence of rural-urban fields corresponding to the wards in year 2014 for Tanzania, we do not use the rural field for analysis at the wider-district

level aggregation and use a variable indicating how rural the consumer's vicinity is by counting the number of rural wards in the 6 km vicinity (or the district boundary).

The physical access to education is measured using the distance to the nearest secondary school. More specifically, a score `secondary_schools` for each district represents the number of wards that have a secondary school within the 6-km area. A district with a higher score has more wards that have a secondary school in close vicinity (within 6-km vicinity). This measure takes care of the issue that there are many wards (and enumerated-areas within) having consumer entries for assets and consumption that the school-availability (primary as well as secondary) is not recorded for. The aggregation over districts takes care of this issues there are no district with consumer entries for assets and consumption that we don't have the total number (score) of secondary-schools for. This score evidently also factors in the high-population density in urban areas since densely populated districts would have more wards within the same geographical distance as a less populated district.

The next important control is the occupation of the households which is directly available for every individual in the survey. Whether an occupational sector is agriculture-related or not determined is based on the main occupation of household head (HH). The number of children in the household, the ages of the children (and other members) are directly observed in the data. Since the typical household unit in Tanzania and Nigeria may not be a nuclear family, the paternal educational level - which is another important control in the study - cannot simply be the education level of the household head(HH). This is because unlike for nuclear families, the individuals on whom the household's reported education expenses are spent, could be the grandchildren (rather than children) of the joint-family household-head. To avoid letting the household's paternal education level point to the generation before the parents of children for certain joint-family households - where the HH has his/her grandchildren attending school - while letting the same variable point to the parents of children in school for nuclear-families (where HH has his/her children attending school), we rely on a more uniform measure for a household's highest education rank `highesteduc` as the highest education in the generation of the parents of the children in household in the survey [30]. While the maternal education level is often a more appropriate control for effects of parental education level on whether the child enters education or not and more generally on the individual education expenses for a child, our focus on the cross-sectional variation of education expenditure aggregated at the household-level offers a preference to the father's education rank solely because other characteristics of the household such as age and occupation are linked with the household heads who are invariably male in the survey. In other words, we do not offer a preference to mother's education level over father's education level as the paternal education control in the analysis only due to normalisation of household characteristics at the household-head level.

To provide a uniform treatment to educational levels across both Tanzania and Nigeria, we

[30]Notice that only relations with the household head are reported in the survey. Thus, there is no direct way of locating the parent of a child except through the relation with the household-head.

use 4 education categories - setting the education rank to 1 for all pre-primary levels, to rank 2 for until the completion of primary education, to rank 3 for up to the secondary education level and to 4 for all higher education levels (higher-secondary, university etc.). These mappings are provided in Tables 29 and 30 (see Appendix 7.1). For instance, the primary school leaving exam is taken at D7 in Tanzania and corresponds to the education rank 2 ; all standards above D7 and before O course (Diploma) correspond to the education rank 3 and all education levels at or above O course correspond to education rank 4 for Tanzania. Similarly, the levels above P6 - which is the highest primary education level in Nigeria - implies an education rank 3 (corresponding to P6 in Tanzania) and all education levels at or above the technical training level (treated equivalent to Diploma level in Tanzania) are correspond to education rank 4. As access to private education may be significant in the urban-areas, we also use a control for private-education that is available in the survey.

The languages spoken (`has_english` in Tanzania) and the religious-identity for household members (`religion` in Nigeria) - used as controls - are recorded at the individual's level in the survey data for the countries and are directly used as controls in the survey. To control for the price of consumption in the household's district, we also use the non-durable consumption averages `log_mean_cost_ne` (excluding the costs associated with ownership of assets in the survey). The measurement `log_mean_cost_ne` of price is akin to the price-deflator (unit-value) methods used in the demand literature (Deaton, 1988; Tafere et al., 2010). The rationale for excluding costs associated with assets is that the costs of maintaining assets (fees, repair costs, usage costs etc.) are not levied on those who don't own the related assets and should not be included in the general price of non-durable consumption.

Dependent Variable	Variable Name	Description
w_{educ}	`w_educ`	budget share of the educational expenses
$log(x_{educ})$	`log_educ`	logarithm of the educational expenses

Table 20: Dependent Variables

Controls / Explanatory Variables	Variable Name	Description	Ind/Ref Level
$log(x_t)$	ln_tot_exp	logarithm of the sum of total expenditure by the household	Individual
$log(A_t)$	ln_tot_assets	logarithm of the sum of total durable goods owned by the household	Individual
r	localwealth	logarithm of mean asset-ownership level in the 6-km radius area around the household district	Area
r_Γ	localwealth_educ	logarithm of mean asset-ownership level of consumers in the 6-km radius area having primary education level (or higher than primary) if the household has a primary (or higher than primary) education level	Area
r_Ω	localwealth_agri	logarithm of mean asset-ownership level among the consumers of having the same value of the indicator variable denoting agrarian /non-agrarian occupation (i.e. the variable Ω) in the 6-km radius area around the household district	Area
p_{ne}	log_mean_cost_ne	logarithm of the mean non-durable expenditure - excluding asset costs in the area surrounding the consumer	Area
γ	rural_wards	the percentage of rural wards/eas in the district of consumer's residence	Area
s	secondary_schools	the number of wards in a district with an accessible secondary school (within 6-km)	Area
π	educpriv	factor variable indicating if any of the child attends a privately owned educational institution	Individual
μ_1	is_primaryage	factor variable indicating if there is a family member in the household of age between the mean primary school attendance ages (i.e. between 7 and 13 inclusive)	Individual
μ_2	is_secondaryage	factor variable indicating if there is a family member in the household of age between the mean secondary school attendance ages (i.e. between 15 and 23 inclusive)	Individual
μ_3	is_tertiaryage	factor variable indicating if there is a family member in the household of age between the mean secondary school attendance ages (i.e. between 23 and 31 inclusive)	Individual
Ω	agri	factor variable indicating whether the household head's occupation is agrarian or not	Individual
Γ^*	highesteduc	highest education rank in the paternal generation of the children considered in the household	Individual
h	numchild	number of children (≤ 18 years age) in the household	Individual
χ	religion	the religion which most members in the household subscribe to	Individual
ξ	has_english	ability to speak or write English	Individual

Table 21: Control Variables

5 Empirical Analysis and Results

5.1 The context for Education in Tanzania and Nigeria

Despite similarities in geographic areas, age-distribution of populations, public infrastructure and use of English as the primary administrative language, the economies of the two countries on Eastern and Western coasts of Africa exhibit quite a few structural differences - owing primarily to levels of industrialisation and endowments in natural resources. Tanzania is spread over nearly 885,800 sq km of land while Nigeria sprawls over 910,768 sq km of total land. The age-distribution of the populations and the dependency ratio are similar in the two economies while the population growth rate remains under 3%. As of 2020, Nigeria had a population of 206.1 million and Tanzania's population was 59.7 million (source: World Bank). The differences in public infrastructure do not appear to be extremely wide - since Tanzania has a total of 145,203 km length of roadways while Nigeria has 195,000 km (based on a 2022 estimate). With nearly 43% of land in Tanzania being agricultural and over 37% of the land being forest land (based on a 2018 estimate), urban residents are more numerous in Nigeria - where agricultural land is nearly 78% of the total land and leaves only 9.5% of forest land. Only 36.7% of the population is urban in Tanzania whereas urban residents constitute 53.5% of the population in Nigeria (2022 estimates). The differences in urbanisation levels across the two economies seem stable - since the rate of urbanisation is nearly the same in the two economies.

The higher urbanisation in Nigeria is not without severe disparities in living standards. The infant-mortality in Nigeria (56.68 /1000) - for example - is in fact nearly double that of Tanzania (30.87/1000) and the per-capita availability of physicians remains lower in Nigeria (0.38/1000) relative to Tanzania (0.7/1000). The overall literacy is slightly better in Tanzania as well - with 77.9% population considered literate (above 15 years of age having the ability to read or write in any language) compared to 62% in Nigeria. Even though the countries have similar school life expectancy (9 years), the youth unemployment (between ages 15-24) is at a staggering 18.3% in Nigeria compared to 3.9% in Tanzania. The overall poverty also seems more widespread in Nigeria - with 40.1% below the poverty line (estimate) compared to 26.4% of the population living below the poverty line in Tanzania (2017 estimate). This is despite higher income inequality in Tanzania which has a Gini-coefficient for household income of 40.5 (2017 estimate) compared to 35.1 for Nigeria (2018 estimate).

A brief overview of the two economies using data from OECD (Organisation for Economic Cooperation and Development) suggests that of the two economies, a higher percent of the GDP comes from agricultural production in Tanzania. The per-capita GDP 2020 for Nigeria (2396.04 USD) is much higher than that of Tanzania (976.16 USD) (source World Bank). The GDP from manufacturing, mining and construction is nearly double for Nigeria (8749.61 million USD in 2022) compared to Tanzania (4221.41 million USD in 2021). Nigeria's GDP from agriculture is 9311.36 million USD and Tanzania's is 4313.47 million USD.

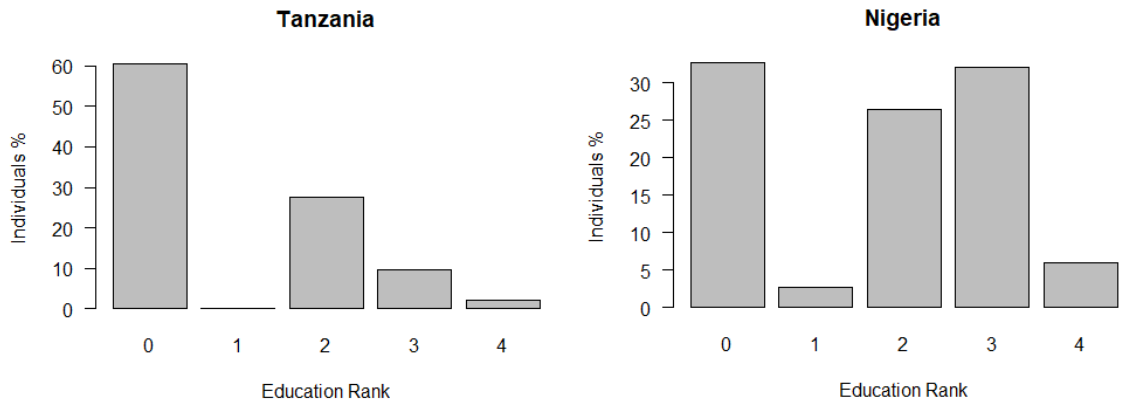


Figure 13: Education ranks for populations in Tanzania Nigeria (see Tables 29 and 30 for rank classification)

The structural differences seem to be due to a higher reliance on energy exports for Nigeria - which, as of 2020, exports $35.9 billion crude-gas and imports $7.75 billion of refined Petroleum apart from cars ($3.03 billion) and wheat ($2 billion). On the other hand, Tanzania's major exports are Gold and other minerals ($3 billion) followed by agricultural products (0.594$ billion). The major imports of Tanzania are refined copper and petroleum ($2.2B). Among the two economies, Nigeria relies less on the agrarian exports and more on the energy-related infrastructure - likely associated with the higher average education levels in Nigeria (see Figure 13).

Of the political systems in the two countries, Tanzania has shown a higher propensity to state control as nearly all land resources in Tanzania are government-owned. Tanzania is also more uniform ethnically - as about 95% of Tanzanian nationals identify as Bantu. On the other hand, the ethnic composition is more varied in Nigeria with Hausa (30%), Yoruba (15.5%) and Igbo (15.2%) as prominent ethnic groups. The policy of education has also seen different directions in the two economies. At a time when other countries in the region were expanding their education sector, the secondary education opportunities had faced some restrictions in Tanzania (Söderbom et al., 2006; Appleton et al., 1999; Buchert, 1992). The private education seems to have flourished with relatively fewer limits in Nigeria. The effects of all of these factors towards education expense should become clearer in further subsections.

5.2 Descriptive Statistics for Education Expenses

Considering the overall distribution of education expenses, we see that the urban residents spend much higher on education than the rural residents in both the economies. The urban residents spend 2.34 times more than rural residents in Nigeria whereas they spend 2.8 times more in Tanzania. The secondary education appears to get much costlier in Tanzania since the mean education expenditure for tertiary education (rank 3) is almost double that of the

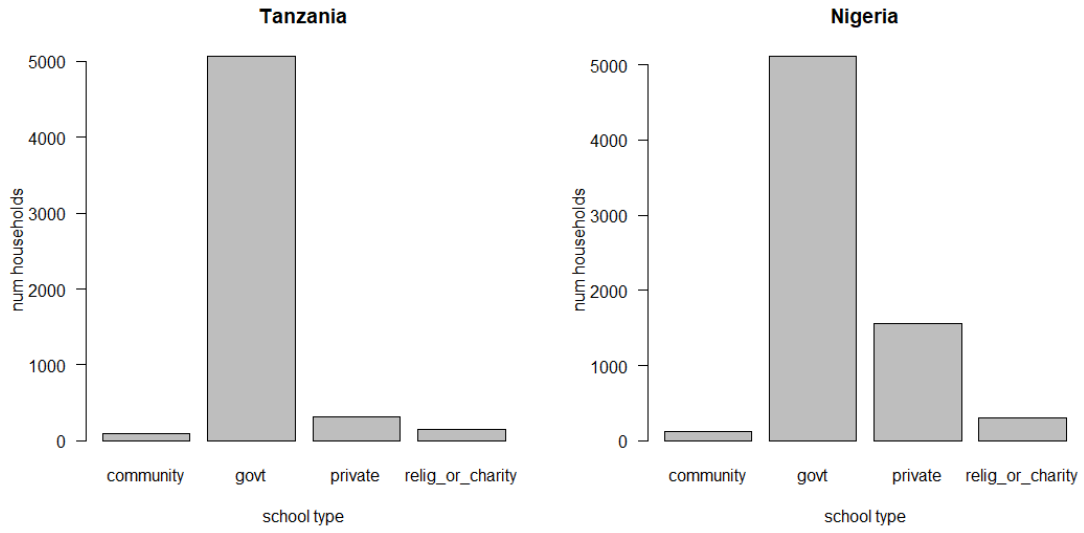


Figure 14: School types for non-adult population in Tanzania and Nigeria

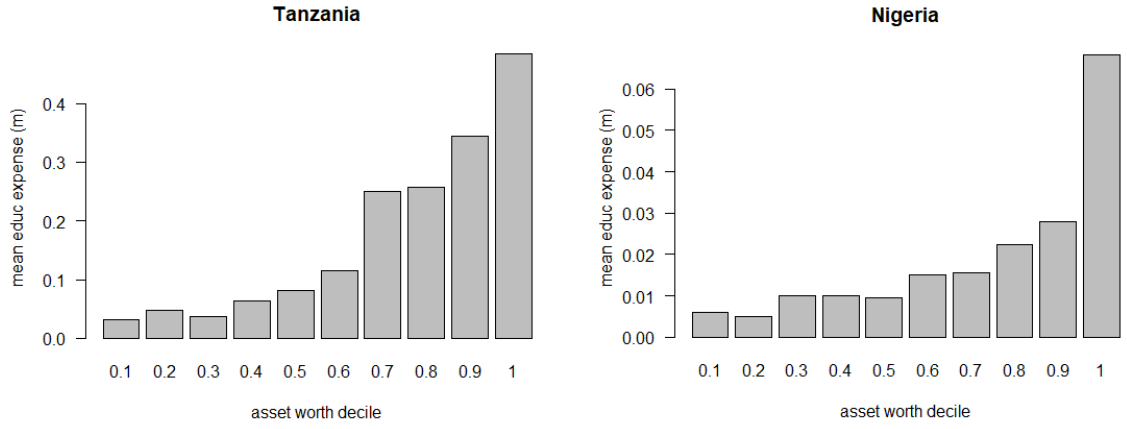


Figure 15: Educational expenses averages over deciles of asset worth in in Tanzania and Nigeria

expenditure on secondary education (rank 2) in Tanzania. In Nigeria, these expenditures rise by about 54%. The educational expenses do not seem to rise much with asset disparities in Nigeria - whereas they seem to be more unequally distributed in Tanzania (see Figure 15)[31]. Comparing education levels of the household heads, we see that the average-education rank of the household heads in Tanzania (1.8) is lower than that in Nigeria(1.9). We also see that the role of private education is much higher in Nigeria (see Figure 14).

	Tanzania			*Nigeria*		
Variable	$mean_{tnz}$	$median_{tnz}$	$stdev_{tnz}$	$mean_{ngr}$	$median_{ngr}$	$stdev_{ngr}$
w_educ	0.04	0.01	0.08	0.04	0	0.1
log_educ	7.77	10.49	5.46	3.97	0	4.94
ln_tot_exp	14.39	14.48	0.83	12.42	12.45	0.72
ln_tot_assets	12.96	12.76	1.93	10.93	11	1.35
localwealth	14.9	14.8	1.13	11.43	11.42	0.74
localwealth_educ	14.71	14.62	1.4	11.42	11.49	0.83
localwealth_agri	14.79	14.82	1.23	11.36	11.39	0.84
log_mean_cost_ne	12.9	12.97	0.53	10.85	10.86	0.53
rural_wards	0.66	0.78	0.33	0.71	1	0.41
secondary_schools	4.56	2	7.59	1.22	1	0.52
educpriv	0.06	0	0.23	0.19	0	0.39
is_primaryage	0.63	1	0.48	0.76	1	0.43
is_secondaryage	0.58	1	0.49	0.61	1	0.49
is_tertiaryage	0.45	0	0.5	0.4	0	0.49
agri	0.66	1	0.47	0.51	1	0.5
highesteduc	1.21	1	1.06	1.53	2	1.47
numchild	3.05	3	2.03	3.55	3	2.16
religion				1.5	1	0.52
has_english	0.36	0	0.48			

Table 22: Summary Statistics for Variables in Tables 12 and 21

Using the average, median and standard deviation values variables listed in Tables 12 and 21) that are shown in Table 22, we note that the measures for average budget shares (w_educ) as well as educational expenses (ln_tot_exp) are not dissimilar in both the economies. The average number of children in households that are of an age appropriate for primary, secondary or tertiary education are not far from each other either (see values for is_primaryage , is_secondaryage and is_tertiaryage in Table 22). A clear difference between Tanzania and Nigeria - however - is that of standard-deviation in local wealth (localwealth) - which is higher for Tanzania. Further, the predominance of agricultural occupations (see stdev values for agri) is higher for Tanzania as well.

It is worth pointing that barring such broad observations, the LSMS data used in our

[31]Despite the uniformity in LSMS data fields corresponding to education expenses, the collection methodology may not be the same across all the countries. This is evident from the much lower educational expenses in Nigeria than in Tanzania (see Table 22). Our comparison of the two economies (Tanzania and Nigeria) therefore does not rely on the measurement of educational expenses relative to one another.

analyses may not permit a cross country comparison for education expenses in a much finer detail. One reason is simply that the distribution of savings, income and employment opportunities may be too different in the two economies. The other reason is that there is also a difference in sizes or granularity of district classifications in the two economies. More specifically, the sizes of districts in Tanzania seem much larger (particularly in the rural parts) than in Nigeria. This is why the per-district density of secondary education facilities - which we measure with the variable `secondary_schools` - is reported lower in Nigeria (see values corresponding to `secondary_schools` in Table 22) - while the education levels of the consumer populations remains higher in Nigeria (see highest_educ in Table 22). The ruralness of districts - measured as number of wards declared as rural in a district (i.e. the variable `rural_wards`) - does not seem affected by this disparity in sizes as much (compare the values for `rural_wards` in Table 22) - since a rural-urban classification that declares boundaries of a district based on areas surrounding an urban (or semi-urban) centre tends to bring a uniformity to the number of rural wards in a district.

It is worth emphasising that the overall distribution of assets (durable goods stocks) in the two countries is disparate - as is clear with a direct comparison of asset values. Contrasting the Lorenz curves for total non-durable expenditure and total assets in the Figure 16 for in Nigeria and Tanzania, we note that distribution of assets is more sparse in Tanzania even though the overall expenditure appears similarly distributed in the two countries.

A geographical distribution of assets further details that higher-income occupations are far fewer in the hinterland and the southern regions of Tanzania (see Figure 17) - relative to the eastern and coastal regions that seem largely urban. The disparities in occupations are less in Nigeria (see Figure 18).

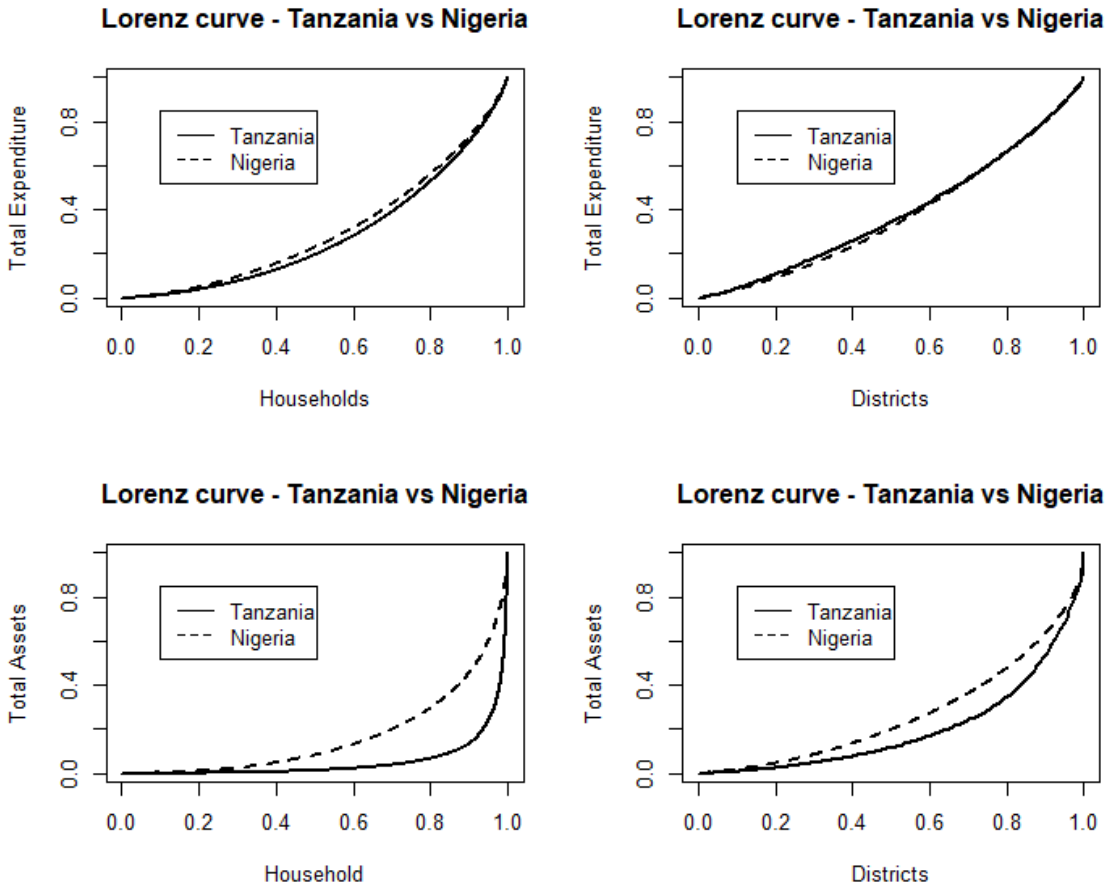


Figure 16: Inequality in assets and non-durable expenditure across households and districts

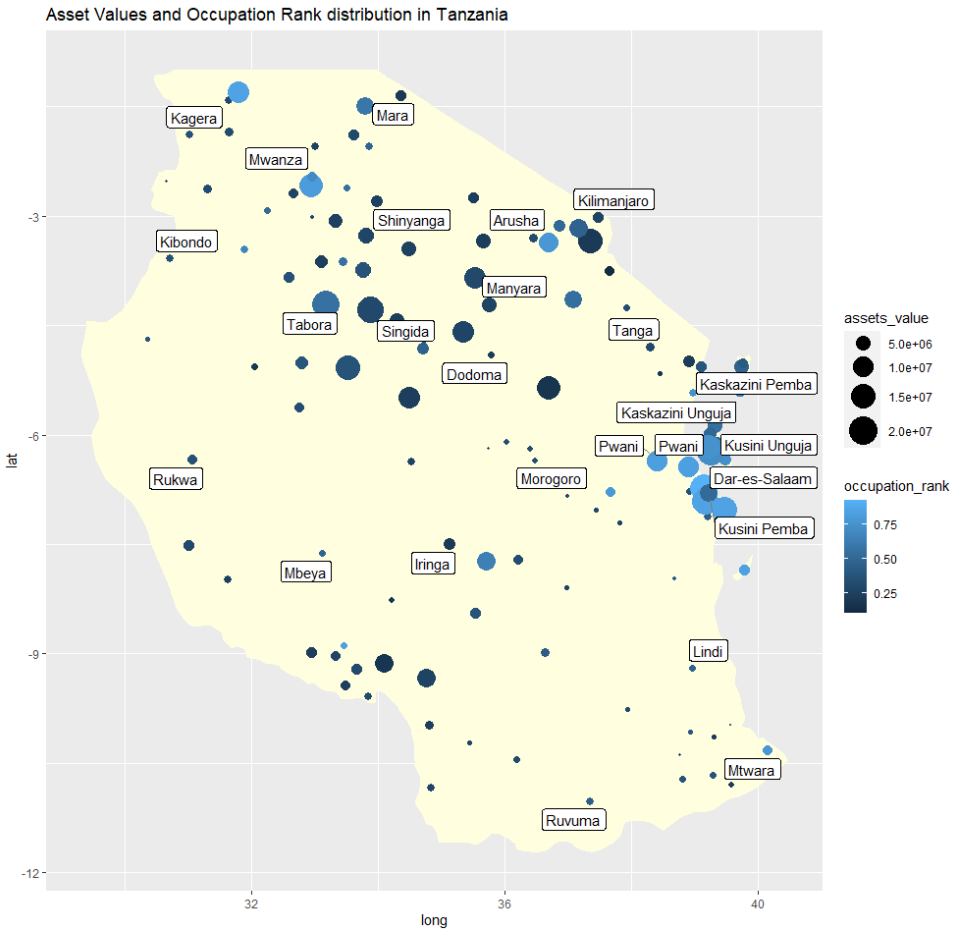


Figure 17: Occupations and consumer-owned assets in Tanzania

Note that the occupation ranks shown in Figures 17 (Tanzania) and Figure 18 (Nigeria) are based on the standardisation of various occupations in Tanzania and Nigeria. Since the range of occupations in the surveys from Tanzania and Nigeria is varied, the different occupations in the survey for Tanzania and Nigeria have been standardised into occupation ranks (ranging from 0 to 3) based on the data on income derived from the occupations in the survey so that a higher occupation rank represents an occupation that provides higher income. The reason why the income itself is not shown in the plots is that the data on income is far sparser than the recall of assets and the reporting of occupations in surveys from both the economies[32]. The mapping from specific occupations to the occupation-ranks for Tanzania and Nigeria is shown in Table 31 and 32 (respectively). A higher occupation rank represents occupations that bring higher income. As described in Section 4, the education levels are also mapped to education ranks denoting primary, secondary and higher education (see Table 29 for Tanzania and Table 30 for Nigeria in the Appendix 7.1). A higher education rank corresponds to a higher education level.

While we undertake a more detailed view on education expenses in the Section 5.3, the

[32]It is not uncommon for the empirical studies in SSA to have the data on occupation or education proxy the income differences (Alesina et al., 2019). The primary reason for this is the sparsity of income data - a problem that may be further aggravated because of a large informal sectors in the economies.

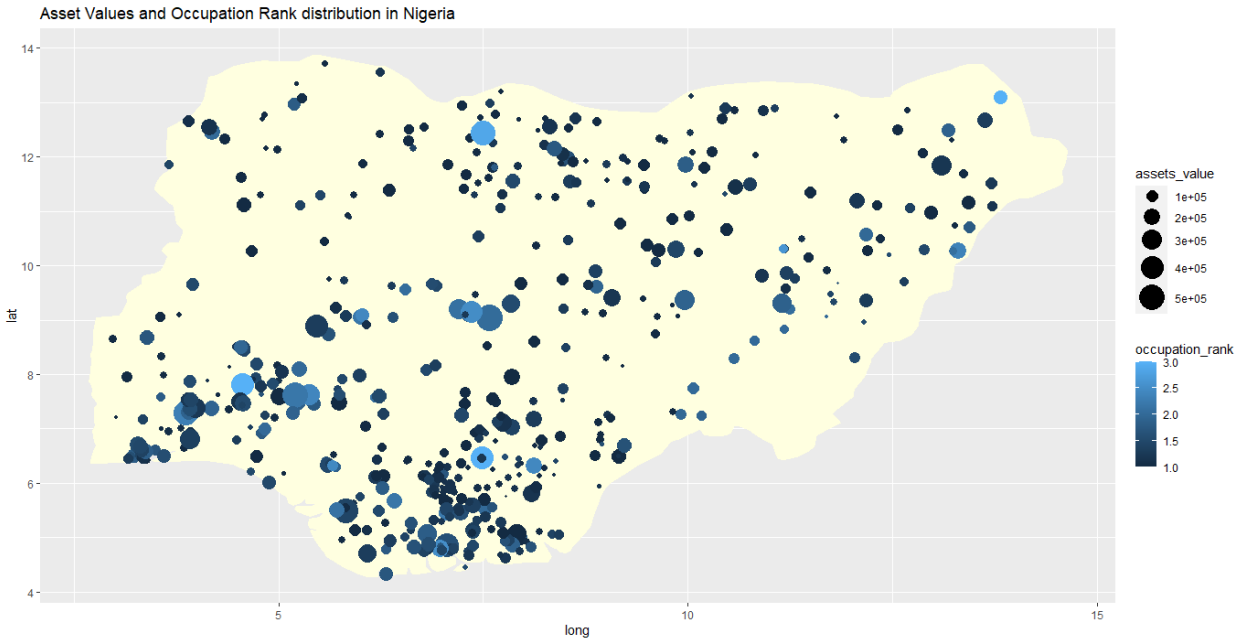


Figure 18: Occupations and consumer-owned assets in Nigeria

key observation from the data is that the disparities in income and durable goods are more localised in Nigeria while they are more extreme across wider geographical regions in Tanzania. More specifically, despite the southern areas of Nigeria having more higher-paid occupations and a higher expenditure on non-durable expenditure overall (see Figure 18), the disparity in durable-goods ownership across regions is not as extreme as in Tanzania. The effect of this non-uniformity on consumption and educational expenses is discussed in Section 5.3.

5.3 Results

The two dependent variables we use in the estimation described in Section 3 (see Equation 23) are the budget shares for educational expenses (w_{educ}) and total expenditures ($log(x_{educ})$) . These dependent variables are listed in Table 20 while controls used for the estimation are listed in Table 21.

The two main explanatory variables used in the study are the logarithm of total expenditure `ln_tot_exp` and the local wealth measure `localwealth`. As discussed before, an alternative formulation is also specified for comparison - where the two explanatory variables are the logarithm of total assets value (`ln_tot_assets`) and the urban/rural indicators (`rural_wards`). The physical access to education (i.e. `secondary_schools`) is an important control in both the formulations. Other controls used in the comparison are the indicator for an agrarian occupation `agri` (Ω), the paternal education level `highesteduc` (Γ^*), the religion and language indicators (`religion` χ used only for Nigeria indicated by 1 for Christianity, 2 for Islam, 3 for Traditional, 4 for other - and English-speaking ability i.e. the variable `has_english` ξ indicating the ability to speak English used only for Tanzania),

the access to private education (`educpriv` or π), the number of children (`numchild` or h) and their ages `is_primaryage` (μ_1), `is_secondaryage` (μ_2), `is_tertiaryage` (μ_3) indicating if they are eligible (due to age) for primary, secondary or tertiary education (respectively) and the non-durable consumption prices (`log_mean_cost_ne` p_{ne}). All of the controls remain the same for the main as well as alternative formulation. Recall that the non-durable consumption prices p_{ne} (see Section 4) - interpreted as the cost of living in the vicinity - is measured as `log_mean_cost_ne` - the logarithm of average consumption excluding the asset-related costs in the consumer's vicinity. Notice that the factor variables are prefixed with respective values `1.` , `2.` etc. in the results.

Of the control variables and explanatory variables discussed above, some are observed at the household level while others are observed at the level of the consumer's district. The logarithm of total expenditure `ln_tot_exp` and of the total assets-owned `ln_tot_assets` are observed at the level of the household. Similarly, the agricultural-occupation factor variable `agri` (indicating whether the household head is employed in agriculture related occupation or not), the number of children in the household `numchild`, the childrens' ages factor variables `is_primaryage` , `is_secondaryage` and `is_tertiaryage`), the private education factor variable `educpriv`, the highest education rank `highesteduc` corresponding to the paternal generations (father-relationship) in a household, the ability to speak English `has_english` (used only for Tanzania) and the household religion `religion` (used only for Nigeria) are all observed for a household. The variables observed at the district (or 6-km vicinity) are the number of wards with an accessible (within 6-km distance) secondary-schools (`secondary_schools`), the rural wards (`rural_wards`) in a district, the respective measures of the local wealth `localwealth` and the logarithm of the per-head cost non-durable consumption `log_mean_cost_ne` .

To reiterate, a significant number of household units in both the economies are made up of joint families where the reported educational expenses correspond to the grandchildren of the household head (HH). The education level of the HH would thus point to the penultimate generation for households where HH has grandchildren and the previous generation for younger households where HH has young children. To resolve this disparity, we avoid using the education rank of the HH and instead use the highest education rank in the generation of the students' parents [33] `highesteduc` instead. The paternal educational level `highesteduc` suffices for our analysis of expenditure spent over all the children in the household (rather than how education budget is distributed among the male/female children). As mentioned in the Section 3, the households with no children are ignored in the analysis. Nearly 22% of households in the Nigeria and 15 % of households in Tanzania are thus excluded from our comparison.

Notice also that the age-variables `is_primaryage`, `is_secondaryage`,

[33] Only relations with the household head are reported in the survey. Thus, there is no direct way of locating the parent of a child except through the relation with the household-head.

`is_tertiaryage` for attendance of primary, secondary and tertiary education have been used in the analysis to get around the unavailability of the current-education level of the household member in the Nigerian LSMS survey data. Instead, we use a combination of the indicator variable telling whether a household member is in school or not (available in the surveys for both Tanzania and Nigeria) and the mean school attendance ages to define variables `is_primaryage`, `is_secondaryage` and `is_tertiaryage` that indicate whether there is a household member of an age appropriate for attending primary, secondary and tertiary education levels (respectively).

The results from the a tobit estimation for educational expenses (w_{educ}) and total expenditures ($log(x_{educ})$) with local wealth `localwealth` as one of the controls are shown in Table 23 for both Tanzania and Nigeria. The tobit uses a lower limit of zero given a number of houses have no expenditure on education. A zero expenditure is reported for nearly 65% households in the survey from Nigeria (2015) and 43% households in the survey from Tanzania (2014). This suggests that either the education expenses are recorded in more detail in Tanzania or that a certain basic level of expenses is not considered in the survey for Nigeria. The particular disparity also limits our ability to compare educational expenses relative to one another for the two economies. A discussion on the cross-section results from the latest available years 2014 for Tanzania and 2015 for Nigeria follows in Section 5.4 (see Table 23). The results from the alternative formulation - where the main explanatory variables are the logarithm of total assets value `ln_tot_assets` (used in place of household total expenditure) and the ruralness-score `rural_wards` derived from the survey's classification of rural and urban areas (used in place of local wealth `localwealth`) - are also discussed in 5.4 (see Table 24).

5.4 Discussion

Results with w_{educ} as the dependent variable

The coefficient of the total expenditure (`ln_tot_exp`) in Table 23 for w_{educ} as the dependent variable suggest that an increase in the total expenditure (`ln_tot_exp`) causes a decrease in the budget share on education (w_{educ}) for both Tanzania and Nigeria. This seems largely due to the educational expenses forming a higher proportion of the total expenditure for poorer households - since the education expenditures $log(x_{educ})$ do seem to rise with the increase in total expenditure (`ln_tot_exp`) in Tanzania (see coefficient of `ln_tot_exp` under columns for $log(x_{educ})$ in Table 23). The effect of non-durable consumption prices (`log_mean_cost_ne`) is weak on education-budget share w_{educ} for both the economies (see coefficient of `log_mean_cost_ne` under w_{educ} in Table 23) . Thus while urban developments must raise the prices of non-durable expenditure `log_mean_cost_ne`, the disparity in budget shares does not appear to be due to higher cost of living (non-durable consumption) alone. In other words, the effect of lower or higher cost of living on education expenditure is weak across the regions in both the economies.

Instead, we see a strong and positive effect of local wealth (`localwealth`) towards educational budget shares (w_{educ}) in both the countries in Table 23 (see coefficients of `localwealth` under w_{educ} in Table 23). This is an indication of the regional disparities in the economies. More particularly, the differences in average assets accumulation indicate the local wealth has a strong effect on education expenses - despite controlling for the effect of physical access. While absence of educational facilities must clearly suppress education expenses in a region, the positive effect of local wealth after controlling for availability of education facilities coul be due to certain environmental factors that contribute to education - such as a competition towards higher income social positions or presence of occupations requiring higher levels of education - factors with an important role after the availability of educational facilities has been considered.

The effects of household total expenditure remain relevant - however - for private education - which is limited to the wealthier households in both the economies regardless of local wealth (`localwealth`). While a higher proportion of young population seems to be in private schools in Nigeria (see Figure 14), private education is slightly higher in budget share in Tanzania as the privately-educated households in Nigeria show higher budget shares w_{educ} (see coefficients of `educpriv` under w_{educ} for the two countries in Table 23). However, since private education only accounts for a minority of the population in both the economies (see Figure 14), the role of access to either or private or government secondary school i.e. we focus on the variable `secondary_schools` for our comparison of education expenses in the two countries.

Viewing the effect of the variable `secondary_schools` on education budget shares w_{educ},

we see a far stronger effect of secondary_schools for Tanzania (compare the coefficients for secondary-school s i.e. secondary_schools in Tables 23) . It also seems that the consumers are more likely to spend on primary education in Tanzania - as the access to secondary and tertiary-education remains much lower (this is further elaborated in discussion below of the results using education expenditure $log(x_{educ})$ as dependent variable).

The results from the alternative formulation using urban-rural indicators suggest a particular limitation of the use of urban-rural indicators from the survey. More specifically, the coefficients of rural-wards (rural_wards) for education budget shares w_{educ} are negative and significant for Tanzania in Table 24. However, the inference of a higher expenditure on education - when using urban-rural indicators - by the rural residents in Nigeria (see significance of rural-wards rural_wards in Table 24) may be improper as it ignores the fact that the wards identified as rural in Nigeria may not be as rural as those identified as rural in Tanzania. This is confirmed with the weaker role of wealth i.e. assets-worth-logarithm (ln_tot_assets in Table 24) and a strong role for local wealth (localwealth in Table 23) for Nigeria in the alternative formulation.

Results with $log(x_{educ})$ as the dependent variable

Comparing the results with education-expenditure logarithm $log(x_{educ})$ as the dependent variable, we see that the education expenditures rise with the increase in total expenditure logarithm (ln_tot_exp) for Tanzania but the effect of total expenditure ln_tot_exp (in Table 23) is not significant for Nigeria. While the higher expenditure on education measured through total education expenditure logarithm $log(x_{educ})$ does seem to respond to the prices of non-durable consumption log_mean_cost_ne in Tanzania (see coefficient of log_mean_cost_ne under column $log(x_{educ})$ in Table 23), the effect of non-durable consumption prices log_mean_cost_ne remains weak on education-budget share w_{educ} for both the economies (see above discussion using budget share w_{educ} as the dependent variable). We also see that the educational expenditure $log(x_{educ})$ rises with the second explanatory variable i.e. the local wealth (localwealth) for Nigeria while the effect of local wealth localwealth is insignificant for Tanzania (compare coefficients for localwealth against both the countries in Table 23 under columns for $log(x_{educ})$). In the alternative formulation where urbanisation is measured with the ruralness score (rural_wards), the ruralness score corresponds to a rise in educational expenditure $log(x_{educ})$ for Nigeria and a drop in educational expenditure $log(x_{educ})$ for Tanzania (see coefficients for rural_wards under columns for $log(x_{educ})$ in Table 24). The disparity in the role of rural_wards is however also due to differences in the extent to which density of educational institutions depends on rural_wards in the respective countries.

Continuing with the comparison of results with education-expenditure $log(x_{educ})$ as the dependent variable, we see that private-education indicator variable educpriv is insignificant for Nigeria (see Table 23). The education expenses in Nigeria therefore depend

less often on whether education is provided through private educational institutions or not. The results also indicate lower attendance (based on age) to secondary and tertiary-education in Tanzania (see coefficients of secondary-school-age `is_secondaryage` and tertiary-school-age `is_tertiaryage` in regression results with education expenditure logarithm $log(x_{educ})$ as the dependent variable in Table 23). In the alternative formulation, the results with education expenditure logarithm $log(x_{educ})$ as dependent variable in Tables 24 show the role of secondary-school access `secondary_schools` to be stronger for Nigeria when taking wealth i.e. total assets-worth logarithm `ln_tot_assets` into account. The permanent income effect thus seems stronger than access in both the economies but is weaker for Nigeria where the enrollment into secondary and tertiary education is higher - regardless of the differences in long-term assets (measured with logarithm of assets-worth `ln_tot_assets`). The decomposed view of expenses on education in terms of secondary and primary education level further clarifies this observation - reflecting the far lower access to secondary education in Tanzania.

Other controls

The role of the number of children (`numchild`) also seems stronger in Tanzania than in Nigeria - but the effect is likely due the education expenses primarily driven by the primary education levels in Tanzania (as well as a higher percentage of households reporting no education expenditure in the survey from Nigeria). Another consequence of most households spending on primary education alone is that the agrarian households appear as spending higher on education after considering the local wealth among households with the same educational-status. The education expenditure is higher for agrarian households in Tanzania after taking into account their poor educational status (see coefficients of agrarian-occupation-indicator `1.is_agri` in Table 23 under column for the education-expenditure-logarithm $log(x_{educ})$ dependent variable).

The effects of permanent income in Tanzania overall seem stronger, but they ought to be considered in light of the much lower expenditure on secondary (and higher) education and higher disparity in asset-ownership (even after taking into account the lower urbanisation levels) in Tanzania. The significance of logarithm of assets-worth (`ln_tot_assets` in Table 24) does indicate a general trend where those with higher inter-generational wealth and in urban areas (see significance of `ln_tot_assets` and `rural_wards` in Table 24 for Tanzania $log(x_{educ})$ as dependent variable) are far more likely to avail the education facilities, but a higher dependence on industrial output in Nigeria also seems to have contributed to a higher demand for secondary and tertiary education in the economy overall (compare with significance of `ln_tot_assets` , `1.is_secondaryage` and `rural_wards` in Table 24 for Nigeria with $log(x_{educ})$ as dependent variable). Another observation - the stronger role of father's education level in Tanzania - also seems related with these disparities in the two economies as well. A comparison of education expenditures across the

two countries must take into account the different stages which the two countries appear to be in terms of urbanisation and economic development.

Finally, the role of social environment is found significant in both the economies. The role of English language seems significant in Tanzania - where the English-speakers spend significantly higher on education (see coefficients of `has_english` in Table 23 under $log(x_{educ})$) . In Nigeria, on the other hand, the educational expenses are higher among households identifying as Christian (see coefficients of `religion` in Table 23 under $log(x_{educ})$) . In both cases, the role of social characteristics may be rather inseparable from differences in educational institutions. In case of Tanzania, the ability to speak English is aligned with attending educational institutions while in Nigeria, the role played by Christian missionaries in establishment of educational institutions may be reflected in the disparity of access to educational facilities.

Table 23: Tanzania -Nigeria - Differences in Local Wealth

	w_{educ}		$log(x_{educ})$	
	Tanzania	*Nigeria*	*Tanzania*	*Nigeria*
ln_tot_exp	-0.00908**	-0.0283***	1.224***	0.133
	(0.004)	(0.000)	(0.000)	(0.744)
localwealth	0.00488*	0.0299***	-0.0368	1.497***
	(0.015)	(0.000)	(0.758)	(0.000)
log_mean_cost_ne	0.0490	0.0479	9.165*	-0.155
	(0.467)	(0.672)	(0.023)	(0.981)
1.agri	0.00735	0.00990	0.953**	0.574
	(0.167)	(0.308)	(0.003)	(0.297)
1.educpriv	0.0756***	0.0475***	2.890***	0.815
	(0.000)	(0.000)	(0.000)	(0.197)
1.highesteduc	0.0109*	0.0829	0.694*	5.899
	(0.030)	(0.301)	(0.021)	(0.202)
2.highesteduc	0.00432	0.00679	0.366	0.363
	(0.711)	(0.585)	(0.598)	(0.607)
3.highesteduc	0.00457	-0.00520	0.473	-0.528
	(0.519)	(0.627)	(0.263)	(0.381)
4.highesteduc	0.0250*	0.0120	1.465*	0.183
	(0.011)	(0.506)	(0.013)	(0.859)
secondary_schools	-0.00138***	0.00738	-0.00569	0.384
	(0.000)	(0.373)	(0.734)	(0.414)
1.is_primaryage	0.0793***	0.0538***	8.842***	3.301***
	(0.000)	(0.000)	(0.000)	(0.000)
1.is_secondaryage	0.0123**	0.0605***	0.143	2.230***
	(0.005)	(0.000)	(0.580)	(0.000)
1.is_tertiaryage	-0.0131**	0.0174	-1.281***	0.931
	(0.001)	(0.056)	(0.000)	(0.071)
1.has_english	0.0584***		2.249***	
	(0.000)		(0.000)	
numchild	0.00433***	-0.00221	0.259***	-0.0499
	(0.000)	(0.360)	(0.000)	(0.714)
2.religion		-0.0455***		-2.043***
		(0.000)		(0.000)
3.religion		-0.0632		-2.160
		(0.170)		(0.396)
_cons	-0.151	-0.231	-41.97***	-22.99
	(0.346)	(0.352)	(0.000)	(0.103)
sigma				
_cons	0.0876***	0.173***	5.361***	9.969***
	(0.000)	(0.000)	(0.000)	(0.000)
N	2410	2192	2410	2192

Table 24: Tanzania-Nigeria - Urban-Rural Differences

	w_{educ}		$log(x_{educ})$	
	Tanzania	*Nigeria*	*Tanzania*	*Nigeria*
ln_tot_assets	-0.00590***	-0.00196	0.392***	0.363
	(0.000)	(0.576)	(0.000)	(0.065)
log_mean_cost_ne	-0.0548	-0.00483	9.559*	8.263
	(0.455)	(0.965)	(0.031)	(0.182)
1.agri	0.0106*	0.00507	0.925**	-0.0768
	(0.048)	(0.617)	(0.004)	(0.892)
1.educpriv	0.0800***	0.0506***	2.730***	1.016
	(0.000)	(0.000)	(0.000)	(0.113)
1.highesteduc	0.0107*	0.0948	0.679*	7.172
	(0.032)	(0.244)	(0.023)	(0.121)
2.highesteduc	0.00199	0.00689	0.512	0.481
	(0.864)	(0.583)	(0.462)	(0.495)
3.highesteduc	0.00359	-0.00767	0.607	-0.551
	(0.608)	(0.478)	(0.149)	(0.361)
4.highesteduc	0.0273**	0.0129	1.320*	0.486
	(0.005)	(0.480)	(0.025)	(0.636)
secondary_schools	-0.00111***	0.0147	0.00265	0.906
	(0.000)	(0.085)	(0.875)	(0.058)
1.is_primaryage	0.0797***	0.0562***	8.812***	3.408***
	(0.000)	(0.000)	(0.000)	(0.000)
1.is_secondaryage	0.0129**	0.0593***	0.145	2.259***
	(0.003)	(0.000)	(0.572)	(0.000)
1.is_tertiaryage	-0.0121**	0.0139	-1.240***	0.878
	(0.003)	(0.132)	(0.000)	(0.088)
rural_wards	-0.0229**	0.0133	-1.184*	1.848**
	(0.007)	(0.291)	(0.021)	(0.009)
1.has_english	0.0599***		2.436***	
	(0.000)		(0.000)	
numchild	0.00426***	-0.00354	0.304***	-0.0588
	(0.000)	(0.143)	(0.000)	(0.662)
2.religion		-0.0445***		-1.716**
		(0.000)		(0.003)
3.religion		-0.0725		-2.411
		(0.120)		(0.343)
_cons	0.143	-0.103	-30.47**	-29.99*
	(0.452)	(0.698)	(0.008)	(0.044)
sigma				
_cons	0.0870***	0.175***	5.357***	9.994***
	(0.000)	(0.000)	(0.000)	(0.000)
N	2410	2192	2410	2192

Table 25: Tanzania-Nigeria - around median assets-value

	w_{educ}		$log(x_{educ})$		w_{educ}		$log(x_{educ})$	
	Above Median *Tanzania*	Above Median *Nigeria*	Above Median *Tanzania*	Above Median *Nigeria*	Below Median *Tanzania*	Below Median *Nigeria*	Below Median *Tanzania*	Below Median *Nigeria*
ln_tot_exp	-0.00338	-0.0168	0.783**	0.411	-0.0121**	-0.0458***	1.272***	-0.582
	(0.483)	(0.113)	(0.005)	(0.511)	(0.006)	(0.000)	(0.000)	(0.321)
localwealth	0.00512	0.0321***	-0.160	1.818***	0.00592	0.0227*	-0.0190	0.819
	(0.061)	(0.000)	(0.311)	(0.000)	(0.052)	(0.024)	(0.919)	(0.131)
log_mean_cost_ne	-0.0900	-0.00476	11.02	-6.579	0.165	0.100	7.127	5.568
	(0.363)	(0.976)	(0.056)	(0.479)	(0.080)	(0.539)	(0.219)	(0.528)
1.agri	0.00286	0.00554	0.758	0.341	0.0129	0.0157	1.151*	0.809
	(0.671)	(0.665)	(0.052)	(0.650)	(0.139)	(0.294)	(0.031)	(0.314)
1.educpriv	0.0850***	0.0348*	3.219***	0.726	0.0527**	0.0711***	2.083	0.896
	(0.000)	(0.010)	(0.000)	(0.364)	(0.004)	(0.000)	(0.074)	(0.406)
1.highesteduc	0.0102	0.0847	0.443	8.091	0.0118	0.107	0.850*	5.336
	(0.156)	(0.482)	(0.292)	(0.263)	(0.089)	(0.323)	(0.047)	(0.368)
2.highesteduc	0.0208	0.00761	0.759	0.178	-0.0235	0.00660	-0.380	0.618
	(0.170)	(0.645)	(0.390)	(0.855)	(0.207)	(0.727)	(0.735)	(0.543)
3.highesteduc	-0.000380	-0.00517	-0.0789	-0.359	0.0114	-0.00197	1.171	-0.594
	(0.967)	(0.711)	(0.882)	(0.661)	(0.310)	(0.906)	(0.092)	(0.508)
4.highesteduc	0.0204	0.00328	0.796	-0.695	0.0368	0.0426	2.572*	2.785
	(0.079)	(0.877)	(0.240)	(0.577)	(0.062)	(0.222)	(0.036)	(0.139)
secondary_schools	-0.00128***	-0.00325	-0.00576	-0.524	-0.00161**	0.0230	0.00286	1.669*
	(0.000)	(0.764)	(0.770)	(0.413)	(0.001)	(0.077)	(0.926)	(0.018)
1.is_primaryage	0.0610***	0.0465**	7.280***	2.766**	0.101***	0.0664***	10.53***	3.983***
	(0.000)	(0.003)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
1.is_secondaryage	0.00978	0.0692***	0.0488	2.637***	0.0161*	0.0502***	0.243	1.706*
	(0.100)	(0.000)	(0.887)	(0.000)	(0.011)	(0.000)	(0.533)	(0.020)
1.is_tertiaryage	-0.0109*	0.0152	-1.381***	0.727	-0.0151*	0.0193	-1.311***	1.108
	(0.043)	(0.197)	(0.000)	(0.292)	(0.015)	(0.181)	(0.001)	(0.152)
1.has_english	0.0496***		1.789***		0.0704***		2.575***	
	(0.000)		(0.000)		(0.000)		(0.000)	
numchild	0.00204	-0.00315	0.267**	-0.115	0.00657***	-0.00143	0.231*	0.0342
	(0.167)	(0.307)	(0.002)	(0.523)	(0.000)	(0.713)	(0.027)	(0.870)
2.religion		-0.0439***		-1.963*		-0.0466**		-2.037*
		(0.001)		(0.011)		(0.003)		(0.015)
3.religion		-0.121		-6.104		-0.0373		-0.447
		(0.144)		(0.192)		(0.516)		(0.882)
_cons	0.145	-0.250	-36.11**	-13.01	-0.447	-0.103	-39.80**	-22.74
	(0.521)	(0.475)	(0.006)	(0.527)	(0.061)	(0.777)	(0.007)	(0.246)
sigma								
_cons	0.0861***	0.169***	5.114***	10.13***	0.0877***	0.176***	5.545***	9.644***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	1233	1192	1233	1192	1177	1000	1177	1000

5.5 Robustness Checks

Assets ownership in poorer and richer sub-samples

To confirm that the assets-ownership does influence the access to education, we inspect the role of assets in lower and higher sub-samples of the surveyed population. We thus re-estimate the results with local wealth `localwealth` as the explanatory variable for those with less than the local median wealth i.e. assets-worth logarithm `ln_tot_assets` and for those above it. The results demonstrate how the education expenditures (coefficients for education budget-shares w_{educ} and education-expenditure logarithm $log(x_{educ})$) differ between the poorer and richer households in a given locality. The results for above and below local median-assets possession using education budget-shares w_{educ} and education-expenditure-logarithm $log(x_{educ})$ as the dependent variables are shown in Table 25 for both above and below median assets-possession levels in the two countries.

While the role that the local wealth `localwealth` plays towards access to education appears similar for the rich and poor halves in a locality as well, we do see a higher economic significance of local wealth `localwealth` for the below-median-assets household in Nigeria - possibly because of high education costs in urban areas where asset-possession may be lower. The access to private education seems more widerspread in Nigeria as the below-median households seem spend a higher budget-share on private-education. In Tanzania, on the other hand, the economic significance of private-education factor variable `educpriv` (as well as tertiary level education) is higher for above-median households. The significance of social factors - on the other hand - remains unaffected by whether the consumers have above or below the median asset ownership level. More specifically, the factor variables for English-speaking ability `has_english` and (i.e. `religion`) are significant in results from Tanzania and Nigeria (respectively) for both richer and poorer households. The factor variable for English-speaking ability `has_english` is significant for consumers both above and below media asset possession level (the coefficient for variable `religion` is significant for Nigeria for above and below median asset possession level in Table 25). The effect of secondary-schools `secondary_schools` is also the same for both richer and poorer halves of a given locality.

Variable	Variable Name	Description	Additional Controls
r	`localwealth`	Average over geographical vicinity (district)	Ω , education-rank
r_Γ	`localwealth_educ`	Separate averages over Γ	Ω
r_Ω	`localwealth_agri`	Separate averages over Ω	education-rank

Table 26: Prosperity levels and control variables used

Asset-density interpretations based on similar neighbours

To ensure that the effect of assets remains strong after accounting for occupation or

Table 27: Tanzania-Nigeria - $localwealth_{educ}$

	w_{educ}		$log(x_{educ})$	
	Tanzania	*Nigeria*	*Tanzania*	*Nigeria*
ln_tot_exp	-0.00927**	-0.0291***	1.229***	0.0988
	(0.004)	(0.000)	(0.000)	(0.809)
localwealth_educ	0.00262	0.0272***	-0.0349	1.323***
	(0.102)	(0.000)	(0.714)	(0.000)
log_mean_cost_ne	0.0535	0.0570	9.210*	0.361
	(0.428)	(0.614)	(0.022)	(0.955)
1.agri	0.00598	0.00747	0.957**	0.449
	(0.256)	(0.440)	(0.002)	(0.412)
1.educpriv	0.0763***	0.0471***	2.887***	0.799
	(0.000)	(0.000)	(0.000)	(0.207)
1.highesteduc	0.0109*	0.0819	0.697*	5.867
	(0.031)	(0.307)	(0.020)	(0.204)
2.highesteduc	0.00402	0.00520	0.369	0.285
	(0.731)	(0.676)	(0.595)	(0.687)
3.highesteduc	0.00452	-0.00734	0.475	-0.635
	(0.524)	(0.492)	(0.261)	(0.293)
4.highesteduc	0.0250*	0.0105	1.469*	0.111
	(0.011)	(0.563)	(0.013)	(0.914)
secondary_schools	-0.00133***	0.00756	-0.00571	0.400
	(0.000)	(0.361)	(0.732)	(0.395)
1.is_primaryage	0.0791***	0.0529***	8.844***	3.258***
	(0.000)	(0.000)	(0.000)	(0.000)
1.is_secondaryage	0.0119**	0.0599***	0.148	2.199***
	(0.006)	(0.000)	(0.566)	(0.000)
1.is_tertiaryage	-0.0133**	0.0177	-1.276***	0.950
	(0.001)	(0.051)	(0.000)	(0.065)
1.has_english	0.0579***		2.253***	
	(0.000)		(0.000)	
numchild	0.00432***	-0.00231	0.260***	-0.0547
	(0.000)	(0.337)	(0.000)	(0.687)
2.religion		-0.0441***		-1.971***
		(0.000)		(0.001)
3.religion		-0.0609		-2.074
		(0.185)		(0.414)
_cons	-0.124	-0.208	-42.21***	-21.67
	(0.438)	(0.401)	(0.000)	(0.124)
sigma				
_cons	0.0877***	0.173***	5.361***	9.974***
	(0.000)	(0.000)	(0.000)	(0.000)
N	2410	2192	2410	2192

Table 28: Tanzania-Nigeria - $localwealth_{occup}$

	w_{educ}		$log(x_{educ})$	
	Tanzania	*Nigeria*	*Tanzania*	*Nigeria*
ln_tot_exp	-0.00924**	-0.0308***	1.147***	0.00661
	(0.003)	(0.000)	(0.000)	(0.987)
localwealth_agri	0.00322	0.0276***	-0.00759	1.352***
	(0.066)	(0.000)	(0.942)	(0.000)
log_mean_cost_ne	0.0225	0.0496	3.952	-0.206
	(0.718)	(0.654)	(0.287)	(0.974)
1.educpriv	0.0761***	0.0449***	2.930***	0.677
	(0.000)	(0.000)	(0.000)	(0.281)
1.highesteduc	0.0109*	0.0882	0.670*	6.186
	(0.029)	(0.273)	(0.025)	(0.182)
2.highesteduc	0.00401	0.00573	0.320	0.310
	(0.732)	(0.644)	(0.646)	(0.660)
3.highesteduc	0.00444	-0.00658	0.433	-0.603
	(0.531)	(0.535)	(0.306)	(0.314)
4.highesteduc	0.0250*	0.0104	1.382*	0.0953
	(0.011)	(0.563)	(0.019)	(0.926)
secondary_schools	-0.00138***	0.00712	-0.0110	0.374
	(0.000)	(0.389)	(0.514)	(0.425)
1.is_primaryage	0.0794***	0.0542***	8.922***	3.328***
	(0.000)	(0.000)	(0.000)	(0.000)
1.is_secondaryage	0.0124**	0.0621***	0.168	2.314***
	(0.004)	(0.000)	(0.516)	(0.000)
1.is_tertiaryage	-0.0130**	0.0179*	-1.287***	0.968
	(0.001)	(0.048)	(0.000)	(0.059)
1.has_english	0.0585***		2.294***	
	(0.000)		(0.000)	
numchild	0.00444***	-0.00169	0.279***	-0.0231
	(0.000)	(0.479)	(0.000)	(0.864)
2.religion		-0.0456***		-2.063***
		(0.000)		(0.000)
3.religion		-0.0641		-2.216
		(0.163)		(0.383)
_cons	-0.0517	-0.172	-27.46***	-19.34
	(0.711)	(0.469)	(0.001)	(0.153)
sigma				
_cons	0.0877***	0.173***	5.372***	9.965***
	(0.000)	(0.000)	(0.000)	(0.000)
N	2410	2192	2410	2192

education level, we recompute the local wealth by considering only those in the household's vicinity with the same value of the indicator for primary/higher-than-primary education-level of the household-head and then again with the same value for the agrarian/non-agrarian occupation indicator. Three types of vicinities thus implied - are summarised in Table 26. The first interpretation is one that we have already described in Section 5.4 - where a simple average of the total assets owned by the households in the consumer's district is used. Both education rank and an agrarian-occupation factor variable Ω indicating whether the household head's occupation is agrarian or not are used as controls in this interpretation of the vicinity. The second interpretation of the vicinity presented as r_{educ} in Table 27 is based on a supra-primary factor variable Γ indicating whether the education level of the household-head is higher than primary school or not. More specifically, this second interpretation uses the prosperity within the educational levels implied by supra-primary factor variable Γ (i.e. primary and higher-than primary educational levels). The agrarian/non-agrarian factor variable Ω is also used as the additional control in this second interpretation. Finally, the third interpretation of vicinity is based on the average prosperity in the agrarian vs non-agrarian occupations (i.e. Ω) and uses educational rank as the additional control variable. The results from this third interpretation of vicinity are presented in Table 28.

Observing the differences in coefficients for local wealth calculated over a supra-primary education vicinity (`localwealth_educ`) and over agrarian/non-agrarian vicinity (`localwealth_agri`) in Tables 27 and 28 (respectively) for Tanzania and Nigeria, we see again that while the local wealth corresponds to a rise in educational expenditure logarithm $log(x_{educ})$ for Nigeria, its effect is insignificant (whether one uses the local wealth in the supra-primary-educated neighbourhood `localwealth_educ` or local wealth in the agrarian-neighbourhood `localwealth_agri`) for Tanzania. Further, since the coefficients for secondary-age indicator `is_secondaryage` seem higher for the agrarian/non-agrarian vicinity (i.e. the coefficient for local wealth in `localwealth_agri`) than for the supra-primary education vicinity (i.e. local wealth in `localwealth_educ`), it can be argued that education expenses are more often clustered by agrarian or non-agrarian occupations than by education levels the populations have attained.

6 Conclusions

We had set ourselves to examine the extent to which educational expenses are influenced by local wealth and household permanent income in Nigeria and Tanzania. While there are idiosyncratic differences between the two countries we've used for our comparison, the weakening of effects of permanent income within high local wealth areas is a notable implication relevant to both the concerns for aspiration consumption and those of education policy in the region.

Since the agrarian consumers maintain a high budget share on education expenses in the poorer economy (Tanzania) and since higher education ranks correspond to higher education expenses in the more developed economy (Nigeria), one can argue that the educational expenses are likely to remain high as more of the population attains education as the regional prosperity rises. This also seems aligned with the recent finding from a cohort-study by Porzio et al. (2021) that the labour with skills added through education is less willing to stay in agricultural activities.

There are two main findings therefore that can be highlighted from the results in the chapter. First, the effects of local wealth are significant on education expenses in both the economies. Second, the lower wealth continues to prevent access to education and mobility in both the economies - even though rises in education expenses seem inevitable when the employment opportunities increase in an economy. In light of both the findings, the effects of permanent income as well as wealth-effects appear less extreme in the higher developed economy (i.e. Nigeria) than in the less developed economy (i.e. Tanzania). One can thus argue that the effects of permanent income appear to become weaker with rise in urban and economic development in the region. Considered as aspirational consumption, the educational expenses may thus continue to rise in the region with more local wealth accumulation. Such a rise associated with weaker effects of permanent income supports a view of conventional pathways to equity in both income and equality (Kuznets, 2019). Based on the arguments from the first chapter, both participation costs and income inequality may drive such changes.

From the view of education policy, the rise of education as a common need does not discount the need to improve access to education for the lower and middle-income populations in the region. While the literature notes a higher significance of wealth effects, the consideration of local wealth suggests that role of physical access may continue to be important in economies with wide disparities in urban development. This is of crucial importance also because the private educational institutions - which are more common in Nigeria - do not yet play a major role in the provision of education in the region. Instead, a more significant role continues to be played by the state-institutions in both the economies. The role of social differences indicated by the significance of social characteristics in education expenses for both the economies - further emphasises the need to improve wider access to education. These findings seem aligned with the wider observations made by

Alesina et al. (2019) who elaborate on the significant role of religious and colonial institutions in the educational mobility across sub-Saharan Africa.

In summary, the development of the education sectors seems to be at different stages of development in the Tanzania and Nigeria. The privatisation of education in Nigeria seems to have coincided with attendance of higher level education in a way that is yet to be seen for Tanzania - where education has not expanded as much yet. The effects of permanent income do seem strong in both the economies, but the relevance of local wealth in the population having higher education levels points to a promising albeit precarious direction where more urban developments could improve access to education and employment opportunities in the region while increasing expenses on education overall. Unlike the typical high quality consumption, therefore, the education expenses may continue to evolve - subject to participation costs and inequalities - as a an expensive common need for the masses.

7 Appendices

7.1 Education and Occupation mappings

Table 29: Education levels as ranks in Tanzania

	education code	education rank
PP	1	1
ADULT	2	2
D1	11	2
D2	12	2
D3	13	2
D4	14	2
D5	15	2
D6	16	2
D7	17	2
D8	18	3
OSC	19	3
MS COURSE	20	3
F1	21	3
F2	22	3
F3	23	3
F4	24	3
O COURSE	25	4
F5	31	4
F6	32	4
A COURSE	33	4
DIPLOMA	34	4
U1	41	4
U2	42	4
U3	43	4
U4	44	4
U5&	45	4

Table 30: Education levels as ranks in Nigeria

	education code	education rank
None	0	0
N1	1	1
N2	2	1
P1	11	2
P2	12	2
P3	13	2
P4	14	2
P5	15	2
P6	16	2
JS1	21	3
JS2	22	3
JS3	23	3
SS1	24	3
SS2	25	3
SS3	26	3
Lower 6	27	3
Upper 6	28	3
Teacher training	31	4
Vocational/Technical	32	4
Modern school	33	3
NCE	34	3
Poly/prof	41	4
1st degree	42	4
Higher degree	43	4
Quranic	51	3
Integrated Quaranic	52	4
Adult Education	61	3

Table 31: Occupations as ranks in Tanzania

code	occupation area	occupation rank
14	Student	0
13	Job Seeker	0
12	Paid Family Work	0
16	Unemployed	0
11	Unpaid Family Work	0
17	Unemployed (too young)	0
1	Livestock/Agriculture	1
2	Fishing	1
3	Mining	1
4	Tourism	1
7	Private Sector	2
9	Non-Agricultural (w Employer)	2
10	Non-Agricultural (w/o Employer)	2
8	Non-Government/Religious Org	3
5	Government	3
6	Parastatal	3

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
3461	decorators and commercial designers	1
7321	potters and related clay and abrasive formers	1
8251	printing machine operators	1
2144	electronic and telecommunications engineers	1
3139	other optical and electronics equipment controllers not elsewh	1
6122	poultry products	1
7432	weavers, knitters and other hand textile products makers	1
9111	street foods vendors	1
7424	basketry weavers, brush markers and related workers	1
3114	mechanical engineering technicians	1
3441	custom and border professionals	1
7331	handicraft workers in wood and related materials	1
7313	jewelry and precious metal trade workers	1
7413	food beverage testers and graders	1
8277	tea coffee cocoa and chocolate preparing and producing machine o	1
8152	cooking, roosting and related heat - treating plant operators	1
7224	metal grinder, polishers and tool sharpeners	1
8279	brewers, wine and other beverage machine operators	1
7412	bakers, pastry cooks and confectionery makers	1
8221	pharmaceutical and toiletry products machine operators	1
8263	sewing and knitting machine operators	1
7121	builders traditional materials	1
313	building construction labourers	1
2429	other legal professionals	1
1312	general managers in manufacturing	1
7332	handicraft workers in textile, leather and related materials	1
5149	other personal services workers not elsewhere classified	1
7442	shoe makers and related good workers	1
8151	crushing mixing and grinding equipment operators	1
7211	metal moulds and core makers	1
8284	metal, rubber and plastic products assemblers	1
7129	other building frames and related workers	1
9332	hand and pedal vehicle drivers	1
3340	other teaching associate professionals	1
5141	hairdressers, barbers, beauticians and related workers	1

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
9162	sweepers and related labourers	1
6114	mixed crop growers	1
5133	home-based personal care workers	1
6151	aquatic liege cultivation workers	1
9321	assembling labourers	1
3116	chemical engineering technicians	1
3421	trade brokers	1
8275	baked goods producing and cereals processing machine operators	1
8113	well drillers and borers and related workers	1
3213	farming and forestry advisers	1
7433	tailors, dress makers and hatters	1
3416	buyers	1
5230	stall and market salespersons	1
2332	pre-primary education teaching professionals	1
3131	photographers and image and sound-recording equipment controller	1
6141	forestry worker and loggers	1
7411	meat and fish butchers and preparers	1
8229	other chemical products machine operators	1
7123	concrete placers, concrete finishers and terrazzo-workers	1
8321	motorcycle drivers	1
1313	general managers in construction	1
7136	building and related electricians	1
3417	appraisers and values	1
5112	transport conductors	1
8123	metal heat - treating plant operators	1
8264	textile bleaching, dyeing and cleaning machine operators	1
2455	film, stage and related actors and directors	1
5121	house stewards and house keepers	1
5220	shop sales persons and demonstrators	1
6111	field crops and vegetable growers	1
6210	subsistence agricultural and fishery workers	1
1228	research and development managers	1
3444	government licensing officials	1
6152	inland and coastal waters fishery workers	1
7131	roofers	1
6121	dairy and livestock producers	1

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
6154	hunters and trappers	1
7222	tool maker, metal patter makers and metal makers	1
1130	traditional chiefs and head of villages	1
3229	other health associate professionals (except nursing)	1
1316	general managers in transportation	1
7213	sheet-metal workers	1
7422	cabinet makers and related workers	1
9213	fishery, hunting and tapping labourers	1
1229	other specialized managers	1
7214	structural metal prepares and erector	1
6130	market oriented crop and animal producers	1
8224	photographic products machine operators	1
7135	plumbers and pipe fitters	1
5210	fashion and other models	1
9322	hand packers and other manufacturing labourers	1
7122	bricklayers, stonemason and tile setters	1
2143	electrical engineers	1
9133	hand launderers and pressers	1
9131	domestice helpers and cleaners	1
7124	carpenter and jointers	1
8240	wood products machine operators	1
5139	other personal care workers	1
2359	other teaching professionals not elsewhere classified	1
8285	wood related materials products assemblers	1
9211	farmland and labourers	1
9212	forestry labourers	1
3442	government tax and excise officials	1
4133	transport clerks	1
7423	wood working machine setter operators	1
1315	general managers in resturants and hotels	1
3222	sanitarian	1
7322	glass formers, cutters grinder and finishers	1
9112	street vendors, other products	1
2146	chemical engineers	1
7221	blacksmiths, hammersmith's, forging-press workers	1
3227	veterinary assistants	1

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
8332	earth-moving and related machinery operators	1
9132	helpers and cleaners in offices and hotels and related workers	1
7435	textile patternmakers and cutters	1
9152	watchers and doorkeepers	1
1210	directors and chief executives	1
8239	other rubber and plastics machine operators	1
2113	chemists	1
5113	travel guides and ground hosts	1
9153	private security guards	1
2460	religion professionals	1
3212	agronomy and forestry technicians	1
6123	mixed animal producers	1
3449	other government associate professionals	1
9333	drivers and operators of animal-drawn vehicles and machinery	1
9141	building caretakers	1
1223	personel and industrial relations managers	1
6113	gardeners, horticultural; nursery growers	1
7243	radio and television service	1
7231	motor vehicle mechanics and filters	1
7437	upholsterers and related workers	1
7133	insulators	1
7241	electrical mechanics	1
7311	precision instrument makers	1
5122	waiters	1
2453	musicians	1
8269	textile machine operators	1
7312	musicians (acoustic)	1
8223	metal finishers	1
7324	ceramic painters	1
4144	scribes	1
3320	education specialists(1)	1
7341	type setters	1
3418	auctioneers	1
3122	computer equipment operators	1
7216	under-water workers	1
3223	dieticians and nutritionists	1

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
7345	textile printers	1
3224	optometrists	1
8312	railway workers	1
7436	embroiderers	1
8122	metal melters	1
3141	ship engineers	1
7344	bookbinders	1
3226	physiotherapists	1
7111	miners	1
8132	ceramic plant operators	1
8334	lift-truck operators	1
9120	shoe-cleaners	1
8159	chemical-plant operators	1
3228	pharma assistants	1
8282	elec and machinery assemblers	1
3151	building and fire inspectors	1
812	cement and materials processing machine operators	1
9161	garbage collectors	1
8143	paper plant operators	1
2133	computer programmers	2
6112	tree shrub crop growers	2
3221	medical assistants	2
4143	coding, proof-reading and related clerks	2
1314	general managers in retail and wholesale trade	2
3422	clearing and fowarding agents	2
7113	stone-splitters, cutters and carvers	2
2412	personnel and careers professionals	2
3465	athletes and related workers	2
4122	statistical and finance clerks	2
8262	weaving and knitting machine operators	2
3310	primary education teaching associate professionals	2
3415	technical and commercials sales representatives	2
2340	special education teaching professionals	2
8141	sawmill, wood panel and related wood-processing plant operat	2
7212	welders and flame-cutters	2
8322	cart, taxi and light van drivers	2

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
8323	bus and train drivers	2
9151	messengers package and luggage	2
4142	mail carriers and sorting clerks	2
8169	other power generating and related operators	2
1224	sales and marketing managers	2
8290	other stationery machine operators and assemblers	2
9142	windows cleaners	2
3411	securities, finance dealers and brokers	2
4141	library and filling clerks	2
7141	painters and paperhangers	2
8272	dairy products machine operators	2
7421	wood treaters	2
2145	mechanical engineers	2
2331	primary education teaching professionals	2
3118	other physical science and engineering technicians	2
8324	heavy truck drivers	2
2229	other health professionals (except nursing)	2
3429	other business services agent and trade brokers	2
3330	special education teaching associate professionals	2
3113	electrical engineering technicians	2
3439	other administrative associate professionals	2
8231	type making and vulcanizing machine operators	2
2446	social work professionals	2
3121	computer assistants	2
7112	short fires and blasters	2
4132	production clerks	2
8161	power-generating plant operators	2
7242	electronic fitters and services	2
3152	safety, health and quality inspectors (vehicles, processes	2
7434	fur tailor and related workers	2
1222	finance and administration managers	2
2230	nursing and midwifery professionals	2
3423	labour contractors and equipment agents	2
5143	undertakers and embalmers	2
6142	charcoal burners and related workers	2
7223	machine tool setter operators	2

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
9311	mining and related labourers	2
9312	construction and maintenance labourers road, dams and similar co	2
3414	travel consultants organisers	2
5131	institution-based personal care workers	2
3431	administrative and related associate professionals	2
4131	stock clerks	2
2139	other computing professionals	2
2452	sculptors, painters and related artists	2
3463	street, nightclub and related musicians, singers and dancers	2
2141	architects, town and traffic planners	2
1120	senior government officials	2
1226	supply and distribution managers	2
2211	biologists	2
3470	religion associate	2
2351	education specialists(2)	2
2441	economists	2
3132	broadcasting equipment controllers	2
3412	insurance representatives	2
1317	business managers	2
3145	air-traffic safety technicians	2
2422	judges	2
2147	mining engineers	2
2213	agronomists and related professionals	3
8153	filtering and separating equipment operators	3
3413	estate agents	3
1311	general managers in agriculture	3
2142	civil engineers	3
2320	secondary education teaching professionals	3
6153	deep-sea fishery workers	3
1141	senior officials of political party organisation	3
3445	commissioned police officers and detectives	3
2223	veterinarians	3
2421	lawyers	3
3143	aircraft pilot and related workers	3
1227	computing services managers	3
3419	other finance and sales associate professionals	3

Table 32: Occupations as ranks in Nigeria

code	occupation name	occupation rank
3450	social work associate professionals	3
2149	other architects, engineers and related professionals	3
3443	government welfare and pension officials	3
2352	school inspectors	3
3462	radio, television and other announcers	3
2411	accountants	3
2419	other business professionals	3
2451	authors, journalist and other writers	3
2122	statisticians	3
1318	general managers in personnel care, cleaning repairs and rel	3
2310	colleges, university and higher education teaching professiona	3
8155	petroleum refining plant operators	3
1221	production and operations managers	3
2224	pharmacists	3
2148	cartographers and surveyors	3
2221	medical doctors	3
3112	civil engineering technicians	3
5111	flight attendants and travel stewards	3
1110	legislators	3
1142	senior business officers	3
3432	legal professionals	3

Part V

Epilogue

The presented study had set out to reassess claims of rise in aspirational consumption with a view of changing consumers needs in the developing economies. With the theoretical implications of a short-term increase in aspirational investments under income inequalities and a varied role of household permanent income on food vs education expenses observed for SSA, we argue that aspirational consumption as investments towards future expected economic gains under an environment of inequalities may form an important part of the rise in consumption seen in developing economies that is not satisfactorily explained with the futility view of aspirational consumption. While the role of aspirational consumption in building social networks has been identified in both sociology and economics, empirical studies in economics have more often narrowed their focus to a visible appeal from aspirational consumption alone without considering the perceived future economic value from such consumption to the consumer. The presented study argues in favour of expectation-based approaches for analysis of aspirational consumption in the developing economies and urges for a more circumspect approach towards attribution of compensatory or excessive behaviour to the poor in particular.

The link between economic and social goals elaborated in the study serves to critique two commonly held views pertaining to empirical study of aspirational consumption. First is the generic view of aspirational goods in empirical studies for measurement of conspicuous consumption and second is the "pure-status" view that prevails more generally across empirical studies. The generic view of aspirational goods - treating all clothing, personal items etc. into a "conspicuous" category - not only ignores the quality variation in goods but in its attempt to achieve separability conditions for demand analysis, it also ends up treating all aspirational consumption as fundamentally futile. There are evident issues with such attributions across varied social contexts without a widely applicable criterion for maximal consumer needs - but the more serious limitations from the view are empirical. While pure-status view may be acceptable in settings connotating a rural or backward social fabric, the view faces theoretical concerns in relation to consumption (Moldovanu et al., 2007; Truyts, 2010) and poses an empirical limitation when combined with the visible consumption view. More particularly, an approach that treats social goals as independent of economic goals (vis-a-vis the pure-status view) while also failing to demarcate private and visible utility appropriately (e.g., with all clothing treated being of a conspicuous value relevant only for social goals) is susceptible to improper inferences of wasteful social tendencies. The unintended consequence of trying to fit consumption behaviours into assumed static social groups could - for example - be an inappropriate branding of certain social group as profligate or unwise (Burger et al., 2015). If such issues have been avoided in

empirical studies, then it is only through direct consumer surveys of status value of items within a particular social context (Heffetz, 2011) - a crucial part of the analysis that isn't often repeated under disparate social contexts. The views from Veblen (1899) seem to have thus outlasted more recent sociological research on status and prevailed in empirical studies since they go hand in hand with the futility assumption. Aspirational consumption can be safely assumed to offer only a futile value to the consumer when social goals - as the motivator of status consumption - have already been assumed separate from economic concerns. It is by acknowledging the more nuanced views on social goals offered by recent sociological research, that the presented study aims for a different approach based on a broader view of aspirational consumption.

The approach in the current study avoids limitations from both the views - first by concerning itself with the variation of high-quality or exclusive consumption without any *a priori* judgments of a generic aspirational basket of items and second, by considering an explicit link between economic and social goals for questions pertaining to effects of income inequalities on consumption (rather than adopting a pure-status view). The empirical chapters are careful to avoid any conclusions on profligate expenditure from consumption observed in food and education categories. In explaining how items of aspirational value could become a common need, the theoretical chapter focuses on a specific mechanism (i.e. small-scale competitions) for how aspirational consumption may benefit the consumer.

It is worth highlighting that many of such issues with studies in conspicuous consumption have already been highlighted in the literature. The theoretical challenges to the zero-sum view of pure-status prevalent in empirical studies are well-substantiated in the sociology literature (Wegener, 1992) while the empirical literature has also made attempts to incorporate status implication of asset-possession and income-classes (Nieftagodien et al., 2007; Burger et al., 2015). Instead of assigning higher status-signaling behaviour to specific social groups, Burger et al. (2015) - for example - have argued that the catch-up behaviour of less affluent groups in South Africa could be driven by historical asset deficits. The link between social and economic goals that we have argued for in the study fits within the attempts to incorporate such catch-up behaviours and aims to model the variation in segregation levels across different societies that may be ignored with a pure-status view.

There is no doubt that the objective reality of caste in India (Dugar, Bhattacharya, & Reiley, 2012) or race in US (Charles et al., 2009) makes a strong case for the pure-status view. A dynamics for status under such conditions is unlikely to exhibit any material change in social positions - thus allowing an empirical analysis to ignore mobility perceptions and focus on how social status - in a pure objective sense - influences consumer choices. Our contention in the presented study - on the other hand - has only been that the implications for objective notions of status may not offer a complete view without considering the subjective aspects of status that are difficult to sideline when discussing effects of status on individual decisions on aspirational consumption. While the role of race or caste is incontrovertible in

their respective contexts, it also makes sense to reconcile the subjective notions of such realities when considering consumption in an environment where higher-population density has changed lifestyles, consumerism has picked up an unprecedented degree of momentum and newer social identities are shaping up (Schroeder, 2012; Laher & Singoei, 2014).

The goal of empirical analyses in the study has been to examine the role of permanent income on consumption of aspirational value - a role that also helps in assessing future rise in aspirational consumption given that high-consumption being limited to richer households limits the rise in such consumption. That the role of household income is less important for education in urbanised areas than for high quality food can be viewed in light of the theoretical findings of the first chapter on how persistent household income inequalities influence demand for status-goods varying in their potential mobility value. More specifically, the assumption of social goals being tied with economic goals allows us to contrast the status value of high-quality food against that of education while considering their future economic benefits. It is after all more often that one's prestigious degrees are more rewarding than one's fine-dining experiences - even though both may have some likelihood of raising the chances of gaining a wealthier social/economic position through social interactions. In terms of the theoretical model from the first chapter, the expenditure on education offers higher returns than high food quality for a given participation cost.

Recall that the two equilibria in the long-run comprise of one where only the rich participate (Case I) and another where both rich and poor can participate (Case III). In case of a binding constraint when the poor are not in a position to participate, only the rich participate with their expenditures increasing with wider income differences. This seems aligned with the case for education in the poorer economy (Tanzania) as well as for high-quality food where participation costs (for a given return) seem beyond the means of the poor. With increased urbanisation, on the other hand, both the role of social capital becomes higher and the participation of poor increases due to lowered participation costs - resulting in an equilibrium where both rich and poor participate (Case III). The relation with inequality remains the same nevertheless in both types of equilibria so that high inequalities may be maintained with high enough participation costs. The increased income differences thus contribute to rise in expenditures in the latter case as well when lower participation costs allow both the rich and poor to participate. This seems aligned with empirical results for education expenditure in the urbanised regions of Nigeria.

We have argued that views on how inequality influences status consumption that focus primarily on notions of visibility may ignore the variation in mobility appeal from disparate status items. As the most primal necessity, food is a likely consumption category to exhibit significant quality differences in consumer markets. While a higher quality of food may help let a rich consumer in SSA distinguish herself from the poorer consumer (as is evidenced by sociological surveys), such quality competes with other newer consumption categories for signal-indication through consumption. With much fewer population in Tanzania having had

education and a much higher population having had education in Nigeria - it is not unlikely that education has a stronger bandwagon effect than food. But what a focus on the visible appeal from the newer consumption types - electronics, cosmetic products, fashionable items etc. - may fail to highlight is that consumer motivations towards social capital investments - expenditures such as elite sports, social activities, education etc. - may offer perceptions of future economic benefits to the consumer as well as a status value that neither food nor new fashion might. What the study argues is that such needs - which are also of aspirational value - can neither be considered basic needs of the consumer nor can they be explained by the dichotomy of private-public utility used in the visible consumption models. The dual purpose of this important element of aspirational consumption - of indicating a superiority in status while also increasing higher expected future wealth - may be better understood in an expected value framework where status gains are plausibly linked with economic gains to the consumer.

Part VI

Bibliography

References

Abay, K. A., Asnake, W., Ayalew, H., Chamberlin, J., & Sumberg, J. (2021). Landscapes of opportunity: Patterns of young people's engagement with the rural economy in sub-Saharan Africa. *The Journal of Development Studies*, *57*(4), 594–613.

Acar, E. Ö., Günalp, B., & Cilasun, S. M. (2016). An empirical analysis of household education expenditures in turkey. *International Journal of Educational Development*, *51*, 23–35.

Acker, J. R. (1980). Women and stratification: A review of recent literature. *Contemporary sociology*, *9*(1), 25–35.

Adler, I., & Kraus, V. (1985). Components of occupational prestige evaluations. *Work and Occupations*, *12*(1), 23–39.

Agada, M. O., Igbokwe, E. M., et al. (2015). Dietary diversity of rural households in north central Nigeria. *European Journal of Nutrition and Food Safety*, *5*(3), 150–5.

Alesina, A., Hohmann, S., Michalopoulos, S., & Papaioannou, E. (2019, February). *Intergenerational mobility in Africa* (Working Paper No. 25534). National Bureau of Economic Research.

Allais, S. (2017). Towards measuring the economic value of higher education: Lessons from South Africa. *Comparative Education*, *53*(1), 147–163.

Amegashie, J. A. (1999). The design of rent-seeking competitions: committees, preliminary and final contests. *Public Choice*, 63–76.

Andalon, M., & Gibson, J. (2017). 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*, *10765*.

Anderson, E., & McKay, A. (2004). Why is inequality so high, but also so variable in sub-Saharan Africa. *Poverty and Public Policy Group, Overseas Development Institute, London*.

Anderson, G., & Blundell, R. (1982, Nov). Estimation and hypothesis testing in dynamic singular equation systems. *Econometrica*, *50*(6).

Anderson, G., & Blundell, R. (1983, Jul). Testing restrictions in a flexible dynamic demand system: An application to consumers' expenditure in Canada. *The Review of Economic Studies*, *50*(3), 397-410.

Appleton, S., Bigsten, A., & Manda, D. (1999). *Educational expansion and economic decline: returns to education in Kenya, 1978-1995*. University of Oxford.

Arbache, J. S., & Page, J. (2009, 08). How Fragile Is Africa's Recent Growth? *Journal of African Economies*, *19*(1), 1-24.

Arimond, M., & Ruel, M. T. (2004). Dietary diversity is associated with child nutritional status: evidence from 11 demographic and health surveys. *The Journal of nutrition*, *134*(10), 2579–2585.

Asche, F. (1996). A system approach to the demand for salmon in the European Union. *Applied Economics*, *28*(1), 97–101.

Atsebi, J. M. B., & i Carbonell, A. F. (2019, October). *Relative deprivation in Tanzania* (Working Papers No. 1124). Barcelona Graduate School of Economics. Retrieved from `https://ideas.repec.org/p/bge/wpaper/1124.html`

Babalola, J. B. (1998). Cost and financing of university education in Nigeria. *Higher Education*, *36*(1), 43–66.

Barnett, J. M. (2005). Shopping for gucci on canal street: reflections on status consumption, intellectual property, and the incentive thesis. *Va. L. Rev.*, *91*, 1381.

Barr, N. (2004). Higher education funding. *Oxford review of economic policy*, *20*(2), 264–283.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407–443.

Barro, R. J. (1999). Human capital and growth in cross-country regressions. *Swedish economic policy review*, *6*(2), 237–277.

Barten, A. P. (1964, Jan-Apr). Consumer demand functions under conditions of almost additive preferences. *Econometrica*, *32*(1-2).

Barten, A. P. (1969). Maximum likelihood estimation of a complete system of demand equations. *European economic review*, *1*(1), 7–73.

Bayar, A. A., & İlhan, B. Y. (2016). Determinants of household education expenditures: Do poor spend less on education. *Topics in Middle Eastern and North African Economies*, *18*, 83–111.

Baye, M. R., Kovenock, D., & De Vries, C. G. (1993). Rigging the lobbying process: an application of the all-pay auction. *The american economic review*, *83*(1), 289–294.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, *87*(6), 1153–1189. Retrieved from `http://www.jstor.org/stable/1833328`

Behrman, J. R., & Knowles, J. C. (1999). Household income and child schooling in Vietnam. *The World Bank Economic Review*, *13*(2), 211–256.

Bekir, I., El Harbi, S., & Grolleau, G. (2013). How a luxury monopolist might benefit from the aspirational utility effect of counterfeiting? *European Journal of Law and Economics*, *36*, 169–182.

Benson, C. S. (1961). *The economics of public education*. Boston: Houghton Mifflin.

Booysen, F., Van Der Berg, S., Burger, R., Von Maltitz, M., & Du Rand, G. (2008). Using an

asset index to assess trends in poverty in seven sub-Saharan African countries. *World Development*, *36*(6), 1113–1130.

Bossuroy, T., & Cogneau, D. (2013). Social mobility in five African countries. *The Review of income and wealth*, *59*(S1), S84-S110.

Bourdieu, P. (2018). Cultural reproduction and social reproduction. In *Knowledge, education, and cultural change* (pp. 71–112). Routledge.

Brusco, S., & Lopomo, G. (2008). Budget constraints and demand reduction in simultaneous ascending-bid auctions. *The Journal of Industrial Economics*, *56*(1), 113–142.

Buchert, L. (1992). Basic and higher education in Tanzania 1919-1990: Quantitative expansion for the many or qualitative improvement for the few? *Eastern Africa Social Science Research Review*, *8*(1), 62.

Burger, R., Louw, M., Pegado, B. B. I. d. O., & van der Berg, S. (2015). Understanding consumption patterns of the established and emerging South African black middle class. *Development Southern Africa*, *32*(1), 41–56.

Burke, T. (1996). *Lifebuoy men, lux women*. Duke University Press.

Byun, S.-y. (2007). *Cultural capital and school success: The case of south Korea*. University of Minnesota.

Castro-Leal, F., Dayton, J., Demery, L., & Mehra, K. (1999). Public social spending in Africa: do the poor benefit? *The World Bank Research Observer*, *14*(1), 49–72.

Charles, K. K., Hurst, E., & Roussanov, N. (2009). Conspicuous consumption and race. *The Quarterly Journal of Economics*. (http://www.nber.org/papers/w13392)

Che, Y.-K., & Gale, I. (1996). Expected revenue of all-pay auctions and first-price sealed-bid auctions with budget constraints. *Economics Letters*, *50*(3), 373–379.

Che, Y.-K., & Gale, I. (1997). Rent dissipation when rent seekers are budget constrained. *Public Choice*, *92*, 109–126.

Che, Y.-K., & Gale, I. (1998). Standard auctions with financially constrained bidders. *The Review of Economic Studies*, *65*(1), 1–21.

Chernichovsky, D. (1985). Socioeconomic and demographic aspects of school enrollment and attendance in rural Botswana. *Economic development and cultural change*, *33*(2), 319–332.

Clementi, F., Fabiani, M., & Molini, V. (2019, 03). The Devil is in the Detail: Growth, Inequality and Poverty Reduction in Africa in the Last Two Decades. *Journal of African Economies*, *28*(4), 408-434.

Cole, H. L., Mailath, G. J., & Postlewaite, A. (1992). Social norms, savings behavior, and growth. *Journal of Political Economy*, *100*(6), 1092–1125. Retrieved from `http://www.jstor.org/stable/2138828`

Coleman, J. (1990). *Foundations of social theory*. Harvard University Press.

Cooray, A., & Potrafke, N. (2011). Gender inequality in education: Political institutions or culture and religion? *European Journal of Political Economy*, *27*(2), 268-280.

Cox, T. L., & Wohlgenant, M. K. (1986, November). Prices and quality effects in cross-sectional demand analysis. *American Journal of Agricultural Economics*, *68*(4), 908-919.

Cramer, J. S. (1973, Jun). Interaction of income and price in consumer demand. *International Economic Review*, *14*(2), 351-363.

Dabalen, A., Oni, B., & Adekola, O. A. (2001). Labor market prospects for university graduates in Nigeria. *Higher Education Policy*, *14*(2), 141–159.

Deaton, A. (1988, June). Quality, quantity, and spatial variation of price. *The American Economic Review*, *78*(3), 418-430.

Deaton, A. (1993). *Understanding consumption*. Clarendon Press.

Deaton, A. (1999). Commodity prices and growth in Africa. *Journal of economic Perspectives*, *13*(3), 23–40.

Deaton, A., & Muellbauer, J. (1980a, Jun). An almost ideal demand system. *The American Economic Review*, *70*(3), 312-326.

Deaton, A., & Muellbauer, J. (1980b). *Economics and consumer behavior*. Cambridge University Press.

De Graaf, N. D., De Graaf, P. M., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of the cultural capital perspective. *Sociology of education*, 92–111.

Dhami, S. (2016). *Foundations of behavioural economics*. Oxford University Press.

Donkoh, S., & Amikuzuno, J. (2011, August). The determinants of household education expenditure in Ghana. *Educational Research and Reviews*, *6*(8), 570–579.

Doyal, L., & Gough, I. (1991). *A theory of human need*. Macmillan Publishers.

Duflo, E., & Banerjee, A. (2011). *Poor economics* (Vol. 619). PublicAffairs New York, NY, USA.

Dugar, S., Bhattacharya, H., & Reiley, D. (2012). Can't buy me love? a field experiment exploring the trade-off between income and caste-status in an Indian matrimonial market. *Economic Inquiry*, *50*(2), 534–550.

Ebaidalla, E. M. (2018). Understanding household education expenditure in sudan: do poor and rural households spend less on education? *African Journal of Economic Review*, *6*(1), 160–178.

Falk, A., & Knell, M. (2004). Choosing the joneses: Endogenous goals and reference standards. *The Scandinavian Journal of Economics*, *106*(3), 417–435.

Filmer, D., & Pritchett, L. (1999). The effect of household wealth on educational attainment: Evidence from 35 countries. *Population and Development Review*, *25*(1), 85-120.

Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data - or tears: an application to educational enrollments in states of India. *Demography*, *38*(1), 115–132.

Fisher, F. M., & Shell, K. (1968). *Value, capital and growth* (J. N. Wolfe, Ed.). Transaction

Publishers.

Frank, R. H. (1993). *Choosing the right pond*. OUP USA.

Friedman, M. (1957). *A theory of the consumption function*. National Bureau of Economic Research, Inc.

Friehe, T., & Mechtel, M. (2014). Conspicuous consumption and political regimes: Evidence from east and west Germany. *European Economic Review*.

Fujii, E., Khaled, M., & Mak, J. (1985). The exportability of hotel occupancy and other tourist taxes. *National Tax Journal*, *38*(2), 169–177.

Fuller, W. P. (1976). More evidence supporting the demise of pre-employment vocational trade training: a case study of a factory in India. *Comparative Education Review*, *20*(1), 30–41.

Galbraith, J. K. (1958). *The affluent society*.

Galbraith, J. K. (1994). Beyond contentment. *The Telegraph*, *14*.

Galbraith, J. K. (2001). *The Essential Galbraith* (A. D. Williams, Ed.). Houghton Mifflin Company.

Geda, A. (2019). *The historical origin of the African economic crisis: From colonialism to China*. Cambridge Scholars Publishing.

Gertler, P., & Glewwe, P. (1990). The willingness to pay for education in developing countries: Evidence from rural Peru. *Journal of Public Economics*, *42*(3), 251-275.

Gibson, J., & Kim, B. (2019). Quality, quantity, and spatial variation of price: Back to the bog. *Journal of Development Economics*, *137*, 66-77.

Glewwe, P., & Jacoby, H. G. (2004). Economic growth and the demand for education: Is there a wealth effect? *Journal of development Economics*, *74*(1), 33–51.

Goldthorpe, J. H. (2000). *On sociology: Numbers, narratives, and the integration of research and theory*. Oxford University Press, USA.

Goldthorpe, J. H., & Hope, K. (1974). *The social grading of occupations. a new scale and approach*. Clarendon Press.

Goode, W. J. (1978). *The celebration of heroes: Prestige as a social control system*. Univ of California Press.

Gorman, W. M. (1953, January). Community preference fields. *Econometrica*, *21*(1), 63-80.

Gould, E. D., Simhon, A., & Weinberg, B. A. (2020). Does parental quality matter? evidence on the transmission of human capital using variation in parental influence from death, divorce, and family size. *Journal of Labor Economics*, *38*(2), 569–610.

Guillen-Royo, M. (2011). Reference group consumption and the subjective wellbeing of the poor in Peru. *Journal of Economic Psychology*, *32*(2), 259–272.

Hailu, S., & Woldemichael, B. (2019). Dietary diversity and associated factors among pregnant women attending antenatal care at public health facilities in Bale zone, southeast Ethiopia. *Nutrition and Dietary Supplements*, *11*, 1–8.

Haller, M., Jowell, R., & Smith, T. W. (2009). *The international social survey programme*

1984-2009: Charting the globe. Routledge.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, *60*(1), 126-142.

Harrod, R. F. (1958). *The possibility of economic satiety - use of economic growth for improving the quality of education and leisure*. Committee for Economic Development.

Hashimoto, K., & Heath, J. A. (1995). Income elasticities of educational expenditure by income class: The case of Japanese households. *Economics of education review*, *14*(1), 63–71.

Hatløy, A., Hallund, J., Diarra, M. M., & Oshaug, A. (2000). Food variety, socioeconomic status and nutritional status in urban and rural areas in Koutiala (Mali). *Public health nutrition*, *3*(1), 57–65.

Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *The Review of Economics and Statistics*, *XCIII*(4).

Heien, D., & Wessells, C. R. (1990, Jul). Demand systems estimation with microdata: A censored regression approach. *Journal of Business and Economic Statistics*, *8*(3), 365-371.

Heyneman, S. P. (1980a). Differences between developed and developing countries: Comment on Simmons and Alexander's" determinants of school achievement". *Economic development and cultural change*, *28*(2), 403–406.

Heyneman, S. P. (1980b). Investment in Indian education: uneconomic? *World development*, *8*(2), 145–163.

Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International journal of educational development*, *23*(3), 315–337.

Heyneman, S. P., & Lee, B. (2014). The impact of international studies of academic achievement on policy and research. *Handbook of international large-scale assessment: Background, technical issues and methods of data analysis*, 37–72.

Hicks, J. R. (1946). Value and capital, 1939. *Mathematical Appendix*, *311312*.

Hirsch, F. (1977). *Social limits to growth*. Routledge and Kegan Paul Ltd.

Hoddinott, J., & Yohannes, Y. (2002). *Dietary diversity as a food security indicator* (Tech. Rep.).

Houthakker, H. S. (1952). Compensated changes in quantities and qualities consumed. *The Review of Economic Studies*, *19*(3), 155-164.

Howe, L. D., Hargreaves, J. R., & Huttly, S. R. (2008). Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerging Themes in Epidemiology*, *5*(3).

Howe, L. D., Hargreaves, J. R., Ploubidis, G. B., De Stavola, B. L., & Huttly, S. R. A. (2010, 09). Subjective measures of socio-economic position and the wealth index: a comparative analysis. *Health Policy and Planning*, *26*(3), 223-232.

Ilie, S., & Rose, P. (2018). Who benefits from public spending on higher education in south Asia and sub-Saharan Africa? *Compare: A Journal of Comparative and International Education*, *48*(4), 630–647.

Ireland, N. J. (1994). On limiting the market for status signals. *Journal of Public Economics*.

Jaikumar, S., & Sarin, A. (2015). Conspicuous consumption and income inequality in an emerging economy: evidence from India. *Marketing Letters*.

Jaikumar, S., Singh, R., & Sarin, A. (2018). "I show off, so I am well off": Subjective economic well-being and conspicuous consumption in an emerging economy. *Journal of Business Research*, *86*, 386–393.

Jee-Peng, T., Jimenez, E., & Psacharopoulos, G. (1986). *Financing education in developing countries: an exploration of policy options*. The World Bank.

Jin, Y., Li, H., & Wu, B. (2011). Income inequality, consumption, and social-status seeking. *Journal of Comparative Economics*, *39*(2), 191–204.

Jorgenson, D. W., Slesnick, D. T., & Stoker, T. M. (1988, Jul). Two-stage budgeting and exact aggregation. *Journal of Business & Economic Statistics*, *6*(3), 313-325.

K'akumu, O. A., & Olima, W. H. (2007). The dynamics and implications of residential segregation in Nairobi. *Habitat International*, *31*(1), 87–99.

Kambon, A., & Busby, L. (2000). *Education and its impact on poverty: Equity or exclusion*. Office of the UNESCO Representative.

Kaplinsky, R., & Morris, M. (2016). Chinese FDI in sub-Saharan Africa: engaging with large dragons. In *The power of the chinese dragon* (pp. 123–148). Springer.

Katsillis, J., & Rubinson, R. (1990). Cultural capital, student achievement, and educational reproduction: The case of Greece. *American sociological review*, 270–279.

Kaus, W. (2013). Conspicuous consumption and "race": Evidence from South Africa. *Journal of Development Economics*, *100*(1), 63-73. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0304387812000648`

Kay, J. A., & King, M. A. (2020). *Radical uncertainty*. Bridge Street Press Decision-making beyond the numbers.

Kennedy, G., Ballard, T., & Dop, M. C. (2011). *Guidelines for measuring household and individual dietary diversity*. Food and Agriculture Organization of the United Nations.

Khamis, M., Prakash, N., & Siddique, Z. (2012). Consumption and social identity: Evidence from India. *Journal of Economic Behavior and Organization*, *83*(3), 353-371. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0167268112001345`

Knell, M. (1999). Social comparisons, inequality, and growth. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift fÃŒr die gesamte Staatswissenschaft*, *155*(4), 664–695.

Knight, J., & Shi, L. (1996). Educational attainment and the rural-urban divide in China. *Oxford Bulletin of Economics and Statistics*, *58*(1), 83–117.

Kosec, K., Ghebru, H., Holtemeyer, B., Mueller, V., & Schmidt, E. (2018). The effect of land access on youth employment and migration decisions: Evidence from rural Ethiopia. *American Journal of Agricultural Economics*, *100*(3), 931-954.

Kuznets, S. (2019). Economic growth and income inequality. In *The gap between rich and poor* (pp. 25–37). Routledge.

Laher, R., & Singoei, K. (Eds.). (2014). *Indigenous people in Africa: Contestations, empowerment and group rights*. Africa Institute of South Africa.

Languille, S. (2019). The politics of the education budget: Financing mass secondary education in Tanzania (2004–2012). *International Journal of Educational Development*, *66*, 96–104.

Lazaridis, P. (2003). Household meat demand in Greece: a demand systems approach using microdata. *Agribusiness: An International Journal*, *19*(1), 43–59.

Li, G., Song, H., & Witt, S. F. (2004). Modeling tourism demand: A dynamic linear AIDS approach. *Journal of travel research*, *43*(2), 141–150.

Linssen, R., Van Kempen, L., & Kraaykamp, G. (2011). Subjective well-being in rural India: The curse of conspicuous consumption. *Social indicators research*, *101*, 57–72.

Manski, C. F. (2000). Economic analysis of social interactions. *The Journal of Economic Perspectives*, *14*(3), 115-136.

Manyanga, T., Tremblay, M. S., Chaput, J.-P., Katzmarzyk, P. T., Fogelholm, M., Hu, G., … others (2017). Socioeconomic status and dietary patterns in children from around the world: different associations by levels of country human development? *BMC Public Health*, *17*(1), 1–11.

Mayen, A.-L., Marques-Vidal, P., Paccaud, F., Bovet, P., & Stringhini, S. (2014). Socioeconomic determinants of dietary patterns in low-and middle-income countries: a systematic review. *The American journal of clinical nutrition*, *100*(6), 1520–1531.

McDonald, C. M., McLean, J., Kroeun, H., Talukder, A., Lynd, L. D., & Green, T. J. (2015). Correlates of household food insecurity and low dietary diversity in rural Cambodia. *Asia Pacific journal of clinical nutrition*, *24*(4), 720–730.

McKelvey, C. (2011). Price, unit value, and quality demanded. *Journal of Development Economics*, *95*, 157-169.

McKenzie, D. J. (2005). Measuring inequality with asset indicators. *Journal of Population Economics*, *18*(2), 229-260.

Mekuria, G., Wubneh, Y., & Tewabe, T. (2017). Household dietary diversity and associated factors among residents of finote selam town, north west Ethiopia: a cross sectional study. *BMC nutrition*, *3*(1), 1–6.

Mijs, J. J. B. (2019, 01). The paradox of inequality: income inequality and belief in meritocracy go hand in hand. *Socio-Economic Review*, *19*(1), 7-35.

Minnis, J. R. (2006). Nonformal education and informal economies in sub-Saharan Africa: Finding the right match. *Adult Education Quarterly*, *56*(2), 119–133.

M'Kaibi, F. K., Steyn, N. P., Ochola, S. A., & Du Plessis, L. (2017). The relationship between agricultural biodiversity, dietary diversity, household food security, and stunting of children in rural Kenya. *Food science & nutrition*, *5*(2), 243–254.

Moav and, O., & Neeman, Z. (2012). Saving rates and poverty: The role of conspicuous consumption and human capital. *The Economic Journal*, *122*(563), 933–956.

Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data. *Franco Modigliani*, *1*(1), 388–436.

Modjadji, P., Molokwane, D., & Ukegbu, P. O. (2020). Dietary diversity and nutritional status of preschool children in north west province, South Africa: A cross sectional study. *Children*, *7*(10), 174.

Moldovanu, B., Sela, A., & Shi, X. (2007). Contests for status. *Journal of political Economy*, *115*(2), 338–363.

Mouw, T. (2006). Estimating the causal effect of social capital: A review of recent research. *Annual review of sociology*, 79–102.

Muellbauer, J. (1975, October). Aggregation, income distribution and consumer demand. *The Review of Economic Studies*, *42*(4), 525-543.

Ndyali, L. (2016). Higher education system and jobless graduates in Tanzania. *Journal of Education and Practice*, *7*(4), 116–121.

Nelson, J. A. (1991, November). Quality variation and quantity aggregation in consumer demand for food. *American Journal of Agricultural Economics*, *73*(4), 1204-1212.

Ngware, M. W., Onsomu, E. N., Muthaka, D. I., & Manda, D. K. (2006). Improving access to secondary education in Kenya: what can be done? *Equal Opportunities International*.

Nieftagodien, S., Van Der Berg, S., et al. (2007). Consumption patterns and the black middle class: The role of assets. *Bureau of Economic Research*, *2*, 1–10.

Nti, C. A. (2011). Dietary diversity is associated with nutrient intakes and nutritional status of children in Ghana. *Asian Journal of Medical Sciences*, *2*(2), 105–109.

Nussbaum, M. (2003). Capabilities as fundamental entitlements: Sen and social justice. *Feminist Economics*, *9*(2-3), 33-59.

Obasuyi, F. O. T., & Rasiah, R. (2019). Addressing education inequality in sub-Saharan Africa. *African Journal of Science, Technology, Innovation and Development*, *11*(5), 629–641.

Obayelu, O. A., & Osho, F. R. (2020). How diverse are the diets of low-income urban households in Nigeria? *Journal of Agriculture and Food Research*, *2*, 100018.

Ogechi, U. P., & Chilezie, O. V. (2017). Assessment of dietary diversity score, nutritional status and socio-demographic characteristics of under-5 children in some rural areas of Imo state, Nigeria. *Malaysian Journal of Nutrition*, *23*(3).

Ohna, I., Kaarhus, R., & Kinabo, J. (2012). No meal without ugali? social significance of food

and consumption in a Tanzanian village. *Culture, Agriculture, Food and Environment*, *34*(1), 3–14.

Ohna, I. L. (2007). Food, culture and livelihood in Malinzanga village, Tanzania. *Department of International Environment and Development Studies*.

Oketch, M. (2016). Financing higher education in sub-Saharan Africa: some reflections and implications for sustainable development. *Higher Education*, *72*(4), 525–539.

Omori, M., & Smith, D. T. (2015). The impact of socioeconomic status and metropolitan area racial composition on visible consumption among whites and racial/ethnic minorities. *Race Soc Probl*(7), 169-180.

Onuorah, C., & Ayo, J. (2003). Food taboos and their nutritional implications on developing nations like Nigeria–a review. *Nutrition & Food Science*.

Owusu, G., & Agyei-Mensah, S. (2011). A comparative study of ethnic residential segregation in Ghana's two largest cities, Accra and Kumasi. *Population and Environment*, *32*(4), 332–352.

Parco, J. E., Rapoport, A., & Amaldoss, W. (2005). Two-stage contests with budget constraints: An experimental study. *Journal of Mathematical Psychology*, *49*(4), 320–338.

Paris, Q., & Caracciolo, F. (2014). Testing the adding up condition in demand systems. *Available at SSRN 2443516*.

Parsons, T. (1963). On the concept of influence. *Public opinion quarterly*, *27*(1), 37–62.

Parsons, T. (1971). *The system of modern societies*. Prentice-Hall Englewood Cliffs, NJ.

Pingali, P., & Sunder, N. (2017). Transitioning toward nutrition-sensitive food systems in developing countries. *Annual Review of Resource Economics*, *9*, 439–459.

Pollak, R. A., & Wales, T. J. (1981, Nov). Demographic variables in demand analysis. *Econometrica*, *49*(6), 1533-1551.

Pollard, D. (1982). A central limit theorem for k-means clustering. *The Annals of Probability*, *10*(4), 919–926.

Porzio, T., Rossi, F., & Santangelo, G. V. (2021). *The human side of structural transformation* (Tech. Rep.). National Bureau of Economic Research.

Postlewaite, A. (1998). The social basis of interdependent preferences. *European economic review*, *42*(3-5), 779–800.

Pottier, J., et al. (1999). *Anthropology of food: the social dynamics of food security.* Polity Press.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics*, *12*(2), 111–134.

Roychowdhury, P. (2016). Visible inequality, status competition, and conspicuous consumption: evidence from rural India. *Oxford Economic Papers*, *69*(1), 36–54.

Ruel, M. T. (2003, 11). Operationalizing Dietary Diversity: A Review of Measurement Issues and Research Priorities. *The Journal of Nutrition*, *133*(11), 3911S-3926S.

Schroeder, R. A. (2012). *Africa after apartheid : South Africa, race, and nation in Tanzania*. Indiana University Press.

Sen, A. (1978). *Three notes on the concept of poverty* (ILO Working Papers No. 991757103402676). International Labour Organization. Retrieved from https://ideas.repec.org/p/ilo/ilowps/991757103402676.html

Sen, A. (1985). A sociological approach to the measurement of poverty: A reply to Professor Peter Townsend. *Oxford economic papers*, *37*(4), 669-676.

Sen, A. (1992). *Inequality reexamined*. Oxford University Press.

Shils, E. (1969). Reflections on deference. *Politics, Personality and Social Sciences in the Twentieth Century: Essays in Honor of Harold D. Lasswell*, 297–346.

Silverstein, M. J., Fiske, N., & Butman, J. (2008). *Trading up: why consumers want new luxury goods–and how companies create them*. Penguin.

Sobel, J. (2002). Can we trust social capital? *Journal of economic literature*, *40*(1), 139–154.

Söderbom, M., Teal, F., Wambugu, A., & Kahyarara, G. (2006). The dynamics of returns to education in Kenyan and Tanzanian manufacturing. *Oxford Bulletin of Economics and Statistics*, *68*(3), 261–288.

Spaull, N. (2013). Poverty & privilege: Primary school inequality in South Africa. *International Journal of Educational Development*, *33*(5), 436–447.

Srivastava, A., Mukherjee, S., & Jebarajakirthy, C. (2020). Aspirational consumption at the bottom of pyramid: A review of literature and future research directions. *Journal of Business Research*, *110*, 246–259.

Stein, W. E., & Rapoport, A. (2005). Symmetric two-stage contests with budget constraints. *Public Choice*, *124*(3-4), 309–328.

Subrahmanyan, S., & Tomas Gomez-Arias, J. (2008). Integrated approach to understanding consumer behavior at bottom of pyramid. *Journal of Consumer Marketing*, *25*(7), 402–412.

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

Tansel, A. (1997). Schooling attainment, parental education, and gender in Cote d'Ivoire and Ghana. *Economic Development and Cultural Change*, *45*(4), 825–856.

Thomas, V., Wang, Y., & Fan, X. (1999). *Measuring education inequality: Gini coefficients of education*. The World Bank.

Tilak, J. B. (2002). Education and poverty. *Journal of human development*, *3*(2), 191–207.

Tilak, J. B. (2007). Post-elementary education, poverty and development in India. *International journal of educational development*, *27*(4), 435–445.

Tilak, J. B., et al. (2002). *Determinants of household expenditure on education in rural India* (No. 88). Citeseer.

Townsend, P. (1985). A sociological approach to the measurement of poverty - a rejoinder to Professor Amartya Sen. *Oxford economic papers*, *37*(4), 659-668.

Truyts, T. (2010). Social status in economic theory. *Journal of Economic Surveys*, *24*(1), 137–169.

Tullock, G. (1967). The welfare costs of tariffs, monopolies, and theft. *Economic inquiry*, *5*(3), 224–232.

Veblen, T. (1899). *Theory of the leisure class*.

Weber, M. (1925). *Wirtschaft und Gesellschaft*. JCB Mohr.

Wegener, B. (1992). Concepts and measurement of prestige. *Annual review of sociology*, *18*(1), 253–280.

Weismantel, M. J., et al. (1989). *Food, gender, and poverty in the Ecuadorian Andes.* University of Pennsylvania Press.

Wenban-Smith, H., Fasse, A., & Grote, U. (2016). Food security in Tanzania: The challenge of rapid urbanisation. *Food security*, *8*(5), 973–984.

Wiggins, S., & Proctor, S. (2001). How special are rural areas? the economic implications of location for rural development. *Development policy review*, *19*(4), 427–436.

Wrong, D. (1979). *Power: Its forms, bases and uses*. Routledge.

Wu, F., Zhang, D., & Zhang, J. (2008). Unequal education, poverty and low growth-a theoretical framework for rural education of China. *Economics of Education Review*, *27*(3), 308-318.

Yamamoto, Y., & Brinton, M. C. (2010). Cultural capital in east Asian educational systems: The case of Japan. *Sociology of Education*, *83*(1), 67–83.

Yan, G., Peng, Y., Hao, Y., Irfan, M., & Wu, H. (2021). Household head's educational level and household education expenditure in China: The mediating effect of social class identification. *International Journal of Educational Development*, *83*, 102400.

Zellner, A., & Theil, H. (1992). Three-stage least squares: Simultaneous estimation of simultaneous equations. In B. Raj & J. Koerts (Eds.), *Henri theil's contributions to economics and econometrics: Econometric theory and methodology* (pp. 147–178). Dordrecht: Springer Netherlands.