THREE ESSAYS ON THE EVOLVING AGRIFOOD SYSTEM IN TANZANIA

By

Christine Marie Sauer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics – Doctor of Philosophy

2021

ABSTRACT

THREE ESSAYS ON THE EVOLVING AGRIFOOD SYSTEM IN TANZANIA

By

Christine Marie Sauer

Broadly, my dissertation focuses on changes in the midstream and downstream of the agrifood value chain in Tanzania. The first essay examines the patterns and determinants of household-level consumption expenditure on processed food and meals away from home. I use a detailed food consumption diary from Tanzania to explore the relationship between the budget share spent on more convenient foods, such as highly processed food and food away from home, and income levels. Additionally, I use (i) geo-spatial data to analyze how these relationships change over space, and (ii) detailed labor data to analyze the correlation between men's and women's non-farm labor force participation and the budget share spent on higher value-added foods.

In my second essay, I revisit the old debate of whether the poor pay more for food, using the same spatial and food diary data as in the first essay. I find that, surprisingly, the poor generally are not more likely to buy in smaller quantities, the rich are not more likely to buy non-perishables in larger quantities, and that bulk discounts are modest at best for most food products we study. Most intriguingly, we find that the poor do not pay more than richer households.

Finally, my third essay uses primary data from maize flour retailers to explore the modernization of the maize flour value chain in Tanzania. I use various measures of value chain structure, conduct, and performance, and I disaggregate by retail type (traditional shops, transitional mini-supermarkets, and modern supermarkets) and town size, to study where changes are occurring. I find a rapid proliferation of maize flour brands, a move toward disintermediation

(especially in the secondary cities) and longer supply chains, and an emerging adoption of mobile money by traditional shops in smaller towns. These findings point toward a supply chain in flux.

ACKNOWLEDGMENTS

The PhD journey cannot be completed in isolation. Thank you to my advisor, Tom Reardon, for guiding me through the process. From your mentoring I have honed my critical thinking skills, nurtured my curiosity, and overall become a strong and disciplined researcher. Thank you also to Dave Tschirley, for your guidance and for funding my research assistantship and my many trips to Tanzania. The on-the-ground fieldwork was indispensable to my training as a researcher, and I am grateful for the opportunity to repeatedly visit such a beautiful country. Thank you to my other committee members, Nicky Mason-Wardell and Jeff Wooldridge, for helping me see my research in a different light and providing helpful comments along the way. For general logistics support, I extend my gratitude to Ashleigh Booth and Nancy Creed, without whom I am sure the AFRE department would descend into chaos.

One of my favorite aspects of a graduate program is the friends you make along the way. The Toucans helped me through econometrics, micro, and all other academic and life challenges since then. April Athnos, Sudha Kannan, Laura Medwid, Aakanksha Melkani, and Edouard Mensah – thank you for your steadfast friendship and support. One of my best memories from my time at MSU is traveling to Hangzhou, China to compete in the 2019 IFAMA Graduate Case Study Competition – and taking second place! Simone Faas, Salim Nuhu, and Ziwei Ye – you guys are the best teammates. I am so glad I got to travel with you and get to know you better in both China and (for Simone) Thailand. Many other friends in classes and around Cook Hall helped me get through this program – thank you to Abubakr Ayesh, Rahul Dhar, Ramyani Mukhopadhyay, Sam Padilla, Charuta Parkhi, and Danielle Ufer.

I would also like to thank my family members and friends who always listened to my rants and helped me to stay grounded: my parents and step-parents, my grandmothers, my aunt Laura, and Nick. Thank you to Rachael for always being a shoulder I could lean on and a fellow foodie with whom I can share the best food. To Jill and Chloe, thank you for sharing your home and being great quarantine buddies.

And finally, thank you to my dude Özkan Binici, for all the laughs, love, and pide. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ESSAY 1: CONSUMPTION OF PROCESSED FOOD & FOOD AWAY FROM HOME IN BIG CITIES, SMALL TOWNS, AND RURAL AREAS OF TANZANIA

1.1 Introduction

Over the past century there has been rapid growth of demand by consumers for purchased-processed food (as opposed to home-processed food) and for prepared "meals away from home" (MAFH). The growth started in developed regions in the late 1800s and increased over the 20th century. The demand for "convenience foods" was driven on the demand side in large part by the rise of the opportunity cost of women's time as women took up employment outside the home (Kinsey 1983). This was consistent with the theory of Mincer (1963) and Becker (1965) of consumption decisions being affected by own value of time.

In the past half century in Africa, Asia, and Latin America, there has been a parallel increase in consumption of processed food. In general, this has been driven by factors similar to those in developed countries: the rise of opportunity costs of time to home-process and home-prepare food for women and to return home to eat by men and women. These costs are correlated with incomes and lifestyle and employment changes brought by urbanization (Senauer et al. 1986 for Sri Lanka), as well as, we posit, the rise of rural nonfarm employment among women and men.

Women first sought low-processed products such as purchased-flour or milled rice (such as in Burkina Faso, Reardon et al., 1989) to avoid hand-pounding of grain, which was taking around 4 hours a day per woman in Africa in the 1980s (Barrett and Browne 1994). Households next sought high-processed products, at first unpackaged then eventually packaged, ready-to-eat products such as bread in Sri Lanka (Senauer et al. 1986) or Kenya (Kennedy and Reardon 1994) or heat and eat, such as *enjera* in Ethiopia (Minten et al. 2016). The purchase of processed foods and meals away from home had grown substantial by the 2000s in Asia (Pingali 2006) and 2010s in Africa (Tschirley et al. 2015a).

A subset of the high-processed products increasingly came to include ultra-processed foods such as sugar-sweetened beverages (SSBs) and snacks high in sugars and salt and fats. The latter in particular alarmed nutritionists. Popkin (1994) highlighted the health risks (including obesity and increased risk of non-communicable diseases (NCDs) such as heart disease) of a “nutrition transition” in developing regions into consumption of ultra-processed foods. Alarm over this transition spread in Africa in the past decades as research showed that overweight and obesity are rapidly increasing (Popkin et al. 2019). Some survey-based studies in Africa showed that ultra-processed food intake increases obesity, and that some poor consumers seek their macronutrients such as calories and proteins by consuming ultra-processed foods (Khonje et al. 2020).

In Africa, the literature on processed food consumption has focused mainly on urban areas. In general, the latter have been treated as homogeneous, with the exception of recent research on processed food in secondary/tertiary cities per se (Demmler et al. 2018, Kimenju et al. 2015, Rischke et al. 2015) or on primary cities per se (Khonje and Qaim 2019). To our knowledge there has been no systematic analysis over city types in one country for all categories of processed food and MAFH.

However, we would expect the consumption of processed food and MAFH to differ over city size, controlling for household income and other characteristics. The meso-level density of micro characteristics such as the agglomeration of purchasing power may affect household choices. Cities of different sizes may differ in the density and supply of prepared food vendors. Larger cities entail longer commutes and more inducement to save time returning to home to eat; such congestion is a manifestation of lifestyle factors (noted by Rischke et al. 2015, for Kenya).

Different city sizes may also correlate with different food environments and advertising contexts that influence consumption of convenience and "junk" foods.

Moreover, when processed food consumption research in Africa has included rural areas, it has largely treated them as one homogeneous group (e.g. Tschirley et al. 2015a) and has not separately treated MAFH. Nor has there been research on the effect of distance from urban areas on rural uptake of processed foods, including ultra-processed and MAFH; yet urban distance has been studied with respect to other rural behavior such as farm input use and intensification (e.g. Vandercasteelen et al. 2018 for Ethiopia). However, we expect distance from cities to affect processed and prepared food purchase for reasons parallel to those for differences in city size, with respect to access to these foods, lifestyle differences, and food environment variation.

Finally, research on the determinants of consumption of the different categories of processed food and MAFH has dwelt largely on incomes and demographic factors, including recent migration experience (Cockx et al. 2018). To our knowledge there has been no survey-based study of income sources of men and women, in particular employment outside the home and outside farming. However, we expect differing employment profiles to affect the quest for convenience foods as discussed above.

To address these three gaps in literature on processed food consumption in Africa with respect to city sizes, rural distance to cities, and employment categories of consumers, we undertake an analysis of a uniquely detailed household consumption data set for urban and rural Tanzania. The data come from a survey of 9788 households by the Tanzania Household Budget Survey (HBS) in 2011/2012 (the most recent HBS). We use HBS instead of LSMS to get the needed food item detail, as HBS distinguished 199 food items while LSMS followed 59.

The paper proceeds as follows. Section 2 lays out our general conceptual framework. Section 3 specifies definitions and discusses the data. Section 4 presents descriptive results on processed and prepared food consumption. Section 5 provides regression results. Section 6 concludes.

1.2 General Model and Hypotheses

Our analysis uses an Engel curve model, following Lewbel (2006).[1] The shares of household consumption of five categories of food, with varying levels of processing, are functions of the total expenditure of the household (modeled as total consumption, summing purchases and home-consumption of own production) and other household characteristics:

$$w_i = h_i[\log(y), \mathbf{z}] \qquad (1)$$

where w_i is the share of the budget spent on processing category *i* (unprocessed; low processed; unpackaged high processed; packaged high processed; and meals away from home (MAFH)); y is total expenditure; and **z** is a vector of household demographic variables. We hypothesize that as income (proxied by total expenditure) increases, the opportunity cost of time for home processing and home food preparation increases, and the demand for processed food and MAFH increases.

We augment the basic Engel curve analysis with a vector **z** of household demographic and human capital variables, as is common (Banks et al. 1997, Lewbel 2006). These include the age, gender, marital status, and education of the household head, the household's dependency ratio, and its size in adult equivalents. We also include variables that would influence the food environment affecting diet preferences, and assets that save time in home processing and preparation and transport: televisions, mobile phones, various cooking appliances, and motorized vehicles.

[1] Estimation of a full demand system such as QUAIDS or EASI is not possible due to a lack of price data.

Second, to model the household's opportunity cost of time, we include employment variables for women (that would affect their incentive and capacity to spend time in home-processing and preparation) and men (that affect their incentive and capacity to wait for breakfast at home in the early morning and to return home for lunch or dinner depending on their work schedule). Women's employment effects are commonly discussed in the consumption literature but the effects of men's employment patterns are rarely analyzed, especially their effects on consumption of MAFH.

For rural areas, we augment the Engel curve model with a measure (full time equivalents, FTEs) of the time spent in non-farm self-employment and non-farm wage work, separately for men and women. We hypothesize that an increase in women's time spent in the non-farm labor force (be it self-employment or wage) will increase consumption of highly processed food and MAFH. For urban areas we also model employment outside the home in self-employment and wage employment. We expect the magnitude of the effect to be higher for wage work, as that is usually done outside of the home. But self-employment may be done from home, which would permit more flexibility to cook at home. We expect men's time working away from the farm and its nearby household, and in urban work away from the home premises, to be correlated with MAFH.

Finally, we augment the basic Engel curve model with geo-spatial variables, with the hypotheses justified in the introduction. For rural households, we use the distance from the village to the nearest urban area. We expect that as such distance increases, the consumption shares of unprocessed and low-processed food will increase, and the shares of high processed food and MAFH will decrease. For urban households, we use the size of the city (in order of increasing population size, these are towns, secondary cities, and primary cities, defined below). We hypothesize that the size of the city is positively correlated with shares of the more highly processed types of foods (packaged high processed and meals away from home).

1.3 Survey Data and Definitions of Product and Spatial Categories

1.3.1 Survey Data

We use the consumption data from the Tanzania Household Budget Survey (HBS), a detailed cross-sectional, nationally representative survey conducted from October 2011 to October 2012; this is the most recent HBS available. The HBS sample includes 9788 households of which 5628 (57%) are rural, and 4149 (43%) are urban. The distinction between rural and urban is based on HBS's using the government's criteria of population and infrastructure density (National Bureau of Statistics 2014).

We use the HBS data set instead of LSMS (Living Standards Measurement Survey) data for Tanzania because to compose our processed food categories we need the greatest disaggregation possible in the food category. HBS has 199 food items in its consumption diary while LSMS only has 59 food items in its consumption recall.

Moreover, HBS, unlike LSMS, provides a breakdown of food away from home into meals away from home (MAFH) versus products that are bought away from home that are not meals, like a sweet cake at a street vendor. Using the category MAFH draws a sharper line between food service (restaurants, food stalls) and retail; it highlights MAFH as a substitute for home cooking. By contrast, other food away from home could be considered just food bought from a retailer or a food service stall and taken back home or on the road or eaten in situ. Of course, MAFH can also be take-away.

The HBS collected information on, inter alia, household members' employment, assets, and consumption of nonfood and food products. The latter are recorded as to their source: (1) from own production on the household farm; (2) from purchases of food to consume at home; (3) from purchases of meals consumed away from home; (4) from purchases of other food products consumed away from home.

The household-level food consumption data are from a 28-day diary. Respondents were instructed to record all food consumed (in unit and value terms) and the four sources noted above by all members of the household during the day, for 28 consecutive days. For illiterate households, enumerators visited daily to record consumption; for other households, enumerators checked in every few days. For food that was home-produced, the household was asked to estimate the monetary value of the food in Tanzanian shillings. The survey implementation was undertaken so that diaries were being done all year over the sample; thus there is little bias where the consumption in a certain part of the country is correlated with only one season.

1.3.2 Definitions of Variables

1.3.2.1 Processed Food Categories

We consider five levels of food processing characterizing the product in order of increasing convenience and value added. Our five categories are the following (with examples given in Table 1.1):

(1) unprocessed: raw fruits and vegetables, grains, beans, and live animals, that have undergone no physical or chemical processing;

(2) low-processed: a food with a single unprocessed product that has undergone one physical process such as milling or grating (what may also be termed first-stage processing or minimally processed). In the Tanzanian context, the most common are maize flour and milled rice. Unprocessed and low-processed foods require some preparation (peeling, chopping, cooking, etc.) before consumption;

(3) highly processed unpackaged: a food comprising multiple low-processed products, with the set undergoing further processing such as frying or baking. A main example of this category in Tanzania is the mandazi, a fried wheat bun like a donut that is a common snack;

(4) highly processed packaged: we refer to this as an ultra-processed food, and includes items such as biscuits/cookies, bread, soda, and alcoholic beverages;

(5) meals consumed away from home (MAFH): typical examples include rice with beans, chips (French fries) and eggs, and purchases from cafés. Because these products are home meal replacements, they are the category of greatest convenience.

We categorized the item at its state at acquisition, before any further home processing or preparation. Thus, all home-produced products are acquired unprocessed (like grain from the farm). Beans and fruit are usually acquired unprocessed. All acquired-processed products are purchased. Note that FAFH (food away from home) is the aggregate of MAFH and any of 1-4 that is purchased and consumed away from home.

After classifying each food product, we calculated food consumption shares using the consumption diary data. The shares are calculated by dividing the total consumption (whether purchased or home-produced) of each processing category by total food consumption. The latter is the total in value (shilling) terms of own production and food purchases for at-home consumption and FAFH. We exclude transfers and hunted and gathered food (a tiny part of consumption).[2]

[2] Together, transfers and hunted and gathered food account for 1.5% of the observations in the diary of food consumed at home; this rises to 2.5% for rural diary entries, and is just 0.37% of urban diary entries.

1.3.2.2 Spatial Categories

To analyze how processed food consumption varies spatially, we created geographic variables using GPS coordinates obtained from the Tanzanian National Bureau of Statistics for the surveyed enumeration areas (EA). We took the following steps.

First, we used the 2012 census data to obtain the population and administrative classification of all 33 urban areas in Tanzania. The government of Tanzania classes urban areas into: (1) primary cities, with population above 500,000; (2) municipalities which we call "secondary cities", with population between 100,000 and 500,000; and (3) towns, with population between 20,000 and 100,000.

Second, using the HERE API (Application Programming Interface) (www.developer.here.com) we calculated the distance in kilometers (km) and travel time in hours of each of the 395 EAs in the full sample from the centers of all the 33 urban areas, as follows: (a) we calculated the average radius of primary and secondary cities and towns using Google Maps; (b) we classed urban EAs into one of the three urban categories above based on which the EA was in; (c) we classed EAs that straddled the border of urban areas as follows: if it was within 10km of the border of the urban area it was classed as being in that type of city; this accounts for 23% of mixed EAs; the other 77% of mixed EAs was classed as rural.

Third, in the spatial analysis of rural EAs, we further classed them into three sets: (a) peri-urban, within 1 hour of a city or town; (b) intermediate, between 1 and 3 hours, and hinterland, 3 or more hours from an urban area.

1.3.2.3 Employment Variables

A key hypothesized determinant of consumption of processed food is employment. For example, urban women who work outside the home, and rural women engaged in RNFE (rural nonfarm employment), would have higher opportunity costs of time to stay home and home-process and home-prepare food. Thus we calculated from the HBS labor module the FTEs (full time equivalents) of employment by working-age (ages 15-64) men and women in the household working in nonfarm self-employment and wage employment. Aggregating over the course of a year and dividing by 2,016 gives the FTE spent in each type of work.[3] Therefore, an FTE value of 1 corresponds to an average of 40 hours worked per week over the year; an FTE value of 0.5 corresponds to an average of 20 hours per week, and so on.

1.4 Descriptive Results on Processed Food Consumption

Below we present descriptive results for rural compared with urban processed food consumption. It is striking that whereas the conventional image has the urban areas and especially the urban middle class as the focal point of processed food penetration, we find that rural and urban areas are not very far apart in their processed food consumption, with a gradation and continuum over space from intermediate to peri-urban to small towns and secondary cities and then a jump up to primary cities, but a continuum and not an abrupt leap from nearly nothing to a high share in urban areas. Moreover, we find that hinterland areas (where 11% of rural Tanzanians live) far from the cities show a spike in ultra-processed food, showing SSBs and snacks and MAFH penetrating these areas. Finally, we find that in each area, the processed food profile of the poor is somewhat below that of the upper tercile, but again not with an abrupt jump. These results imply that the rise in the opportunity cost of time in rural and urban areas, and among rich and poor alike, has driven a much broader diffusion of processed food than we expected. The details are presented below.

[3] We define "full time" to be 2,016 hours per year (= 40 hours/week * 4.2 weeks/month * 12 months/year).

1.4.1 Rural Results

Table 1.2 shows processed food consumption by rural tercile of total expenditure per adult equivalent (a proxy for income). Terciles are calculated within the overall rural sample. Several points stand out.

First, annual food consumption per AE is 55% of total expenditure in rural areas. That share is 1.5 times higher for tercile 1 (68%) compared with tercile 3 (44%). This pattern is expected from Engels' Law. Total food consumption per AE (adult equivalent) is 276 USD per year, about 70% of the urban level shown below.

Second, on average about 60% of rural food consumption is purchased (in value terms).[4] This is noteworthy when viewed in light of the common image of rural households in Africa depending mostly on own-farming for food, and even more striking when one notes that even the poorest tercile buys 50% of their food, just a little below the upper tercile, which buys 75% of their food. This finding of the importance of purchases in the diet of rural areas, including the rural poor, is similar to that of Sibhatu and Qaim (2017) in Ethiopia, and Tschirley et al. (2015) for a set of Eastern and Southern African countries (Ethiopia, Malawi, Mozambique, Tanzania, and Uganda).

Third, it is not surprising that the share of own-production averages 44% of unprocessed food consumption (see Table 1.7 in Appendix B). The share for the lowest tercile (55%) is nearly double that for the highest tercile (30%). The poorest rely more on their own farming for raw products, and the rural upper tercile, more on purchases. The complement is that the purchased share of raw products is 71% in tercile 3 and only 45% in tercile 1.

[4] The value of home production was estimated by the household. The survey questionnaire did not specify to the respondent whether to use the consumer price or producer price in this valuation. We interpret the value of home production as the average opportunity cost of buying the product instead of the producer price of selling the product.

Fourth, it is striking that even in rural areas the share of purchased-processed food in all rural food consumption is high – 47% overall, varying from 39% to 47% to 59% over the three terciles (Table 1.2). This is below the 78% share of processed food in urban areas shown below, but the rural share is still high compared to what we perceive as conventional wisdom regarding African rural areas. We observe that in food debates in Tanzania and Africa in general, purchased-processed food is considered an urban phenomenon, and middle class at that. Our findings show that it is neither only urban nor only middle class.

Fifth, purchases of low-processed food form 77% of purchases of all processed food, and form 56% of all rural food consumption, with strikingly little variation over terciles (Table 1.2). The purchased low-processed food is 58% from milled grains – 34% is maize flour, 19% is milled rice, 5% is milled wheat (Table 1.7).

However, own-produced low-processed food is 37% of low-processed food consumption in rural areas, nearly doubling from high (24%) to low income tercile (47%) (Table 1.7). This implies a strong reliance on rural households to grow or purchase a raw product and then process it in a custom-milling service. The HBS consumption data have no information on how the raw product was acquired that was consumed low-processed (such as whether the household bought or grew maize grain that they then had ground to use in the main dish, ugali). The data do not show whether the household hand-processed the raw ingredient or took it to be custom-milled at the village mill. Key informants told us that it is rare now for hand-pounding of grain to occur in rural areas, so we assume most of the "home" low-processing is use of custom milling services, itself a time-saving convenience.

Figure 1.1 shows the shares of low-processed wheat, maize, and rice in total processed and purchased low-processed food by the rural households' distance from the nearest city. The distances are grouped into zones: peri-urban (less than 1 hour from urban areas); intermediate (from 1 to 3 hours from urban); and hinterland (3 and more hours from urban). Based on our representative sample, 30% of the rural population lives in the peri-urban, 59% in the intermediate, and 11% in the hinterland zone. The most striking result is that maize flour shares plummet 2-fold from the intermediate to the hinterland zone, from about 30% to 15%. The lower right quadrant explains what took its place: tubers. This fits with the image of the hinterland zones being mainly in the hilly and mountain areas where tuber production and consumption are much more important than in the other zones. Interestingly, the share of rice in purchased low-processed food is relatively steady at a fifth of low processed food, even into the hinterland zone where it is not being displaced by tubers.

Sixth, 18% of processed food in rural areas is high-processed, including 2% high processed unpackaged, 9% ultra-processed (high processed packaged eaten at home or at a kiosk), and 7% MAFH (Table 1.2). Thus, while processed food in general has become important in rural areas, the penetration of sugary/oily snacks and sweets in the rural areas sharply lags the urban areas (discussed below), at present.

The food service sector is less developed in rural areas than in urban areas, as is one of its correlates, commuting to work. Yet MAFH form 8% of purchased food in rural areas (Table 1.2). This depicts a continuity with the patterns of the towns' use of the food service sector, rather than an abrupt and large difference in patterns in rural areas. Upper tercile rural households have almost the same MAFH share in purchases (12%) as do towns and secondary cities (14% and 13% respectively, as we show below). As with urban areas, the rural poor rely less on MAFH, with their MAFH share in purchases being only 6%.

Table 1.3 and Figure 1.2 show the distribution over rural zones of consumption of processed food. Table 1.3 shows the composition of ultra-processed food consumption in rural areas by zone. SSBs occupy first rank in ultra-processed food, averaging a quarter of ultra-processed, with a U-shaped curve over rural space. The shares of bread and dairy products fall quickly from peri-urban to hinterland zones; perhaps this is related to greater commuting in peri-urban areas where workers rise early to go to work and have bread and coffee with milk en route. Another striking finding is the importance of alcoholic beverages, the share of which rises steeply from peri-urban (13%) to hinterland zones (around 31%). This is linked to a steep rise of the share of hard liquor as the distance from the urban areas increases. Interestingly, the share of chocolate and confectionary products is low in all the rural zones, perhaps because of the popularity of the traditional *mandazi* (local sweet donuts).

Table 1.9 in Appendix B focuses on the spatial distribution and levels of SSBs, showing unconditional and conditional averages of expenditure per person on SSBs in shillings and milliliters per month of SSBs. We discuss the table in terms of the equivalents of cans of Coca Cola. A can of Coca Cola is 330ml. The table shows that SSB consumption in urban and rural areas is similar: 2.25 cans of Coca Cola per month per urban adult equivalent; a bit less than 2 cans in rural areas. However, within rural areas there is a lot of heterogeneity, with a steady decrease from the peri-urban through the intermediate zone, and then a large increase in the 3-4 hour part of the hinterland, and then a sharp fall beyond 4 hours from cities.

Figure 1.2 presents rural consumption shares of processed food categories over hours of distance from urban areas, grouped into zones as discussed above. The upper left quadrant shows that the share of unprocessed food stays interestingly steady at around 30% across the heterogeneous rural zones, only changing abruptly starting 4 hours from urban areas, in the hinterland zone. The complement of this is that the share of processed food stays steady at around 70%, again with the drop in processed food only in the hinterland zone, where only 11% of the rural population live. As we explained above, this is because grain as a staple gives way to roots and tubers as the main staples far from the cities, essentially in the mountain areas in the west and center south of the country.

The upper right quadrant of Figure 1.2 shows that low-processed food stays remarkably steady as a share of processed food over rural zones, hovering around the mid 80 percentiles after a small dip up from the peri-urban area closest to the city. This shows that the great majority of processed food consumption in rural areas is still the “first stage” of processing evolution, where women buy flour and rice and oil and cut up beef and so on and home-prepare meals. This is a step “beyond” hand pounding of cereals their mothers did 20-30 years ago, but a step behind their urban counterparts who as Table 1.4 shows, buy around 62% of their processed food as low processed and have thus moved to nearly 40% of their processed food as ready to eat or heat and eat, eschewing both home processing and home preparation.

The lower left quadrant shows that throughout the rural zones unpackaged high-processed food is only 1-2% of consumption and processed food. Thus, while a snack like traditional mandazi is often held up as the archetype of what rural households might buy as a snack, it is far less important than the “non-traditional” ultra-processed foods we show in the lower right quadrant – even in rural areas.

Packaged ultra-processed food shows a slight U curve over space, from 11% of processed food in peri-urban areas to a low of an average 8% in other rural areas, but interestingly jumping back to 10% in hinterland areas. That shows the reach of marketing of packaged ultra-processed foods, belying the popular image of them as urban foods. Figure 1.2 surprisingly shows MAFH traces a similar U curve, with around 10% in the peri-urban and hinterland areas, dipping to 5% in the middle of the U. The penetration of MAFH in even hinterland areas is fascinating.

1.4.2 Urban Results

Table 1.4 shows processed food in urban food consumption, by city type and by total expenditure tercile. Terciles are calculated within each urban category (towns, secondary cities, primary cities). Several points stand out.

First, annual total food consumption per AE is roughly 40% of total expenditure per urban category. That share is thrice higher for tercile 1 (the poorest) compared with tercile 3 (the richest) in towns, and twice higher in secondary and primary cities, as expected from Engels' Law.

Second, on average 98% of urban food consumption is purchased, with little variation over urban category and tercile. Only 2% of food comes from own-farming, whether from farms operated by urban consumers in rural areas or "urban agriculture."

Third, processed food purchases are 78% of food consumption in Tanzanian urban areas. Interestingly that share barely differs over towns and secondary cities (at 75%) and primary cities (at 80%). The corollary is that the share of food consumption acquired unprocessed (22%) is much lower than one would expect if we were in the traditional situation where households buy raw products and home-process and home-prepare them.

Fourth, purchased low-processed products form 62% of purchased-processed food in urban areas (varying from 68% in towns and secondary cities to only 54% in primary cities). Online Table 1.8 shows that 58% of low-processed food in urban areas is milled cereal – 21% maize flour, rice, 24%, and wheat flour, 13%. The share of maize flour drops from towns to primary cities, from 24% to 17%. The share in the lowest tercile is about twice that of the highest tercile. Those general patterns hold also for rice and wheat, but less sharply. This shows the importance of maize, but instead of buying it in grain form and then hand pounding or even custom milling, it is now common even in small towns and among the poor to buy it already milled.

Fifth, in urban areas, high-processed packaged (ultra-processed) foods, such as potato chips and sweetened beverages, are more than 4 times as important as high processed unpackaged foods, such as traditional fritters and cakes (such as mandazi) bought mainly to eat at home. In particular, the average budget share of ultra-processed is 9%, and the average share of high processed unpackaged products is just 2%.

Table 1.3 examines the composition of ultra-processed food (as packaged and highly processed) in urban areas. Two product sets dominate with two-thirds of the category, each with about a third with similar distribution over urban categories. First-ranked are sugar-sweetened beverages (SSBs) that average about a third of ultra-processed food consumption. The second-ranked is bread, also about a third of ultra-processed; it is not specified as packaged in the data set but key informants noted that it is nearly always sold that way. The last third of the category is highly fragmented over a wide range of products with modest differences over urban categories.

Table 1.8 shows that ultra-processed (packaged) food bought away from home (not as MAFH but just from kiosks to consume at the kiosks or take back home, such as drinking a sugary soda at a kiosk) forms 35% of ultra-processed food consumption, rising from 31% in towns/secondary cities to 42% in primary cities. This attests to a shift in street kiosks selling traditional snacks to selling packaged snacks today (as found also in Namibia by Nickanor et al. 2019). Note that there is also a strong income correlation in consumption of FAFH packaged snacks and beverages, with the share nearly doubling between the lower and upper terciles in all urban categories (in the towns, 23% to 39%; in the secondary cities, 30% to 45%; in the primary cities, 28% to 55%).

Sixth, MAFH is a surprisingly developed phenomenon. In urban Tanzania, it is twice as important as high and ultra-processed packaged foods eaten at home or in kiosks (the MAFH share is 18% and the ultra-processed share is 9%) (Table 1.4). MAFH are fully 13% of all food consumption in towns and secondary cities and a stunning 26% in primary cities, with 18% of urban food consumption overall. The upper tercile is particularly reliant on MAFH with 21%, 25%, and 39% from MAFH over the three city types, thrice the shares for the lowest tercile. The emergence of the "*mama ntilie*" (street food vendor) is a lynchpin of the urban food system, where commuting and women's work outside the home have spurred the rise of a large food service sector.

In sum, high and ultra-processed and MAFH together, as a share of processed food, form 30% in towns/secondary cities and 49% in primary cities. The counterpart is that low-processed is 70% of processed food in towns/secondary cities and only 51% in primary cities. This makes sense as usually the first step in purchasing processed food is to acquire time-saving forms of basic ingredients like flour that form the base of home cooking. The second step is to largely replace home cooking by buying ready to eat products and meals, and that step is taken "earlier" by richer consumers and the bigger cities.

There is an interesting contrast with findings of other studies, such as for Ouagadougou, Burkina Faso, in Reardon et al. (1989), where MAFH were much more important for the lowest tercile compared with the highest. In that case, poor workers commuted and could not easily return home to lunch, while the schedules and means of transport of the upper tercile consumers allowed lunch at home. In Tanzania, it is the opposite. The richer consumers, working in office buildings, lunch near them, perhaps because the development of the food service sector has made it much easier to do that now than several decades ago.

1.5 Regression Models and Results

1.5.1 Regression Models and Estimation Methods

We estimate the following augmented Engel curve models, following Lewbel (2006). To allow for differences in the parameters between urban and rural households, we estimate the model separately for the urban and rural samples. In particular, there may be fundamental differences between urban and rural households that are not adequately captured by distance or city size dummies, such as the food environment, employment opportunities, agro-climatic conditions, and lifestyle patterns. For rural households, the regression is:

$$Share_i = \alpha_0 + \alpha_1 ln(total\_exp) + \alpha_2 (ln(total\_exp))^2 + \alpha_3 dist + \alpha_4 dist^2 + \boldsymbol{labor\delta} + \boldsymbol{other\_controls\zeta} + \varepsilon_i \quad (2)$$

- $Share_i$ is the share of the household's food consumption (from purchases plus own production) on processing category *i*. As discussed above, the processing categories are unprocessed, low-processed, unpackaged high-processed, ultra-processed, and MAFH.
- *ln(total_exp)* is the natural log of the household's total monthly consumption of food and non-food goods and services per adult equivalent. Total expenditure enters the model quadratically (Banks et al. 1997, Deaton and Zaidi 2002). We hypothesize that the higher the income, the greater the share of consumption to higher processed foods because the opportunity cost of time rises with income.
- *dist* is the distance (in kilometers) from the household's village to the nearest urban area (either a town, secondary city, or primary city). Distance enters the model quadratically. We hypothesize that the greater the distance from urban areas, the lower the share of consumption to high processed foods because the food environment is less amenable to processed foods in deep rural areas and consumers tend to work nearer home.

- *labor* is a vector of four employment-related variables (FTEs in non-farm self-employment and non-farm wage work, each for men and women). We hypothesize that processed food consumption is correlated with non-farm activity, especially wage work that takes consumers further from home and thus increases the opportunity cost of returning to process, prepare, and eat food at home.
- The vector *other_controls* contains household demographics such as the age and education of the household head, a dummy variable equal to one if the household head is female, and a dummy variable equal to one if the household head is married. Also included are other dummy variables that could affect preferences including ownership of bicycle and motor vehicle (motorcycle or car), cooking appliance, and communication assets (TV, radio, mobile phone). We hypothesize that married couples eat more at home and consume less processed food. Households with easy transport access and cooking appliances could return home and cook at home easier and so may be less apt to consume processed food. Those with TVs and radios may be more exposed to advertising and consume more processed food.
- We include region dummies and dummies for the month of the interview, to control for spatial and temporal heterogeneity.

The urban model is similar, except that instead of *dist*, we include dummy variables that indicate whether the household lives in a secondary or primary city (with town as the base):

$$Share_i = \beta_0 + \beta_1 ln(total\_exp) + \beta_2 ln(total\_exp)^2 + \beta_3 secondary + \beta_4 primary + \boldsymbol{labor\eta} + \boldsymbol{other\_controls\theta} + u_i \quad (3)$$

The *labor* and *other_controls* vectors contain the same variables as in the rural model.

Both sets of regressions were estimated using ordinary least squares (OLS). We ran two robustness checks: we also estimated the equations using seemingly unrelated regressions (SUR), to account for correlation among the error terms across equations, and fractional multinomial logit (FMNL), to account for the fact that the dependent variables are fractional and together sum to one (Zellner 1962, Papke and Wooldridge 1996). When each equation has the same set of explanatory variables (as is the case in our analysis), the SUR coefficient estimates are identical to the OLS estimates. The FMNL estimator ensures that the values of the predicted dependent variables lie between 0 and 1, and no adjustments are needed for observations at 0 and 1 (Papke and Wooldridge 1996). There are marginal changes in statistical significance in a handful of SUR, FMNL, and OLS estimates, but because the key takeaways (i.e., the coefficient estimates on the total expenditure, distance, and city size variables) remain unchanged, we report and discuss the OLS results.

Tables 1.10-1.12 in the online supplement show descriptive statistics for male and female employment and other regression determinants for rural and urban households.

1.5.2 Econometric Results

The results of the Engel curve regressions are presented in Table 1.5 (for rural households) and Table 1.6 (for urban households).[5] Several results are salient.

1.5.2.1 Total Expenditure

For urban households, total expenditure (TE) is negatively correlated with the consumption shares of unprocessed, low processed, and unpackaged high processed food. These results are significant at the 1% level. TE is positively and significantly correlated with the consumption shares of ultra-processed food and MAFH. This controls for the urban household's employment profile, and thus suggests a correlation between income and opportunity cost of time and the

[5] The SUR and FMNL results are present in tables A7-A10 in the online supplemental appendix.

pursuit of convenience in food consumption. The results are roughly the same in the rural sample, with the exceptions that TE is not significantly correlated with the consumption share of unprocessed food, and is positively correlated with the share of unpackaged high processed food (though the magnitude is very small).

1.5.2.2 Employment, Representing Opportunity Cost of Time

Tables 1.5 and 1.6 show the correlations between a household's engagement in various types of employment and processed food category shares.[6] For nonfarm self-employment (NFSE), a marginal increase in female FTEs has few correlations with the processing categories. In the urban sample, female NFSE is positively correlated with the shares of unprocessed food and unpackaged high processed food, and is negatively correlated with the share of MAFH. This suggests that additional time spent by females on self-employment does not increase demand for more processed food, and may even decrease the share of the budget dedicated to MAFH. This makes sense because self-employment can often be from the home itself or be flexible to allow the woman to return home to cook. In the rural sample, female NFSE is negatively correlated with the consumption share of unprocessed food and positively correlated with the share of unpackaged, high processed food (such as *mandazi*).

Female non-farm wage labor (NFW) in urban areas is strongly and negatively correlated with the consumption shares of unprocessed, low processed, and unpackaged high processed food, and strongly and positively correlated with the consumption share of MAFH. This makes sense because wage work tends to be away from home and inflexible in hours, so both the woman and the rest of the family may need to eat outside the home, at least at lunchtime. However, in rural

[6] We also estimated the regressions using the share of time spent in each employment category, by gender (see tables A11 and A12 in the online supplemental appendix). The results are very similar to those using FTE levels, which are discussed in the text.

areas, female NFW work is largely insignificant in its impact on processed food consumption, likely because the sample of households with females engaged in such work is too small.

The above non-farm employment results support our hypotheses that additional female participation in the non-farm labor force is positively correlated with consumption of high processed food and MAFH, and that the association is stronger with wage work than with self-employment work.

The participation of males in the non-farm labor force is especially strongly correlated with the share of MAFH, as hypothesized. For both urban and rural households, male NFSE FTEs have a positive and significant (at the 5% level) correlation with the share of MAFH, and the magnitude of the correlation is approximately the same in urban and rural areas. Additional FTEs of male non-farm wage (NFW) work exhibit even larger positive correlations with the share of MAFH, and negative correlations with the shares of unprocessed and low-processed food. These results are consistent with our hypothesis of a link between men working away from home in enterprises and their need to commute and thus eat out.

1.5.2.3 Distance and City Size

In the rural regressions, the average marginal effect of the distance to the nearest urban area is negative in the unprocessed and unpackaged high processed regressions, and positive in the low processed regression. This coincides with the descriptive results: the more distant a rural household is from urban areas, i.e., the more in the "hinterland," the more of their food budget is spent on the simple act of replacing laborious hand pounding of staple grains like maize by buying already milled grain, less on unprocessed food (such as fruits and vegetables), and less on buying traditional snack foods (such as sweet buns).

Importantly, distance from urban areas does not drive down ultra-processed food consumption – in fact there is no significant correlation. Recall from the descriptive section that even in the deep hinterland there was substantial consumption, in fact similar to even peri-urban areas, of SSBs, bottled beers and spirits, and packaged sweets.

For urban households, our regressions do not show a significant impact of living in a secondary city compared to living in a town. However, living in a primary city (compared with a town) is strongly, negatively correlated with the share of low-processed food, and strongly, positively correlated with the shares of ultra-processed food and MAFH. These results support our hypotheses and coincide with the descriptive results.

1.5.2.4 Demographic Variables

In both urban and rural regressions, an increase in the dependency ratio is associated with higher average unprocessed and low processed shares, and a lower share of MAFH. This makes sense: with more children and aged to perform home chores for, women need to stay at home more, but they also try to avoid hand-pounding of grain and so buy flour and other first-stage processed foods. Likewise as they tend to cook more, the women buy less MAFH. The same effect is seen with married household heads, who, we surmise, would have a greater tendency to eat at home at least for dinner, compared to a household head who is single.

Finally, as expected, having kitchen assets such as stoves and refrigerators is negatively associated with the consumption share of MAFH in both sets of regressions.

1.6 Conclusions

In this paper we sought to contribute to the literature on the penetration of processed food in diets of developing countries with an exploration of the consumption of a continuum of processed food categories – from minimally processed to ultra-processed, and from processed food purchased and consumed at home to meals away from home (MAFH). We crossed that product perspective with a spatial continuum, from primary to secondary cities to towns, and from peri-urban to intermediate to hinterland rural areas. The processed and prepared food categories continuum has across it the commonality that all those categories are "convenience foods," as they save women's labor time in home processing and preparation, and men and women's time returning home to eat.

Moreover, the spatial categories continuum has across it the commonality that in all those spaces, men and women depend a lot on working outside the home, driving the quest for convenience; in each space there are also strata with higher incomes who have higher opportunity costs of time, and the difference between urban and rural blurs among these strata. For example, the richest households in rural areas have a similar total food budget to that of urban areas. Urban men and women are heavily engaged in work outside the home. In primary cities women's jobs tend to be wage jobs, taking them away from home; in secondary cities, many women are in self-employment, as they are in towns; in towns, many women also work on farms near the town. Rural people also have major job time commitments, not just on their farms: a quarter of rural women and a third of men work off-farm.

Using these product and spatial lenses, with a unique detailed data set from Tanzania, we had several key findings.

First, we found a deep penetration of processed food in general. Instead of processed food penetration being mainly confined to the urban middle class, we found (echoing Tschirley et al. 2015a) that it has spread well into the consumption of the rural areas where 47% of food consumption overall is purchased-processed food; even in the rural poor tercile, it is 39%. In urban areas, purchased-processed food is on average 78% of food consumption.

Second, we found that even ultra-processed food has penetrated diets beyond the urban middle class. For example, in towns, ultra-processed is 15% of processed food consumed by the upper tercile, but even 9% of the low tercile. Much of these are SSBs and packaged bread and bottled alcoholic drinks, but also a wide range of other snack foods.

Third, MAFH have penetrated surprisingly deeply into diets of urban Tanzanians in particular, where it averages 18% of food consumption and 23% of total processed food. MAFH has also penetrated rural diets, averaging 5% of food consumption and 7% of processed food consumption. In both cases MAFH consumption is correlated with income.

Fourth, processed food and prepared food (MAFH) penetration is most intense in primary cities, but has also gone far in secondary cities and is now making headway in towns. But it does not stop at city limits. It occurs over rural zones, in interesting patterns. Ultra-processed food is important in peri-urban areas, is less marked in intermediate zones, and then is again important in hinterland areas far from cities. This shows the far reach of marketing of branded SSBs in particular. MAFH are most important in urban, then second in peri-urban, then less in intermediate rural areas, and least in hinterland areas, but still present.

Fifth, as hypothesized, income, women's and men's employment outside the home, size of city, and rural distance from cities condition the purchase of processed foods and MAFH, linked with opportunity cost of time to home-process and prepare foods implied by those factors.

Our findings demonstrate that processed food supply chains provide convenience foods, both low processed and ultra-processed, where rising incomes, changing employment patterns of men and women, and tastes create demand for them. With this transformation, rooted in basic lifestyle changes similar to those experienced over the decades in developed countries, has come a rise in the consumption of unhealthy highly processed foods, even among the poor and even in hinterland areas distant from the city.

The work points to the need for further research on the nutrition effects of these trends in rural areas and the structure and conduct of supply chains providing these foods. It also points to the need for further research on the employment effects of these changes (see Tschirley et al. 2015b for one treatment of this issue) such as reducing women's home chore time to engage in off-farm employment, and men in the development of commuting patterns.

APPENDICES

APPENDIX A: Tables and Figures

Table 1.1 Example products in each processing category

Category	Example Products
Unprocessed	Raw fruits and vegetables, roots and tubers, grains, beans and lentils, fresh cow milk
Low Processed	Maize flour, rice, cooking oils, butchered meat, sugar, drinking water
High Processed, Unpackaged	Fried wheat buns (*mandazi*), chickpea patties (*bagia*)
Ultra-Processed	Bread, sugar-sweetened beverages, alcohol, pasta, biscuits, cake, coffee and tea

Table 1.2 Average shares (percentages) of processed categories in total food consumption in value terms: rural households

	T1*	T2	T3	All
N =	1,876	1,876	1,876	5,628
Average yearly food consumption per AE (USD)	$156	$240	$420	$276
Yearly total expenditure per AE (USD)	$228	$426	$953	$499
Total Food Budget Shares				
Own-produced share[a]	47.4	36.6	24.1	37.1
Total Purchased Share	48.8	60.5	73.7	59.8
Purchased to eat at home	44.2	53.7	59.8	51.9
FAFH	4.6	6.8	13.9	7.9
Share of purchased processed food (including processed FAFH and MAFH) in total food expenditure	38.6	47.4	58.6	47.3
Unprocessed Share	30.2	29.5	26.0	28.8
Low Processed Share	58.6	56.9	52.9	56.4
High Processed Unpackaged Share	0.6	1.0	1.3	1.0
Ultra-Processed Share	4.4	5.8	7.7	5.8
MAFH Share	2.5	3.9	9.9	5.1
Share of …. In total purchased food				
Purchased unprocessed	24.9	24.0	22.6	24.0
Purchased low processed	60.1	58.8	53.6	57.8
Purchased high processed unpackaged	1.1	1.4	1.6	1.4
Purchased ultra-processed	8.5	9.1	10.1	9.1
MAFH	5.5	6.6	12.1	7.7
Share of …. In total purchased processed food				
Purchased low processed	80.5	78.3	70.9	77.1
Purchased high processed unpackaged	1.5	1.9	2.1	1.8
Purchased ultra-processed	11.1	11.8	12.9	11.8
MAFH	6.8	8.0	14.1	9.2

* Refers to terciles calculated using total expenditure per adult equivalent. [a] These are the shares of each source within the total household food budget (will not exactly sum to 100 down columns due to exclusion of gifts/food aid share). Source: authors' calculations.

Table 1.3 Composition of ultra-processed category by zone (percent of each sub-category in total ultra-processed)

	Urban				Rural						
					Peri-Urban	Intermediate		Hinterland			Overall Rural
	Primary City	Secondary City	Town	Overall Urban	<1 hr	[1hr, 2hr]	[2hr, 3hr]	[3hr, 4hr]	[4hr, 5hr]	[5hr, 6hr]	
Share of **sugar-sweetened beverages (SSBs), including fruit juices**, in ultra-processed	32	31	33	32	28	24	21	22	33	29	25
Share of **bread**	38	33	29	37	21	12	9	9	9	8	14
Share of **tea** and **coffee**	15	17	13	15	13	13	12	11	15	19	13
Share of **biscuits, cake, and ice cream**	3	3	3	3	3	3	3	2	2	2	3
Share of **alcohol (beer, spirits/liquors, and wine)**	3	4	10	3	13	24	24	28	36	29	21
Share of **pasta**	4	4	3	4	4	2	1	1	0	4	4
Share of **chocolate, confectionery products**	1	1	1	1	1	1	1	1	2	1	1
Share of **processed meats** (beef sausage, dried or salted meat, canned fish/shellfish)	0	1	0	0	1	1	3	0	0	0	0
Share of **dairy products** (powdered milk, jibini, clotted milk, butter, margarine)	2	4	5	2	10	15	23	4	0	3	2
Share of **canned fruits and fruit jams**	0	0	0	0	0	0	0	0	0	0	0
Share of **syrup, condiments, chili sauce, vinegar**	0	0	0	0	0	0	0	0	0	0	0
Share of **baking products** (yeast, cocoa/cooking chocolate, cocoa and powdered chocolate)	0	0	0	0	0	0	0	0	1	0	0
Share of **other products** (cornflakes, unidentifiable products)	2	2	3	2	5	5	3	20	1	5	2
TOTAL	100	100	100	100	100	100	100	100	100	100	100

Source: authors' calculations. Includes food consumed at home and away from home (but not MAFH).

Table 1.4 Average shares (percentages) of processed categories in total food consumption in value terms: urban households

	Towns				Secondary Cities				Primary Cities				All Urban
	T1*	T2	T3	All	T1	T2	T3	All	T1	T2	T3	All	
N =	107	107	106	320	338	338	338	1,014	938	939	938	2,815	4,149
Average yearly food consumption per AE (USD)	$216	$360	$636	$408	$192	$312	$540	$348	$264	$444	$708	$480	$408
Average yearly total expenditure per AE (USD)	$364	$726	$2019	$1026	$350	$655	$1473	$790	$534	$998	$2217	$1237	$1032
Total Food Budget Shares													
Own-produced share[a]	5.6	4.3	1.3	3.8	2.5	1.6	0.9	1.7	0.9	0.2	0.4	0.5	2.0
Total Purchased Share	92.3	95.2	98.1	95.2	96.4	98.1	98.8	97.7	98.8	99.7	99.5	99.4	97.5
Purchased to eat at home	82.5	80.5	70.2	77.8	88.4	86.0	66.9	81.1	79.8	66.1	48.9	65.2	74.5
FAFH	9.8	14.7	27.9	17.4	8.0	12.1	31.9	16.6	19.0	33.6	50.6	34.2	23.0
Share of purchased processed food (including processed FAFH and MAFH) in total food expenditure	70.4	73.3	78.1	74.0	72.7	75.4	80.3	75.9	75.7	80.5	84.6	80.2	78.3
Unprocessed Share	25.4	23.1	20.9	23.1	26.0	23.7	19.2	23.2	23.6	19.4	15.1	19.4	21.8
Low Processed Share	54.4	53.3	44.2	50.7	59.9	57.8	45.7	54.9	48.4	41.8	33.5	41.3	48.8
High Processed Unpackaged Share	3.3	2.9	1.4	2.6	2.0	1.8	1.3	1.7	3.3	2.3	1.3	2.3	2.2
Ultra-Processed Share	6.9	9.2	11.6	9.2	5.2	6.8	8.8	6.8	10.0	11.2	11.2	10.8	9.0
FAFH -- Meals share	8.0	11.0	21.2	13.4	5.8	9.5	24.8	12.8	14.4	25.2	38.9	26.0	17.6
Share of …. In total purchased food													
Purchased unprocessed	25.3	23.5	20.8	23.2	25.3	23.5	19.0	22.8	23.7	19.4	15.1	19.4	20.8
Purchased low processed	55.0	52.5	44.5	50.7	61.3	58.3	45.9	55.7	48.5	41.9	33.5	41.4	46.6
Purchased high processed unpackaged	3.6	3.0	1.5	2.7	2.0	1.8	1.3	1.7	3.3	2.3	1.3	2.3	2.2
Purchased ultra-processed	7.3	9.5	11.8	9.5	5.4	6.8	8.8	6.9	10.0	11.2	11.2	10.8	9.5
MAFH	8.8	11.5	21.4	13.8	5.9	9.6	24.9	12.9	14.5	25.3	39.0	26.1	20.9
Share of …. In total purchased processed food													
Purchased low processed	75.0	69.8	58.6	67.8	82.8	77.5	59.9	74.1	65.4	54.3	42.3	54.2	61.6
Purchased high processed unpackaged	4.8	4.0	2.1	3.6	2.7	2.4	1.8	2.3	4.3	3.0	1.6	3.0	2.9
Purchased ultra-processed	9.6	12.4	15.0	12.4	7.2	8.9	11.0	8.9	13.2	14.1	13.5	13.6	12.0
MAFH	10.6	13.8	24.3	16.2	7.3	11.2	27.3	14.6	17.1	28.7	42.6	29.3	23.5

Table 1.4, cont.

[*] Refers to terciles calculated using total expenditure per adult equivalent. [a] These are the shares of each source within the total household food budget (will not exactly sum to 100 down columns due to exclusion of gifts/food aid share). Source: authors' calculations.

Table 1.5 Econometric results for Engel curve regressions of food budget shares: rural households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	0.369***	0.094	0.035***	-0.019	-0.478***
	(0.089)	(0.082)	(0.010)	(0.031)	(0.099)
Natural log of total household expenditure, squared	-0.017***	-0.007*	-0.002***	0.002	0.024***
	(0.004)	(0.004)	(0.000)	(0.001)	(0.005)
Average marginal effect of ln(total expenditure)[a]	0.001	-0.057***	0.001**	0.019***	0.035***
Household Demographics					
Adult equivalents	0.003**	-0.003**	-0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Dependency ratio	0.008**	0.013***	0.000	-0.006***	-0.015***
	(0.003)	(0.003)	(0.000)	(0.001)	(0.002)
=1 if the head of household is female	0.037***	0.029***	-0.000	-0.018***	-0.048***
	(0.007)	(0.008)	(0.001)	(0.004)	(0.006)
Age of the household head	0.000**	0.000*	0.000	0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	-0.002**	0.001	0.000*	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.024***	0.052***	0.001	-0.016***	-0.060***
	(0.007)	(0.008)	(0.001)	(0.004)	(0.007)
Labor Variables					
Total FTEs of females in nonfarm wage	-0.015*	-0.003	-0.000	0.005	0.013
	(0.008)	(0.012)	(0.001)	(0.004)	(0.010)
Total FTEs of females in nonfarm self-employment	-0.020***	0.010	0.005***	0.004	0.001
	(0.006)	(0.007)	(0.001)	(0.003)	(0.006)
Total FTEs of males in nonfarm wage	-0.019***	-0.017**	-0.000	0.008***	0.028***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.007)
Total FTEs of males in nonfarm self-employment	-0.023***	0.005	0.001	0.002	0.014**
	(0.004)	(0.010)	(0.001)	(0.003)	(0.006)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	0.002	0.040**	-0.001	-0.006	-0.035*
	(0.016)	(0.018)	(0.003)	(0.007)	(0.019)

Table 1.5, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
=1 if HH has charcoal stove	-0.021***	0.021***	0.005***	-0.002	-0.002
	(0.005)	(0.006)	(0.001)	(0.003)	(0.004)
=1 if HH has firewood and coal stove	-0.005	0.021***	-0.002**	-0.006**	-0.009***
	(0.005)	(0.006)	(0.001)	(0.003)	(0.003)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.026**	0.016	-0.008***	-0.002	-0.031***
	(0.011)	(0.014)	(0.002)	(0.006)	(0.011)
=1 if household has either a car or motorcycle	-0.022**	0.053***	0.007***	-0.012***	-0.026***
	(0.010)	(0.011)	(0.002)	(0.004)	(0.010)
=1 if household has a bicycle	-0.001	0.004	0.000	-0.000	-0.002
	(0.005)	(0.005)	(0.001)	(0.002)	(0.004)
Tropical Livestock Units	0.001**	-0.001*	-0.000**	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hectares of land cultivated by HH	0.001*	-0.001*	-0.000*	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
=1 if HH owns a television	-0.016**	0.009	0.002	0.003	0.002
	(0.008)	(0.010)	(0.002)	(0.004)	(0.010)
=1 if HH owns a radio	-0.071*	0.017	0.009*	-0.017	0.062
	(0.039)	(0.052)	(0.005)	(0.017)	(0.045)
=1 if HH has mobile phone	-0.010*	0.012**	0.001*	-0.004*	0.001
	(0.005)	(0.006)	(0.001)	(0.002)	(0.004)
Spatial Variables					
Distance (in 100's of km) to nearest urban center	-0.116***	0.125***	-0.006***	0.001	-0.004
	(0.014)	(0.015)	(0.002)	(0.006)	(0.009)
Distance, squared	0.038***	-0.039***	0.002***	-0.002	0.001
	(0.006)	(0.006)	(0.001)	(0.002)	(0.003)
Average marginal effect of distance[a]	-0.046***	0.054***	-0.003***	-0.003	-0.002
Constant	-1.676***	0.170	-0.187***	0.177	2.516***
	(0.493)	(0.456)	(0.055)	(0.173)	(0.534)
Observations	5,054	5,054	5,054	5,054	5,054
R-squared	0.336	0.318	0.249	0.201	0.223

Notes: Estimated with OLS. Robust standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.6 Econometric results for Engel curve regressions of food budget shares: urban households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	-0.115*	-0.041	-0.007	0.165***	-0.002
	(0.060)	(0.080)	(0.011)	(0.023)	(0.125)
Natural log of total household expenditure, squared	0.004	-0.001	0.000	-0.006***	0.004
	(0.002)	(0.003)	(0.000)	(0.001)	(0.005)
Average marginal effect of ln(total expenditure)[a]	-0.032***	-0.058***	-0.006***	0.015***	0.080***
Household Demographics					
Adult equivalents	-0.001	0.008***	0.000	0.002***	-0.009***
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)
Dependency ratio	0.018***	0.044***	0.004***	-0.002	-0.064***
	(0.003)	(0.004)	(0.001)	(0.002)	(0.005)
=1 if the head of household is female	0.050***	0.068***	0.003	0.001	-0.123***
	(0.006)	(0.009)	(0.002)	(0.003)	(0.011)
Age of the household head	0.001***	0.001***	0.000***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	-0.000	0.000	-0.000*	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.047***	0.084***	0.004***	-0.001	-0.132***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.010)
Labor Variables					
Total FTEs of females in nonfarm wage	-0.011***	-0.013***	-0.002***	0.001	0.025***
	(0.003)	(0.005)	(0.001)	(0.002)	(0.006)
Total FTEs of females in nonfarm self-employment	0.008**	0.004	0.002*	-0.001	-0.012**
	(0.004)	(0.005)	(0.001)	(0.002)	(0.006)
Total FTEs of males in nonfarm wage	-0.009***	-0.013***	-0.001**	0.000	0.023***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)
Total FTEs of males in nonfarm self-employment	-0.007**	-0.007	-0.001	0.001	0.014**
	(0.003)	(0.005)	(0.001)	(0.002)	(0.006)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	0.014**	0.016*	-0.003**	0.008*	-0.035***
	(0.007)	(0.009)	(0.002)	(0.004)	(0.012)
=1 if HH has charcoal stove	0.064***	0.091***	0.005***	-0.000	-0.160***

Table 1.6, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
	(0.006)	(0.009)	(0.002)	(0.004)	(0.013)
=1 if HH has firewood and coal stove	0.008	-0.004	-0.002	-0.005	0.003
	(0.007)	(0.010)	(0.002)	(0.004)	(0.013)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.021***	0.027***	0.001	-0.003	-0.046***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.009)
=1 if household has either a car or motorcycle	0.013*	0.022**	-0.000	-0.005	-0.030**
	(0.007)	(0.010)	(0.002)	(0.004)	(0.013)
=1 if household has a bicycle	-0.006	-0.002	0.001	0.005	0.002
	(0.005)	(0.008)	(0.002)	(0.003)	(0.010)
Tropical Livestock Units	-0.000	0.001	-0.000	-0.001	0.001
	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)
Hectares of land cultivated by HH	0.001	-0.008***	-0.001**	0.001	0.007**
	(0.002)	(0.002)	(0.000)	(0.001)	(0.003)
=1 if HH owns a television	0.002	0.008	-0.002	0.001	-0.009
	(0.004)	(0.006)	(0.001)	(0.002)	(0.008)
=1 if HH owns a radio	0.030	0.039	-0.001	-0.031**	-0.036
	(0.035)	(0.032)	(0.011)	(0.015)	(0.052)
=1 if HH has mobile phone	-0.005	-0.007	0.003**	-0.003	0.012
	(0.006)	(0.008)	(0.002)	(0.003)	(0.010)
Spatial Variables					
=1 if HH is urban and lives in a secondary city	-0.001	-0.002	-0.001	-0.004	0.007
	(0.009)	(0.016)	(0.002)	(0.006)	(0.020)
=1 if HH is urban and lives in a primary city	-0.010	-0.071***	-0.006	0.016**	0.070***
	(0.012)	(0.020)	(0.004)	(0.007)	(0.025)
Constant	0.849**	0.773	0.078	-0.890***	0.191
	(0.359)	(0.477)	(0.071)	(0.144)	(0.735)
Observations	4,098	4,098	4,098	4,098	4,098
R-squared	0.343	0.444	0.160	0.191	0.462

Notes: Estimated with OLS. Robust standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Figure 1.1 Rural average consumption shares of processed maize, rice, processed wheat, and roots and tubers by travel time to nearest urban center

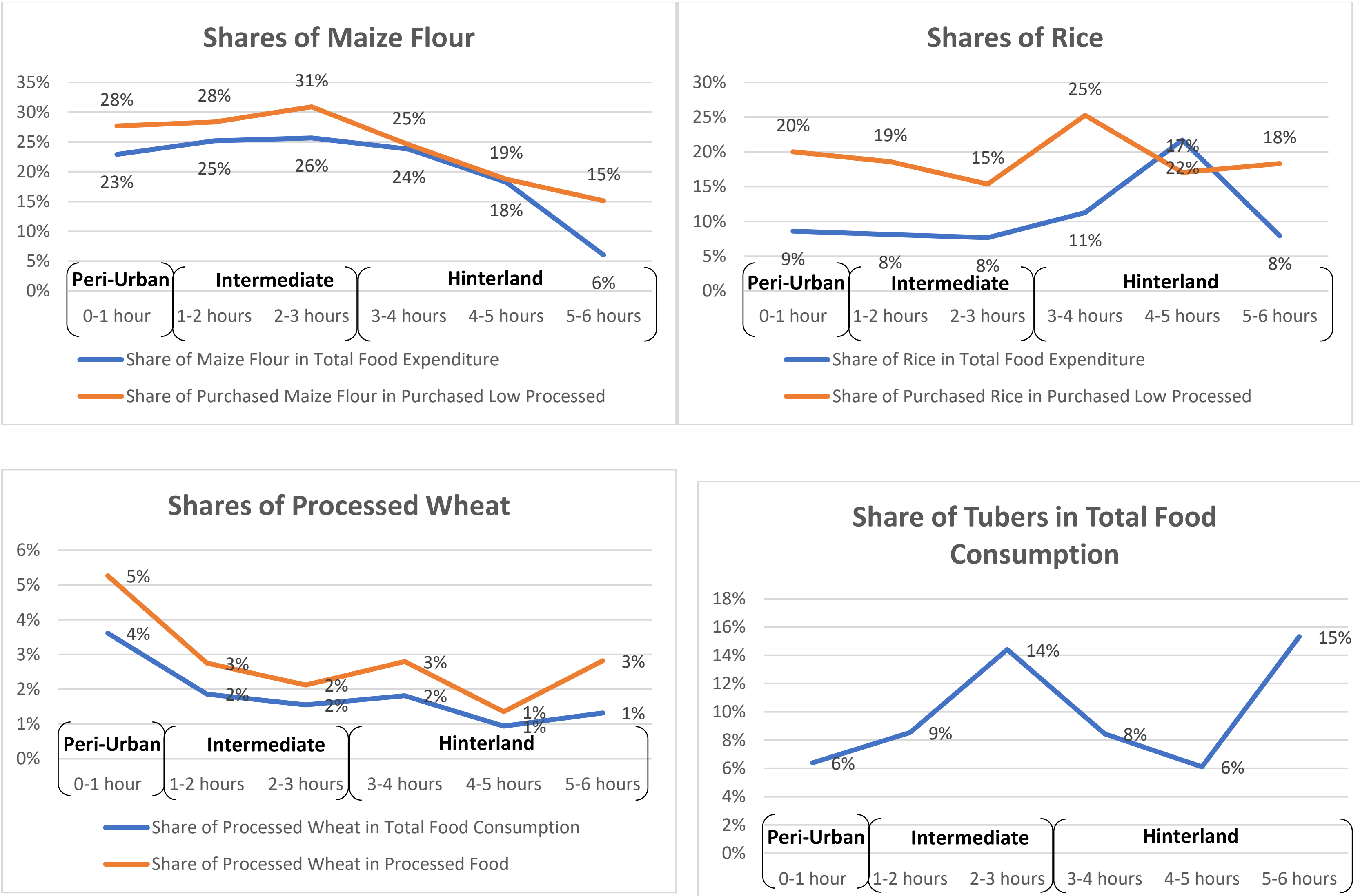


Figure 1.2 Rural average consumption shares of processed categories (regardless of source) by travel time to nearest urban center

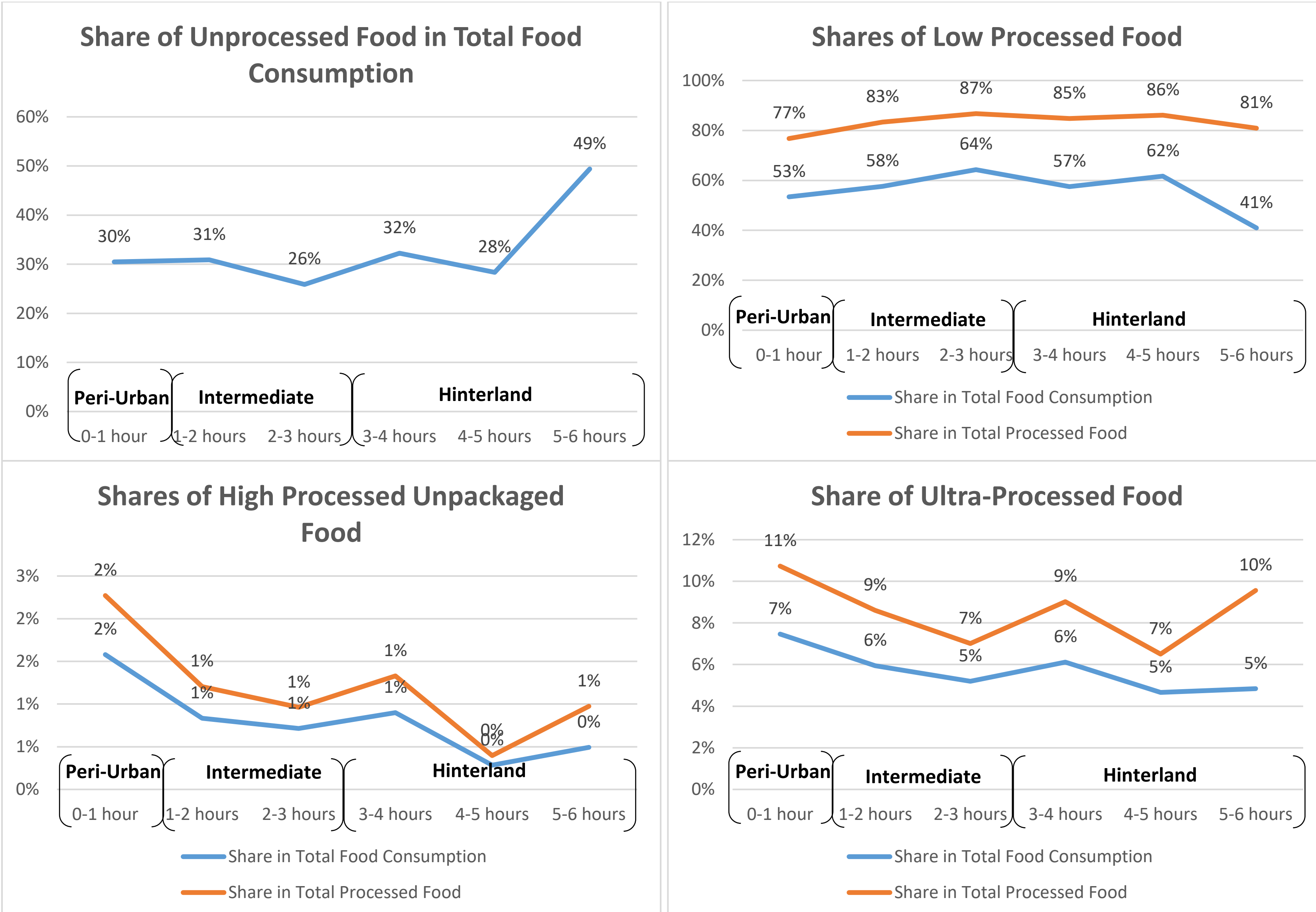
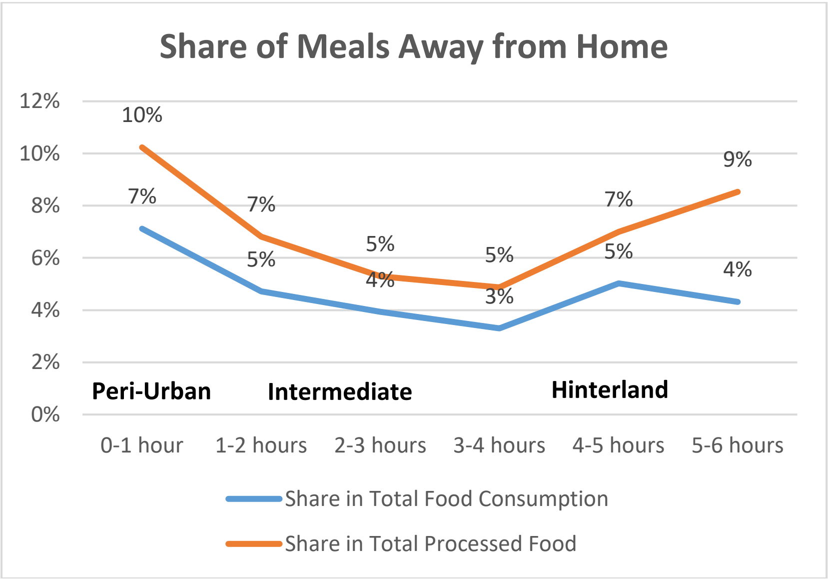


Figure 1.2, cont.



APPENDIX B: Supplemental Tables

Table 1.7 Average shares (percentages) of processed categories in total food consumption in value terms, and shares of processed grains in total food consumption: rural households

	T1*	T2	T3	Overall
N =	1,876	1,876	1,876	5,628
Average yearly food consumption per AE (USD)	$156	$240	$420	$276
Yearly total expenditure per AE (USD)	$228	$426	$953	$499
Total Food Budget Shares				
Own-produced share[a]	47.4	36.6	24.1	37.1
Total Purchased Share	48.8	60.5	73.7	59.8
Purchased to eat at home	44.2	53.7	59.8	51.9
FAFH	4.6	6.8	13.9	7.9
Unprocessed Share	**30.2**	**29.5**	**26.0**	**28.8**
Own-produced share[b]	54.8	43.8	29.5	43.9
Total Purchased Share	45.2	56.3	70.5	56.1
Purchased to eat at home	44.4	55.2	68.0	54.7
FAFH	0.8	1.1	2.5	1.4
Low Processed Share	**58.6**	**56.9**	**52.9**	**56.4**
Own-produced share	46.8	35.7	24.4	36.7
Total Purchased Share	53.2	64.2	75.6	63.2
Purchased to eat at home	53.1	63.9	73.2	62.4
FAFH	0.1	0.3	2.4	0.8
High Processed Unpackaged Share	**0.6**	**1.0**	**1.3**	**1.0**
Own-produced share	0.9	1.2	0.9	1.0
Total Purchased Share	99.1	98.8	99.1	98.9
Purchased to eat at home	92.1	95.3	95.7	94.5
FAFH	7.0	3.5	3.4	4.4

Table 1.7, cont.

	T1*	T2	T3	Overall
Ultra-Processed Share	**4.4**	**5.8**	**7.7**	**5.8**
Own-produced share	10.2	8.2	5.4	8.0
Total Purchased Share	89.8	91.9	94.6	92
Purchased to eat at home	49.7	52.3	53.3	51.7
FAFH	40.1	39.6	41.3	40.3
FAFH Meals Share	**2.5**	**3.9**	**9.9**	**5.1**
Shares of Processed Maize, Rice, and Wheat in Total Food Consumption				
Processed maize (i.e. maize flour)	**31.0**	**23.4**	**16.1**	**24.2**
Share of purchased processed maize	43.2	49.6	60.2	50.2
Share of own-produced processed maize	53.8	48.6	38.0	47.5
Share of ‘other’ (e.g. gifts) processed maize	3.0	1.8	1.9	2.3
Rice	**5.1**	**8.5**	**10.8**	**7.9**
Share of purchased rice	82.4	85.1	85.1	84.3
Share of own-produced rice	15.4	12.5	12.7	13.4
Share of ‘other’ (e.g. gifts) rice	2.3	2.4	2.2	2.3
Processed wheat	**1.1**	**2.1**	**3.2**	**2.0**
Share of purchased processed wheat	93.6	94.0	95.2	94.3
Share of own-produced processed wheat	3.2	2.9	2.4	2.8
Share of ‘other’ (e.g. gifts) processed wheat	3.3	3.0	2.4	2.9

[*] Refers to terciles calculated using total expenditure per adult equivalent. [a] These are the shares of each source within the total household food budget (will not exactly sum to 100 down columns due to exclusion of gifts/food aid share). [b] These are the shares of each source within each processed category. Source: authors’ calculations.

Table 1.8 Average shares (percentages) of processed categories in total food consumption in value terms: urban households

	Towns				Secondary Cities				Primary Cities				All Urban
	T1*	**T2**	**T3**	**All**	**T1**	**T2**	**T3**	**All**	**T1**	**T2**	**T3**	**All**	
N =	107	107	106	320	338	338	338	1,014	938	939	938	2,815	4,149
Average yearly food consumption per AE (USD)	$216	$360	$636	$408	$192	$312	$540	$348	$264	$444	$708	$480	$408
Average yearly total expenditure per AE (USD)	$364	$726	$2019	$1026	$350	$655	$1473	$790	$534	$998	$2217	$1237	$1032
Total Food Budget Shares													
Own-produced share[a]	5.6	4.3	1.3	3.8	2.5	1.6	0.9	1.7	0.9	0.2	0.4	0.5	2.0
Total Purchased Share	92.3	95.2	98.1	95.2	96.4	98.1	98.8	97.7	98.8	99.7	99.5	99.4	97.5
Purchased to eat at home	82.5	80.5	70.2	77.8	88.4	86.0	66.9	81.1	79.8	66.1	48.9	65.2	74.5
FAFH	9.8	14.7	27.9	17.4	8.0	12.1	31.9	16.6	19.0	33.6	50.6	34.2	23.0
Unprocessed Share	**25.4**	**23.1**	**20.9**	**23.1**	**26.0**	**23.7**	**19.2**	**23.2**	**23.6**	**19.4**	**15.1**	**19.4**	**21.8**
Own-produced share[b]	7.2	3.9	2.5	4.5	3.8	2.3	1.3	2.6	1.2	0.5	0.8	0.8	2.6
Total Purchased Share	92.8	96.1	97.5	95.5	96.2	97.7	98.7	97.4	98.8	99.6	99.2	99.2	97.4
Purchased to eat at home	91.5	95.5	94.6	93.9	95.4	96.7	94.0	95.4	96.4	92.2	84.7	91.4	93.5
FAFH	1.3	0.6	2.9	1.6	0.8	1.0	4.7	2.0	2.4	7.4	14.5	7.8	3.9
Low Processed Share	**54.4**	**53.3**	**44.2**	**50.7**	**59.9**	**57.8**	**45.7**	**54.9**	**48.4**	**41.8**	**33.5**	**41.3**	**48.8**
Own-produced share	6.3	5.9	1.4	4.5	1.4	1.2	1.1	1.2	0.8	0.2	0.4	0.5	2.1
Total Purchased Share	93.7	94.1	98.6	95.5	98.6	98.9	99	98.8	99.2	99.8	99.6	99.5	98
Purchased to eat at home	91.6	91.4	87.8	90.3	97.6	96.4	86.3	93.8	93.6	83.9	68.0	82.0	88.6
FAFH	2.1	2.7	10.8	5.2	1.0	2.5	12.7	5.0	5.6	15.9	31.6	17.5	9.4
Share of maize in low processed	35.5	22.1	12.9	24.3	29.4	19.9	15.1	22.6	22.3	14.6	9.9	16.5	21.3
Share of rice in low processed	17.1	24.8	22.4	21.5	23.0	27.0	25.2	25.0	26.7	26.6	21.2	25.2	23.8
Share of processed wheat in low processed	12.5	14.4	13.3	13.4	7.8	9.5	10.3	9.1	18.6	17.7	14.6	17.2	12.9
High Processed Unpackaged Share	**3.3**	**2.9**	**1.4**	**2.6**	**2.0**	**1.8**	**1.3**	**1.7**	**3.3**	**2.3**	**1.3**	**2.3**	**2.2**
Own-produced share	1.7	2.0	0.0	1.3	0.2	0.6	0.1	0.3	0.4	0.0	0.4	0.2	0.6
Total Purchased Share	98.3	98	100	98.6	99.8	99.4	99.9	99.7	99.6	100	99.6	99.7	99.4
Purchased to eat at home	92.7	95.7	99.1	95.7	93.8	98.0	93.1	95.0	95.2	96.5	92.4	94.9	95.2
FAFH	5.6	2.3	0.9	2.9	6.0	1.4	6.8	4.7	4.4	3.5	7.2	4.8	4.2
Ultra-Processed Share	**6.9**	**9.2**	**11.6**	**9.2**	**5.2**	**6.8**	**8.8**	**6.8**	**10.0**	**11.2**	**11.2**	**10.8**	**9.0**
Own-produced share	1.7	0.6	0.6	0.9	0.2	0.7	0.1	0.3	0.4	0.1	0.3	0.3	0.5

Table 1.8, cont.

	Towns				Secondary Cities				Primary Cities				All Urban
	T1*	**T2**	**T3**	**All**	**T1**	**T2**	**T3**	**All**	**T1**	**T2**	**T3**	**All**	
Total Purchased Share	98.3	99.4	99.4	99.1	99.8	99.3	99.9	99.7	99.6	99.8	99.7	99.7	99.5
Purchased to eat at home	75.4	70.3	60.6	68.7	70.0	72.3	55.3	66.2	71.5	57.4	44.5	58.0	64.2
FAFH	22.9	29.1	38.8	30.4	29.8	27.0	44.6	33.5	28.1	42.4	55.2	41.7	35.3
FAFH -- Meals share	**8.0**	**11.0**	**21.2**	**13.4**	**5.8**	**9.5**	**24.8**	**12.8**	**14.4**	**25.2**	**38.9**	**26.0**	**17.6**

[*] Refers to terciles calculated using total expenditure per adult equivalent. [a] These are the shares of each source within the total household food budget (will not exactly sum to 100 down columns due to exclusion of gifts/food aid share). [b] These are the shares of each source within each processed category. Source: authors' calculations.

Table 1.9 Average monthly mL and Tanzanian shillings consumed of sugar-sweetened beverages (SSBs) per adult equivalent

	Urban				Rural						
					Peri-urban	Intermediate		Hinterland			
	Primary City	**Secondary City**	**Town**	**Overall Urban**	**<1 hr**	**[1hr, 2hr]**	**[2hr, 3hr]**	**[3hr, 4hr]**	**[4hr, 5hr]**	**[5hr, 6hr]**	**Overall Rural**
Average monthly mL – unconditional	257	162	371	243	197	105	78	225	19	60	128
(Coefficient of Variation)	(2.84)	(2.76)	(3.58)	(3.06)	(3.33)	(3.89)	(4.57)	(4.95)	(4.54)	(2.50)	(4.16)
Average monthly mL – conditional on SSB consumption	792	574	830	750	696	528	494	1134	215	274	602
	(1.40)	(1.20)	(2.28)	(1.53)	(1.56)	(1.49)	(1.57)	(2.03)	(0.98)	(0.77)	(1.70)
Average shillings spent – unconditional	416	269	596	394	304	172	121	350	28	92	202
	(2.87)	(2.72)	(2.72)	(2.90)	(3.29)	(3.93)	(4.78)	(4.70)	(4.77)	(2.68)	(4.09)
Average shillings spent – conditional on SSB consumption	1280	950	1333	1215	1071	863	768	1761	318	422	952
	(1.41)	(1.17)	(1.66)	(1.43)	(1.54)	(1.51)	(1.67)	(1.91)	(1.08)	(0.89)	(1.66)

Source: authors' calculations.

Table 1.10 Employment variables, by gender and by rural/urban category - household averages

	Rural	Urban		
		Towns	**Secondary Cities**	**Primary Cities**
N =	5,628	320	1,014	2,815
Share of households with females in the labor force (i.e., with >0 FTEs in any of own-farm, farm wage, non-farm self-employment or non-farm wage work)	87	62	48	40
If any females in labor force, total full time equivalents (FTEs)	0.80	1.07	1.02	1.17
(Coefficient of Variation)	(0.84)	(0.74)	(0.72)	(0.73)
Share of households with males in the labor force	79	66	53	62
If any males in labor force, total FTEs	0.95	1.41	1.28	1.42
	(0.86)	(0.59)	(0.64)	(0.55)
FEMALES				
Share of households with females doing own-farm	**95**	**53**	**27**	**11**
If doing own-farm, total FTEs in own-farm	0.69	0.52	0.47	0.43
	(0.86)	(0.89)	(0.91)	(1.16)
If doing own-farm, share of female FTEs in own-farm	93	68	77	69
Share of households with females doing farm wage work	**2**	**0.3**	**1**	**0.2**
If doing farm wage, total FTEs in farm wage	0.36	0.14	1.04	1.43
	(1.17)	(0.00)	(0.55)	(0.08)
If doing farm wage, share of female FTEs in farm wage	42	27	95	92
Share of households with females doing nonfarm self-employment	**16**	**43**	**55**	**51**
If doing nonfarm self-employment, total FTEs in nonfarm self-employment	0.57	0.91	0.91	1.00
	(1.00)	(0.80)	(0.68)	(0.64)
If doing nonfarm self-employment, share of female FTEs in nonfarm self-employment	49	80	89	93
Share of households with females doing nonfarm wage work	**3**	**33**	**31**	**47**
If doing nonfarm wage work, total FTEs in nonfarm wage	1.07	1.26	1.22	1.30
	(0.66)	(0.58)	(0.66)	(0.67)
If doing nonfarm wage work, share of female FTEs in nonfarm wage work	78	90	93	95

Table 1.10, cont.

	Rural	Towns	Urban Secondary Cities	Primary Cities
MALES				
Share of households with males doing own-farm	**92**	**41**	**22**	**7**
If doing own-farm, total FTEs in own-farm	0.73	0.44	0.63	0.44
	(0.85)	(1.10)	(0.97)	(1.26)
If doing own-farm, share of male FTEs in own-farm	87	56	73	61
Share of households with males doing farm wage	**5**	**3**	**3**	**1**
If doing farm wage, total FTEs in farm wage	0.75	1.05	0.89	1.02
	(1.02)	(0.49)	(0.74)	(0.57)
If doing farm wage, share of male FTEs in farm wage	49	76	92	81
Share of households with males doing nonfarm self-employment	**19**	**40**	**36**	**30**
If doing nonfarm self-employment, total FTEs in nonfarm self-employment	0.75	1.43	1.26	1.29
	(1.05)	(0.40)	(0.55)	(0.49)
If doing nonfarm self-employment, share of male FTEs in nonfarm self-employment	52	80	92	94
Share of households with males doing nonfarm wage	**9**	**49**	**50**	**69**
If doing nonfarm wage, total FTEs in nonfarm wage	1.09	1.28	1.33	1.45
	(0.62)	(0.55)	(0.60)	(0.52)
If doing nonfarm wage, share of male FTEs in nonfarm wage	80	87	96	98

Source: authors' calculations.

Table 1.11 Descriptive statistics for rural households

	T1*	T2	T3	Overall
N =	1,876	1,876	1,876	5,628
Yearly total expenditure per AE (USD)	$228	$426	$953	$499
	(3.30)[a]	(1.85)	(6.56)	(9.79)
Yearly total expenditure per capita (USD)	$182	$343	$795	$410
	(3.59)	(2.37)	(7.41)	(10.68)
Adult equivalents	5.1	4.1	3.1	4.2
	(0.54)	(0.58)	(0.66)	(0.61)
Dependency ratio[b]	1.3	1	0.8	1
	(0.72)	(0.77)	(0.92)	(0.80)
1 if female headed household (HH)	0.25	0.23	0.25	0.24
Age of HH head	47.4	46.3	44.0	46.1
	(0.32)	(0.34)	(0.35)	(0.34)
Education level (years) of HH head	4.3	4.9	6.2	5.0
	(0.76)	(0.69)	(0.60)	(0.70)
1 if HH head is married	0.74	0.73	0.64	0.71
1 if HH has gas or electric stove	0	0	0	0
1 if HH has charcoal stove	0.20	0.35	0.57	0.35
1 if HH has firewood and coal stove	0.31	0.33	0.25	0.30
1 if HH has refrigerator, freezer, or fridge-freezer	0	0	0.06	0
1 if HH has car	0	0	0.02	0.01
1 if HH has motorcycle	0.01	0.02	0.10	0.04
1 if HH has bicycle	0.36	0.40	0.40	0.39
1 if HH farms any land	0.95	0.91	0.79	0.89
Hectares of land farmed	2.6	2.3	2.2	2.4
	(2.60)	(1.31)	(1.45)	(2.05)
1 if HH has livestock	0.40	0.35	0.26	0.34
Amount of livestock (Tropical Livestock Units) (excluding zeros)	6.1	6.7	7.3	6.6
	(2.30)	(2.05)	(2.99)	(2.42)
1 if HH has TV	0.01	0.04	0.16	0.06
1 if HH has radio	1	1	1	1
1 if HH has mobile phone	0.36	0.51	0.66	0.49
Distance to nearest urban center (town, secondary city, or primary city) (km)	93.6	91.9	90.1	92.0
	(0.64)	(0.64)	(0.68)	(0.65)
Distance to nearest primary city (km)	461.0	424.9	386.2	427.4
	(0.59)	(0.63)	(0.69)	(0.63)
Distance to nearest secondary city (km)	147.7	150.2	151.3	149.6

Table 1.11, cont.

	T1*	T2	T3	Overall
	(0.56)	(0.58)	(0.56)	(0.57)
Distance to nearest town (km)	158.0	165.7	166.1	163.0
	(0.61)	(0.62)	(0.64)	(0.62)
1 if HH is within 1 hour travel time of nearest urban center	0.22	0.25	0.30	0.26
1 if HH is between 1 and 2 hours travel time of nearest urban center	0.40	0.37	0.34	0.37
1 if HH is between 2 and 3 hours travel time of nearest urban center	0.26	0.24	0.20	0.23
1 if HH is between 3 and 4 hours travel time of nearest urban center	0.07	0.05	0.07	0.06
1 if HH is between 4 and 5 hours travel time of nearest urban center	0.02	0.05	0.05	0.04
1 if HH is between 5 and 6 hours travel time of nearest urban center	0.03	0.03	0.03	0.03

[*] Refers to terciles created using total expenditure per adult equivalent.
[a] All numbers are averages. Coefficients of variation are reported in parentheses for continuous variables.
[b] Calculated as the ratio of elderly [age 65+] and children [under age 14] to working-age adults [ages 15-64]
Source: authors' calculations.

Table 1.12 Descriptive statistics for urban households

	Towns				Secondary Cities				Primary Cities			
	T1*	T2	T3	All	T1	T2	T3	All	T1	T2	T3	All
N =	107	107	106	320	338	338	338	1,014	938	939	938	2,815
Gini coefficient for monthly total expenditure per AE	0.2	0.1	0.3	0.4	0.2	0.1	0.2	0.4	0.2	0.1	0.2	0.4
Yearly total expenditure per AE (USD)	$364	$726	$2019	$1026	$350	$655	$1473	$790	$534	$998	$2217	$1237
	(3.40)[a]	(2.02)	(9.27)	(13.25)	(3.12)	(1.90)	(5.95)	(9.50)	(3.19)	(1.81)	(6.91)	(9.94)
Yearly total expenditure per capita (USD)	$304	$610	$1812	$898	$293	$551	$1284	$678	$455	$856	$2008	$1095
	(3.87)	(2.58)	(10.11)	(14.46)	(3.47)	(2.48)	(6.26)	(9.97)	(3.42)	(2.43)	(7.22)	(10.54)
Adult equivalents	4.5	3.9	2.9	3.8	4.5	3.9	2.5	3.7	4.6	3.2	2.3	3.4
	(0.56)	(0.61)	(0.68)	(0.63)	(0.48)	(0.60)	(0.64)	(0.60)	(0.48)	(0.56)	(0.72)	(0.63)
Dependency ratio[b]	0.9	0.6	0.4	0.6	0.9	0.8	0.4	0.7	0.7	0.6	0.3	0.5
	(0.88)	(0.79)	(1.16)	(0.99)	(0.83)	(1.04)	(1.28)	(1.03)	(0.88)	(0.97)	(1.56)	(1.10)
1 if female headed household (HH)	0.32	0.26	0.31	0.29	0.33	0.30	0.31	0.31	0.27	0.22	0.25	0.24
Age of HH head	49.8	44.8	39.2	44.6	46.1	43.7	38.5	43	45.5	39.8	36.3	40.6
	(0.34)	(0.32)	(0.32)	(0.34)	(0.31)	(0.31)	(0.35)	(0.33)	(0.30)	(0.32)	(0.30)	(0.32)
Education level (years) of HH head	5.6	7	9.3	7.3	6.4	7.8	8.9	7.6	7.1	8.2	10	8.4
	(0.57)	(0.45)	(0.41)	(0.51)	(0.51)	(0.47)	(0.41)	(0.48)	(0.48)	(0.40)	(0.37)	(0.44)
1 if HH head is married	0.59	0.70	0.62	0.64	0.59	0.63	0.47	0.57	0.67	0.60	0.46	0.58
1 if HH has gas or electric stove	0	0.01	0.17	0.06	0.03	0.04	0.12	0.06	0.05	0.07	0.18	0.10
1 if HH has charcoal stove	0.72	0.87	0.88	0.83	0.88	0.95	0.89	0.91	0.91	0.88	0.76	0.85
1 if HH has firewood and coal stove	0.25	0.21	0.07	0.18	0.10	0.11	0.05	0.09	0.09	0.03	0.03	0.05
1 if HH has refrigerator, freezer, or fridge-freezer	0.04	0.16	0.37	0.19	0.10	0.17	0.26	0.17	0.21	0.31	0.46	0.33
1 if HH has car	0.03	0.01	0.14	0.06	0.01	0.02	0.05	0.02	0.03	0.02	0.11	0.05
1 if HH has motorcycle	0.01	0.09	0.13	0.08	0.01	0.05	0.08	0.04	0.01	0.01	0.03	0.02
1 if HH has bicycle	0.25	0.25	0.25	0.25	0.16	0.22	0.15	0.18	0.11	0.07	0.07	0.09
1 if HH farms any land	0.56	0.48	0.27	0.44	0.24	0.24	0.10	0.20	0.12	0.03	0.04	0.07
Hectares of land farmed	1.1	1.8	1.2	1.4	1.2	1	1.4	1.1	1.4	1.7	2.6	1.7
	(1.00)	(2.22)	(0.81)	(1.87)	(1.31)	(0.94)	(0.92)	(1.15)	(1.65)	(1.43)	(1.42)	(1.58)
1 if HH has livestock	0.07	0.15	0.09	0.11	0.04	0.03	0.02	0.03	0.04	0.01	0.02	0.02
Amount of livestock (Tropical Livestock Units) (excluding zeros)	1.5	5.5	2.7	3.8	0.9	2.2	2.8	1.7	1.6	2.4	4	2.3
	(1.16)	(0.71)	(1.42)	(1.01)	(1.09)	(0.61)	(0.51)	(0.83)	(1.71)	(1.70)	(1.32)	(1.64)
1 if HH has TV	0.13	0.38	0.75	0.41	0.25	0.43	0.54	0.39	0.41	0.56	0.72	0.56
1 if HH has radio	1	1	1	1	1	1	1	1	1	1	1	1
1 if HH has mobile phone	0.64	0.86	0.88	0.80	0.77	0.82	0.83	0.80	0.83	0.87	0.94	0.88

* Refers to terciles calculated using total expenditure per adult equivalent.

[a] All numbers are averages. Coefficients of variation are reported in parentheses for continuous variables.

[b] Calculated as the ratio of elderly [age 65+] and children [under age 14] to working-age adults [ages 15-64]

Source: authors' calculations.

Table 1.13 SUR results for Engel curve regressions of food budget shares: rural households

	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure				
Natural log of total household expenditure on food and non-food per adult equivalent	0.094	0.035***	-0.019	-0.478***
	(0.069)	(0.009)	(0.032)	(0.051)
Natural log of total household expenditure, squared	-0.007**	-0.002***	0.002	0.024***
	(0.003)	(0.000)	(0.001)	(0.002)
Average marginal effect of ln(total expenditure)[a]	-0.057***	0.001**	0.019***	0.035***
Household Demographics				
Adult equivalents	-0.003***	-0.000	-0.000	0.000
	(0.001)	(0.000)	(0.001)	(0.001)
Dependency ratio	0.013***	0.000	-0.006***	-0.015***
	(0.003)	(0.000)	(0.001)	(0.002)
=1 if the head of household is female	0.029***	-0.000	-0.018***	-0.048***
	(0.007)	(0.001)	(0.003)	(0.005)
Age of the household head	0.000*	0.000	0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	0.001	0.000**	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.052***	0.001	-0.016***	-0.060***
	(0.006)	(0.001)	(0.003)	(0.005)
Labor Variables				
Total FTEs of females in nonfarm wage	-0.003	-0.000	0.005	0.013*
	(0.010)	(0.001)	(0.005)	(0.007)
Total FTEs of females in nonfarm self-employment	0.010	0.005***	0.004	0.001
	(0.007)	(0.001)	(0.003)	(0.005)
Total FTEs of males in nonfarm wage	-0.017***	-0.000	0.008***	0.028***
	(0.006)	(0.001)	(0.003)	(0.005)
Total FTEs of males in nonfarm self-employment	0.005	0.001	0.002	0.014***
	(0.005)	(0.001)	(0.002)	(0.004)
Other Preference-Altering Variables				
=1 if HH has gas or electric stove	0.040*	-0.001	-0.006	-0.035**
	(0.024)	(0.003)	(0.011)	(0.017)

Table 1.13, cont.

	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
=1 if HH has charcoal stove	0.021***	0.005***	-0.002	-0.002
	(0.005)	(0.001)	(0.002)	(0.004)
=1 if HH has firewood and coal stove	0.021***	-0.002**	-0.006**	-0.009**
	(0.005)	(0.001)	(0.002)	(0.004)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.016	-0.008***	-0.002	-0.031***
	(0.016)	(0.002)	(0.007)	(0.012)
=1 if household has either a car or motorcycle	0.053***	0.007***	-0.012**	-0.026***
	(0.011)	(0.001)	(0.005)	(0.008)
=1 if household has a bicycle	0.004	0.000	-0.000	-0.002
	(0.005)	(0.001)	(0.002)	(0.003)
Tropical Livestock Units	-0.001***	-0.000*	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Hectares of land cultivated by HH	-0.001***	-0.000*	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH owns a television	0.009	0.002	0.003	0.002
	(0.010)	(0.001)	(0.005)	(0.008)
=1 if HH owns a radio	0.017	0.009	-0.017	0.062
	(0.071)	(0.010)	(0.033)	(0.052)
=1 if HH has mobile phone	0.012**	0.001*	-0.004*	0.001
	(0.005)	(0.001)	(0.002)	(0.004)
Spatial Variables				
Distance (in 100's of km) to nearest urban center	0.125***	-0.006***	0.001	-0.004
	(0.012)	(0.002)	(0.006)	(0.009)
Distance, squared	-0.039***	0.002***	-0.002	0.001
	(0.005)	(0.001)	(0.002)	(0.003)
Average marginal effect of distance[a]	0.054***	-0.003***	-0.003	-0.002
Constant	0.170	-0.187***	0.177	2.516***
	(0.388)	(0.053)	(0.181)	(0.285)
Observations	5,054	5,054	5,054	5,054
R-squared	0.318	0.249	0.201	0.223

Notes: Estimated with SUR. Standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.14 SUR results for Engel curve regressions of food budget shares: urban households

	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure				
Natural log of total household expenditure on food and non-food per adult equivalent	-0.041	-0.007	0.165***	-0.002
	(0.060)	(0.011)	(0.023)	(0.080)
Natural log of total household expenditure, squared	-0.001	0.000	-0.006***	0.004
	(0.002)	(0.000)	(0.001)	(0.003)
Average marginal effect of ln(total expenditure)[a]	-0.058***	-0.006***	0.015***	0.080***
Household Demographics				
Adult equivalents	0.008***	0.000*	0.002***	-0.009***
	(0.002)	(0.000)	(0.001)	(0.002)
Dependency ratio	0.044***	0.004***	-0.002	-0.064***
	(0.004)	(0.001)	(0.002)	(0.005)
=1 if the head of household is female	0.068***	0.003*	0.001	-0.123***
	(0.007)	(0.001)	(0.003)	(0.009)
Age of the household head	0.001***	0.000***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	0.000	-0.000*	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.084***	0.004***	-0.001	-0.132***
	(0.006)	(0.001)	(0.002)	(0.008)
Labor Variables				
Total FTEs of females in nonfarm wage	-0.013***	-0.002***	0.001	0.025***
	(0.004)	(0.001)	(0.002)	(0.005)
Total FTEs of females in nonfarm self-employment	0.004	0.002**	-0.001	-0.012**
	(0.005)	(0.001)	(0.002)	(0.006)
Total FTEs of males in nonfarm wage	-0.013***	-0.001**	0.000	0.023***
	(0.003)	(0.001)	(0.001)	(0.004)
Total FTEs of males in nonfarm self-employment	-0.007*	-0.001	0.001	0.014**
	(0.004)	(0.001)	(0.002)	(0.006)
Other Preference-Altering Variables				
=1 if HH has gas or electric stove	0.016*	-0.003*	0.008**	-0.035***
	(0.009)	(0.002)	(0.004)	(0.013)
=1 if HH has charcoal stove	0.091***	0.005***	-0.000	-0.160***

Table 1.14, cont.

	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
	(0.007)	(0.001)	(0.003)	(0.010)
=1 if HH has firewood and coal stove	-0.004	-0.002	-0.005	0.003
	(0.009)	(0.002)	(0.004)	(0.012)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.027***	0.001	-0.003	-0.046***
	(0.006)	(0.001)	(0.002)	(0.008)
=1 if household has either a car or motorcycle	0.022**	-0.000	-0.005	-0.030**
	(0.010)	(0.002)	(0.004)	(0.013)
=1 if household has a bicycle	-0.002	0.001	0.005*	0.002
	(0.007)	(0.001)	(0.003)	(0.010)
Tropical Livestock Units	0.001	-0.000	-0.001	0.001
	(0.003)	(0.001)	(0.001)	(0.004)
Hectares of land cultivated by HH	-0.008***	-0.001*	0.001	0.007**
	(0.002)	(0.000)	(0.001)	(0.003)
=1 if HH owns a television	0.008	-0.002*	0.001	-0.009
	(0.005)	(0.001)	(0.002)	(0.007)
=1 if HH owns a radio	0.039	-0.001	-0.031	-0.036
	(0.069)	(0.013)	(0.027)	(0.092)
=1 if HH has mobile phone	-0.007	0.003**	-0.003	0.012
	(0.007)	(0.001)	(0.003)	(0.009)
Spatial Variables				
=1 if HH is urban and lives in a secondary city	-0.002	-0.001	-0.004	0.007
	(0.013)	(0.002)	(0.005)	(0.018)
=1 if HH is urban and lives in a primary city	-0.071***	-0.006*	0.016**	0.070***
	(0.018)	(0.003)	(0.007)	(0.024)
Constant	0.773**	0.078	-0.890***	0.191
	(0.364)	(0.067)	(0.141)	(0.488)
Observations	4,098	4,098	4,098	4,098
R-squared	0.444	0.160	0.191	0.462

Notes: Estimated with SUR. Standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.15 FMNL results (average marginal effects) for Engel curve regressions of food budget shares: rural households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	0.003	-0.054***	0.002***	0.020***	0.029***
	(0.005)	(0.005)	(0.001)	(0.002)	(0.004)
Household Demographics					
Adult equivalents	0.003**	-0.004**	0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Dependency ratio	0.009***	0.017***	0.000	-0.006***	-0.021***
	(0.003)	(0.003)	(0.000)	(0.002)	(0.002)
=1 if the head of household is female	0.029***	0.018**	-0.000	-0.017***	-0.030***
	(0.007)	(0.007)	(0.001)	(0.003)	(0.004)
Age of the household head	0.000**	0.000	0.000	0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	-0.001**	0.001	0.000*	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.017***	0.042***	0.001	-0.016***	-0.044***
	(0.006)	(0.007)	(0.001)	(0.004)	(0.005)
Labor Variables					
Total FTEs of females in nonfarm wage	-0.015*	-0.001	0.000	0.005*	0.010**
	(0.009)	(0.009)	(0.001)	(0.003)	(0.004)
Total FTEs of females in nonfarm self-employment	-0.021***	0.008	0.002***	0.003	0.008**
	(0.006)	(0.006)	(0.001)	(0.003)	(0.004)
Total FTEs of males in nonfarm wage	-0.015***	-0.004	-0.000	0.008***	0.012***
	(0.006)	(0.006)	(0.001)	(0.002)	(0.003)
Total FTEs of males in nonfarm self-employment	-0.026***	0.014	0.001	0.003	0.009***
	(0.006)	(0.009)	(0.001)	(0.003)	(0.003)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	-0.004	0.027	-0.001	-0.007	-0.016**
	(0.016)	(0.017)	(0.002)	(0.005)	(0.008)
=1 if HH has charcoal stove	-0.023***	0.019***	0.004***	-0.002	0.001
	(0.005)	(0.006)	(0.001)	(0.003)	(0.004)
=1 if HH has firewood and coal stove	-0.004	0.021***	-0.001	-0.006**	-0.011***
	(0.005)	(0.005)	(0.001)	(0.003)	(0.003)

Table 1.15, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.022*	0.003	-0.005***	-0.005	-0.014***
	(0.012)	(0.013)	(0.001)	(0.005)	(0.005)
=1 if household has either a car or motorcycle	-0.028***	0.041***	0.005***	-0.011***	-0.008
	(0.010)	(0.011)	(0.002)	(0.004)	(0.006)
=1 if household has a bicycle	-0.001	0.003	0.000	-0.001	-0.001
	(0.005)	(0.005)	(0.001)	(0.002)	(0.004)
Tropical Livestock Units	0.001***	-0.000	-0.000***	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hectares of land cultivated by HH	0.001*	-0.001	-0.001***	0.000**	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if HH owns a television	-0.017**	0.016*	-0.000	0.003	-0.002
	(0.007)	(0.008)	(0.001)	(0.004)	(0.005)
=1 if HH owns a radio	-0.080	0.039	0.007***	-0.003	0.036***
	(0.049)	(0.045)	(0.002)	(0.009)	(0.013)
=1 if HH has mobile phone	-0.010*	0.012**	0.001*	-0.005**	0.002
	(0.005)	(0.006)	(0.001)	(0.003)	(0.004)
Spatial Variables					
Distance (in 100's of km) to nearest urban center	-0.047***	0.051***	-0.003***	-0.002	0.001
	(0.005)	(0.006)	(0.001)	(0.003)	(0.004)
Observations	5,054	5,054	5,054	5,054	5,054

Notes: Estimated with fractional multinomial logit (FMNL). Standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.16 FMNL results (average marginal effects) for Engel curve regressions of food budget shares: urban households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	-0.029***	-0.051***	-0.006***	0.018***	0.068***
	(0.004)	(0.006)	(0.001)	(0.002)	(0.007)
Household Demographics					
Adult equivalents	-0.000	0.009***	0.000	0.003***	-0.012***
	(0.001)	(0.002)	(0.000)	(0.001)	(0.003)
Dependency ratio	0.021***	0.048***	0.004***	-0.000	-0.073***
	(0.003)	(0.005)	(0.001)	(0.002)	(0.007)
=1 if the head of household is female	0.045***	0.047***	0.002	-0.006*	-0.088***
	(0.006)	(0.008)	(0.002)	(0.003)	(0.009)
Age of the household head	0.001***	0.001***	0.000***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	-0.000	-0.000	-0.000*	0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.040***	0.060***	0.003**	-0.011***	-0.092***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.009)
Labor Variables					
Total FTEs of females in nonfarm wage	-0.011***	-0.014***	-0.002**	-0.001	0.027***
	(0.003)	(0.004)	(0.001)	(0.002)	(0.005)
Total FTEs of females in nonfarm self-employment	0.007**	0.002	0.001*	-0.001	-0.009
	(0.003)	(0.005)	(0.001)	(0.002)	(0.006)
Total FTEs of males in nonfarm wage	-0.009***	-0.011***	-0.001**	0.000	0.021***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)
Total FTEs of males in nonfarm self-employment	-0.008**	-0.008*	-0.000	0.000	0.016***
	(0.003)	(0.004)	(0.001)	(0.002)	(0.006)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	0.013**	0.014*	-0.004**	0.005	-0.029***
	(0.006)	(0.008)	(0.002)	(0.004)	(0.011)
=1 if HH has charcoal stove	0.057***	0.071***	0.005***	-0.017***	-0.116***
	(0.006)	(0.009)	(0.002)	(0.004)	(0.012)
=1 if HH has firewood and coal stove	0.008	-0.001	-0.002	-0.005	0.001
	(0.006)	(0.010)	(0.002)	(0.004)	(0.014)

Table 1.16, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.019***	0.022***	0.001	-0.004	-0.038***
	(0.004)	(0.006)	(0.001)	(0.003)	(0.007)
=1 if household has either a car or motorcycle	0.009	0.010	0.000	-0.008**	-0.011
	(0.007)	(0.010)	(0.002)	(0.004)	(0.013)
=1 if household has a bicycle	-0.007	-0.004	0.001	0.004	0.005
	(0.005)	(0.009)	(0.001)	(0.004)	(0.011)
Tropical Livestock Units	-0.000	-0.000	0.000	-0.002	0.002
	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)
Hectares of land cultivated by HH	0.001	-0.008***	-0.001*	0.000	0.008**
	(0.001)	(0.002)	(0.000)	(0.001)	(0.003)
=1 if HH owns a television	0.002	0.010*	-0.002	0.001	-0.012
	(0.004)	(0.006)	(0.001)	(0.002)	(0.007)
=1 if HH owns a radio	0.030	0.047	-0.004	-0.037*	-0.035
	(0.039)	(0.038)	(0.011)	(0.022)	(0.072)
=1 if HH has mobile phone	-0.005	-0.006	0.003**	-0.002	0.010
	(0.005)	(0.007)	(0.001)	(0.004)	(0.009)
Spatial Variables					
=1 if HH is urban and lives in a secondary city	0.002	0.001	0.001	-0.002	-0.002
	(0.010)	(0.017)	(0.003)	(0.007)	(0.022)
=1 if HH is urban and lives in a primary city	-0.011	-0.069***	-0.003	0.016**	0.067**
	(0.012)	(0.022)	(0.003)	(0.008)	(0.031)
Observations	4,098	4,098	4,098	4,098	4,098

Notes: Estimated with FMNL. Standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here.
*** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.17 Alternative employment specification: econometric results for Engel curve regressions of food budget shares: rural households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	0.338***	0.087	0.033***	-0.024	-0.433***
	(0.086)	(0.080)	(0.010)	(0.032)	(0.095)
Natural log of total household expenditure, squared	-0.016***	-0.006*	-0.001***	0.002	0.022***
	(0.004)	(0.004)	(0.000)	(0.001)	(0.004)
Average marginal effect of ln(total expenditure)[a]	-0.001	-0.055***	0.002***	0.018***	0.036***
Household Demographics					
Adult equivalents	0.003**	-0.002*	0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Dependency ratio	0.007**	0.012***	0.000	-0.006***	-0.012***
	(0.003)	(0.003)	(0.000)	(0.001)	(0.002)
=1 if the head of household is female	0.039***	0.027***	-0.000	-0.021***	-0.045***
	(0.007)	(0.008)	(0.001)	(0.004)	(0.007)
Age of the household head	0.000	0.000	0.000	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	-0.002**	0.001	0.000**	0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.025***	0.055***	0.001	-0.017***	-0.063***
	(0.007)	(0.008)	(0.001)	(0.004)	(0.007)
Labor Variables					
Share of nonfarm wage FTEs in total female FTEs	-0.028**	0.005	-0.000	0.013*	0.011
	(0.012)	(0.016)	(0.002)	(0.007)	(0.018)
Share of nonfarm self-employment FTEs in total female FTEs	-0.024***	0.015	0.008***	0.009*	-0.007
	(0.008)	(0.010)	(0.002)	(0.004)	(0.008)
Share of nonfarm wage FTEs in total male FTEs	-0.026***	-0.043***	-0.000	0.017***	0.053***
	(0.009)	(0.011)	(0.002)	(0.005)	(0.011)
Share of nonfarm self-employment FTEs in total male FTEs	-0.035***	-0.013	0.002	0.010*	0.036***
	(0.008)	(0.011)	(0.002)	(0.005)	(0.009)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	0.005	0.039**	-0.001	-0.007	-0.036*

Table 1.17, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
	(0.016)	(0.018)	(0.003)	(0.007)	(0.019)
=1 if HH has charcoal stove	-0.022***	0.023***	0.005***	-0.002	-0.004
	(0.005)	(0.006)	(0.001)	(0.003)	(0.004)
=1 if HH has firewood and coal stove	-0.004	0.022***	-0.002**	-0.005**	-0.011***
	(0.005)	(0.005)	(0.001)	(0.003)	(0.003)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.029**	0.016	-0.008***	-0.004	-0.032***
	(0.011)	(0.014)	(0.002)	(0.006)	(0.011)
=1 if household has either a car or motorcycle	-0.023**	0.053***	0.007***	-0.012***	-0.025**
	(0.010)	(0.012)	(0.002)	(0.004)	(0.010)
=1 if household has a bicycle	-0.001	0.004	0.000	0.000	-0.003
	(0.005)	(0.005)	(0.001)	(0.003)	(0.004)
Tropical Livestock Units	0.001**	-0.001*	-0.000**	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hectares of land cultivated by HH	0.001*	-0.001*	-0.000*	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
=1 if HH owns a television	-0.014*	0.011	0.001	0.001	0.000
	(0.008)	(0.010)	(0.002)	(0.004)	(0.010)
=1 if HH owns a radio	-0.064*	0.019	0.008*	-0.023	0.060
	(0.039)	(0.053)	(0.005)	(0.017)	(0.043)
=1 if HH has mobile phone	-0.008	0.013**	0.001	-0.005**	-0.001
	(0.005)	(0.006)	(0.001)	(0.002)	(0.004)
Spatial Variables					
Distance (in 100's of km) to nearest urban center	-0.111***	0.120***	-0.006***	-0.003	-0.001
	(0.014)	(0.015)	(0.002)	(0.006)	(0.009)
Distance, squared	0.036***	-0.037***	0.002***	0.001	-0.000
	(0.006)	(0.006)	(0.001)	(0.002)	(0.003)
Average marginal effect of distance[a]	-0.045***	0.051***	-0.003***	-0.002	-0.001
Constant	-1.495***	0.210	-0.178***	0.212	2.252***
	(0.481)	(0.449)	(0.054)	(0.177)	(0.512)
Observations	5,266	5,266	5,266	5,266	5,266
R-squared	0.323	0.312	0.248	0.190	0.201

Notes: Estimated with OLS. Robust standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

Table 1.18 Alternative employment specification: econometric results for Engel curve regressions of food budget shares: urban households

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
Total Expenditure					
Natural log of total household expenditure on food and non-food per adult equivalent	-0.125**	-0.051	-0.009	0.166***	0.020
	(0.059)	(0.077)	(0.011)	(0.023)	(0.121)
Natural log of total household expenditure, squared	0.004	-0.000	0.000	-0.006***	0.003
	(0.002)	(0.003)	(0.000)	(0.001)	(0.005)
Average marginal effect of ln(total expenditure)[a]	-0.033***	-0.060***	-0.006***	0.016***	0.084***
Household Demographics					
Adult equivalents	-0.003**	0.006***	0.000	0.002***	-0.006***
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)
Dependency ratio	0.018***	0.044***	0.004***	-0.002	-0.064***
	(0.003)	(0.004)	(0.001)	(0.001)	(0.005)
=1 if the head of household is female	0.048***	0.066***	0.001	0.001	-0.116***
	(0.006)	(0.009)	(0.002)	(0.003)	(0.011)
Age of the household head	0.001***	0.001***	0.000***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# years formal schooling of HH head	0.000	0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
=1 if the household head is married	0.046***	0.082***	0.004***	-0.001	-0.131***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.010)
Labor Variables					
Share of nonfarm wage FTEs in total female FTEs	-0.012**	-0.010	-0.002	0.001	0.023**
	(0.005)	(0.008)	(0.001)	(0.003)	(0.010)
Share of nonfarm self-employment FTEs in total female FTEs	0.016***	0.021***	0.003**	-0.002	-0.039***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.008)
Share of nonfarm wage FTEs in total male FTEs	-0.012***	-0.012*	-0.004***	-0.001	0.028***
	(0.004)	(0.006)	(0.001)	(0.003)	(0.008)
Share of nonfarm self-employment FTEs in total male FTEs	-0.007	0.000	-0.001	-0.001	0.009
	(0.005)	(0.008)	(0.002)	(0.003)	(0.010)
Other Preference-Altering Variables					
=1 if HH has gas or electric stove	0.013**	0.016*	-0.003*	0.008*	-0.034***

Table 1.18, cont.

	Unprocessed	Low Processed	High Processed, Unpackaged	Ultra-Processed	Meals Away from Home
	(0.007)	(0.009)	(0.002)	(0.004)	(0.012)
=1 if HH has charcoal stove	0.062***	0.089***	0.005***	-0.000	-0.157***
	(0.006)	(0.009)	(0.002)	(0.003)	(0.012)
=1 if HH has firewood and coal stove	0.011	0.000	-0.003	-0.005	-0.003
	(0.007)	(0.010)	(0.002)	(0.004)	(0.013)
=1 if HH has a refrigerator, freezer, or fridge-freezer	0.022***	0.027***	0.001	-0.003	-0.046***
	(0.005)	(0.007)	(0.001)	(0.003)	(0.009)
=1 if household has either a car or motorcycle	0.013*	0.021**	-0.000	-0.005	-0.029**
	(0.007)	(0.010)	(0.002)	(0.004)	(0.013)
=1 if household has a bicycle	-0.006	-0.003	0.000	0.005	0.005
	(0.005)	(0.008)	(0.002)	(0.003)	(0.010)
Tropical Livestock Units	-0.001	0.000	-0.000	-0.001	0.002
	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)
Hectares of land cultivated by HH	0.001	-0.007***	-0.001**	0.001	0.006*
	(0.002)	(0.002)	(0.000)	(0.001)	(0.003)
=1 if HH owns a television	0.002	0.008	-0.002	0.001	-0.010
	(0.004)	(0.006)	(0.001)	(0.002)	(0.008)
=1 if HH owns a radio	0.033	0.044	0.000	-0.031**	-0.046
	(0.034)	(0.036)	(0.011)	(0.015)	(0.057)
=1 if HH has mobile phone	-0.007	-0.006	0.003**	-0.003	0.013
	(0.006)	(0.008)	(0.002)	(0.003)	(0.010)
Spatial Variables					
=1 if HH is urban and lives in a secondary city	-0.000	-0.003	-0.001	-0.004	0.008
	(0.009)	(0.016)	(0.003)	(0.006)	(0.020)
=1 if HH is urban and lives in a primary city	-0.007	-0.074***	-0.008*	0.013*	0.075***
	(0.012)	(0.019)	(0.004)	(0.007)	(0.024)
Constant	0.911***	0.837*	0.090	-0.898***	0.060
	(0.352)	(0.464)	(0.071)	(0.143)	(0.710)
Observations	4,147	4,147	4,147	4,147	4,147
R-squared	0.343	0.438	0.163	0.190	0.458

Notes: Estimated with OLS. Robust standard errors in parentheses. Region and month dummies were included as regressors, but the results are not presented here. [a] Average marginal effects are calculated using the *margins* command in Stata. *** p<0.01, ** p<0.05, * p<0.10. Source: authors' calculations.

REFERENCES

REFERENCES

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, *79*, 527-539.

Barrett, H.R., & Browne, A.W. (1994). Women's time, labour-saving devices and rural development in Africa. *Community Development Journal, 29,* 203-214.

Becker, G. (1965). A Theory of the Allocation of Time. *The Economic Journal, 75*, 493-517.

Cockx, L., Colen, L., & De Weerdt, J. (2018). From corn to popcorn? Urbanization and dietary change: Evidence from rural-urban migrants in Tanzania. *World Development, 110,* 140-159.

Deaton, A., & Zaidi, S. (2002). Guidelines for constructing consumption aggregates for welfare analysis. Living Standards Measurement Study Working Paper No. 135. Washington, D.C.: The World Bank.

Demmler, K. M., Ecker, O., & Qaim, M. (2018). Supermarket shopping and nutritional outcomes: A panel data analysis for urban Kenya. *World Development*, *102,* 292-303.

Kennedy, E., & Reardon, T. (1994). Shift to non-traditional grains in the diets of East and West Africa: Role of women's opportunity cost of time. *Food Policy, 19*, 45-56.

Khonje, M. G., & Qaim, M. (2019). Modernization of African food retailing and (un)healthy food consumption. *Sustainability*, *11*, 1-18.

Khonje, M.G., Ecker, O., & Qaim, M. (2020). Effects of modern food retailers on adult and child diets and nutrition. *Nutrients, 12*, 1714.

Kimenju, S.C., Rischke, R., Klasen, S., & Qaim, M. (2015). Do supermarkets contribute to the obesity pandemic in developing countries? *Public Health Nutrition, 18,* 3224-3233.

Kinsey, J. (1983). Working wives and the marginal propensity to consume food away from home. *American Journal of Agricultural Economics, 65*, 10-19.

Lewbel, A. (2006). Engel Curves. In S. Durlauf & L.E. Blume (Eds.), *The New Palgrave Dictionary of Economics, 2nd Ed*. Palgrave Macmillan UK.

Mincer, J. (1963). Market prices, opportunity costs, and income effects. In F. Carl (Ed.), Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld (pp. 34-55). Palo Alto: Stanford University Press.

Minten, B., Tamru, S., Engida, E., & Kuma, T. (2016). Transforming staple food value chains in Africa: The case of teff in Ethiopia. *The Journal of Development Studies*, *52,* 627-645.

Monteiro, C. A., Levy, R. B., Claro, R. M., Ribeiro de Castro, I. R., & Cannon, G. (2010). A new classification of foods based on the extent and purpose of their processing. *Cadernos de Saude Publica*, *26*, 2039-2049.

Nickanor, N., Crush, J., & Kazembe, L. (2019). The informal food sector and cohabitation with supermarkets in Windhoek, Namibia. *Urban Form, 30*, 425-442.

Papke, L.E. & Wooldridge, J.M. (1996). Econometric methods for fractional response variables with an application to 401(K) plan participation rates. *Journal of Applied Econometrics*, *11*, 619-632.

Pingali, P. (2006). Westernization of Asian diets and the transformation of food systems: Implications for research and policy. *Food Policy*, *32*, 281-298.

Popkin, B.M. (1994). The nutrition transition in low-income countries: an emerging crisis. *Nutrition Review*, *52*, 285-298.

Popkin, B.M., Corvalan, C., & Grummer-Strawn, L.M. (2019). Dynamics of the double burden of malnutrition and the changing nutrition reality. *The Lancet*, *395*, 65-74.

Reardon, T., Thiombiano, T., & Delgado, C. (1989). L'importance des céréales non-traditionnelles dans la consommation des riches et des pauvres à Ouagadougou. *Economie Rurale*, *190*, 9-14.

Rischke, R., Kimenju, S., Klasen, S., & Qaim, M. (2015). Supermarkets and food consumption patterns: The case of small towns in Kenya. *Food Policy*, *52*, 9-21.

Senauer, B., Sahn, D., & Alderman, H. (1986). The effect of the value of time on food consumption patterns in developing countries: Evidence from Sri Lanka. *American Journal of Agricultural Economics*, *68*, 920-927.

Sibhatu, K.T., & Qaim, M. (2017). Rural food security, subsistence agriculture, and seasonality. *PLoS ONE*, *12*, 1-15.

Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015a). The rise of a middle class in Eastern and Southern Africa: Implications for food system transformation. *Journal of International Development*, *27*, 628-646.

Tschirley, D., Snyder, J., Dolislager, M., Reardon, T., Haggblade, S., Goeb, J., Traub, L., Ejobi, F., & Meyer, F. (2015b). Africa's unfolding diet transformation: Implications for agrifood system employment. *Journal of Agribusiness in Developing and Emerging Economies*, *5*(2), 102-136.

Vandercasteelen, J., Beyene, S.T., Minten, B., & Swinnen, J. (2018). Big cities, small towns, and poor farmers: Evidence from Ethiopia. *World Development*, 106, 393-406.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*, 348-368.

ESSAY 2: THE POOR DO NOT PAY MORE: NEW EVIDENCE FROM TANZANIA

2.1 Introduction

Do the poor pay more? For decades, researchers have pondered this question, considering the prices faced by the poor for both food and non-food goods and services. In this paper, we revisit the question of whether the poor pay more using data collected from Tanzanian food consumers. We analyze purchase patterns and prices and segment households by the size of their city (for urban households) and by urban proximity (for rural households). In short, we find little evidence that the poor are paying more for their food.

The bulk of the 'poor pay more' literature has come in two broad waves. The first asked whether the poor pay more in developed countries, while the second revisited the question in a developing country context. The first strand of literature from developed countries (primarily the United States) had two sub-strands: an early wave in the 1920s and 1940s, and a more recent wave. The early wave (which occurred in a situation that is similar to what developing countries are experiencing today) was written when patterns of employment changed and workers began to be paid monthly, rather than daily, and urbanization was rapidly taking place. Rural areas of the country were still largely traditional (i.e. in a subsistence farming situation). Before the supermarket revolution happened in the US, there was a focus on encouraging families to buy and prepare nutritious food, and analyses looked at how this could be done in the most cost-effective manner (for one example, see Dantzig's 1991 discussion of George Stigler's application of linear programming to nutrition).

After supermarkets came onto the scene, concern arose over so-called food deserts, and hence the second sub-strand of the developed country literature. This strand often jointly considered geography, modern retail accessibility, and consumers' socioeconomic status. More specifically, this research highlights the lack of supermarkets and other large chain stores (which, by economies of scale, often offer lower prices) in urban centers, which in turn limits access and availability of affordable and nutritious food for poor consumers, many of whom lack the ability to travel to the suburbs to shop at such stores. The poor, then, are often forced to pay more for the same products at smaller-format retail or convenience stores, or travel outside their immediate area of residence to seek lower prices at suburban supermarkets. Many of these studies use retail data rather than data on consumer expenditures (Chung and Myers 1999, Frank et al. 1967, Kunreuther 1973, MacDonald and Nelson 1991), but there have also been consumer-level studies such as that of Beatty (2010) in the UK and Orhun and Palazzolo (2018), who study toilet paper purchases in the US. This wave in the US and UK tended to focus only on urban households, and lacked differentiation between different city sizes (which is also a gap in the developing country literature, discussed next).

The second more recent strand of literature from the 2000's and 2010's revisits the question of whether the poor pay more, generally in a developing country context. As in the first strand of literature, the second strand looked at urbanization and employment differentiation as drivers of a price differential between the poor and the rich. As more people started working in professional jobs that paid a monthly salary, and as the food industry developed and packaged food (often sold in larger units) became more common, the concern emerged that the poor were stuck buying small (usually unpackaged) units. Because the poor had low liquidity from hourly or daily wage work, and with a poor or nonexistent credit availability in much of the developing world, it was thought that the poor were forced to buy tiny amounts of food more frequently. Most of these studies focus on whether bulk discounts exist, and if so, whether liquidity constraints prevent the poor from buying in bulk and taking advantage of the resulting lower unit costs (Attanasio and Frayne 2006, Mussa 2015, Rao 2000). Such an inability to benefit from bulk discounts is one driver of what Mendoza (2011) terms the ‘poverty penalty.’[7]

While the second wave of literature often did control for differences in price in rural and urban areas, it did not account for differences in types of rural areas (e.g. zones defined by distance to the nearest urban area). The closest attempt was that of Attanasio and Frayne (2006, p. 17) who included dummies for residence in a rural center and residence in a “rural dispersed area,” but the definitions of each are not clear and do not appear to systematically segment rural areas. This is important because the food environment can change over different rural zones. When a food product is not produced locally, higher transaction costs – including not only transportation costs to move the product from one place to another, but also costs associated with the number of agents along the supply chain that handle the product - likely inflate the price. On the other hand, for products that are produced locally, transport costs are probably lower, and the value chain is likely to be ‘disintermediated’ in the sense that fewer midstream agents (such as traders) are needed to deliver the product.

Similarly, both the developed and developing country literatures neglected to analyze the urban data by city sizes. Just as with rural zones, the size of the city may be correlated with the food environment. For example, the presence of modern retail outlets like supermarkets may be strongly correlated with the size of the city. Additionally, city size could be positively correlated with economies of scale and economies of agglomeration. Greater economies of scale are associated with greater production efficiency, which would be expected to exert downward pressure on prices. In a similar way, economies of agglomeration (in which many firms in the same industry cluster in the same general geographic area) generally improve the efficiency of

[7] Other factors behind the poverty penalty are the size of food stores available to poor consumers and various types of market failures.

production, which would again result in lower prices. Most early literature on developing countries did not differentiate urban areas, but a spate of recent research by Christiaensen (who studied migration, labor markets, and poverty reduction) and Qaim (who studied supermarkets and their nutritional effects) has demonstrated the importance of secondary cities and smaller towns relative to primary cities (Christiaensen and Todo 2014, Christiaensen et al. 2013, Christiaensen and Kanbur 2017; Demmler et al. 2018, Kimenju et al. 2015, Rischke et al. 2015).

A third very recent wave of literature, exemplified by Bai et al. (2021), studies the cost and affordability of nutritionally adequate diets (vs. calorically adequate diets) across the world. As in the 1920's discussion from a century ago, the concern is that the poor cannot afford a nutritious diet, and indeed the conclusion reached by Bai and his colleagues is that nutritionally adequate diets (that meet the daily requirements for micro- and macronutrients) are often unaffordable in low-income countries.

In this paper, we revisit the same drivers that have traditionally motivated the poor pay more debate (i.e., income and unit size purchased) and condition them on the new factors (i.e., urban and rural differentiation) whose importance is increasingly being recognized in development research and policymaking. In this sense, our study complements that of Bai et al. (2021) who study the cost of nutritious diets by levels of national income, rural differentiation (distance to cities), services share of labor, and urban share of the population – all of which are at the meso level. Our analysis adds most of these same angles, but at the micro level. We account for differences across city size (for urban households) and for differences across rural zones (defined by a village's distance to the nearest city), neither of which has been done in the poor pay more literature. We also directly test the correlations between engagement in work associated with different pay frequencies (hourly/daily wage work, monthly wage work, and own-farm work) and the transaction-level quantity purchased.

Our research questions are as follows. First, is there dietary differentiation between the poor and the rich, among the set of food items we study?[8] We find that the answer is no – the food budget shares of ten of the top food items in Tanzania are very similar between rich and poor. The main difference is that the share of rice is slightly higher than the share of maize flour in the urban areas, while the reverse is true in rural areas. The shares of the other nine items are remarkably similar across and between rural and urban zones, and even across income terciles within each zone.

Second, is it true that the poor buy smaller units on average? We find that, generally, no – the poor do not buy smaller units. Descriptive statistics show that for most products, the average purchase quantity is the same for households in the third and first total expenditure terciles. Additionally, though the econometric findings suggest the rich do statistically buy more on average, it is not a meaningfully larger amount. Fascinatingly, for rice and maize flour, the rich in both rural and urban areas buy a greater share of total quantity in smaller unit sizes (i.e., 1kg or less), which defies conventional wisdom.

Third, controlling for income, does the timing of pay affect the unit size purchased? Our results suggest that the answer is no, with the following exceptions. Rural households purchase more tomatoes and onions on average with higher shares of daily wage, monthly wage, and own-farm work (relative to weekly wage work). Urban households purchase more rice on average with higher shares of monthly wage and own-farm work, and more maize flour with higher shares of own-farm work, relative to weekly wage work.

[8] The set of food products includes rice, maize flour, cooking oil, tomatoes, onions, spinach/lettuce, other leafy vegetables, dried fish, beef, and fresh cow milk. See sections 3 and 4.2 for further information.

Fourth, do bulk discounts exist? The literature is quite robust in its findings that purchasing a larger unit size, all else equal, reduces the unit price paid. We find that, beyond a certain unit size threshold, there is very little additional bulk discounting (with the exception of green leafy vegetables and certain animal proteins).

Fifth, controlling for unit size, spatiality (for rural households, the distance to the nearest urban area; for urban households, the size of the city), and product, do the poor pay more? Surprisingly, we find the poor do *not* pay more. In fact, the results suggest that the poor actually pay less on average than the rich, but again, the magnitudes are so small that we conclude the poor basically pay the same.

The paper is organized as follows. Section 2 describes the data we use. Section 3 presents descriptive results. Section 4 outlines our econometric strategy and section 5 presents the econometric results. Section 6 concludes.

2.2 Data

Our data come from the most recent Tanzania Household Budget Survey (HBS), a detailed cross-sectional and nationally representative survey conducted from October 2011 to October 2012. A total of 10,186 households in mainland Tanzania were interviewed; of those, HBS classified 59% rural and 41% urban. For the sample we ended up using, first we eliminated households from the original sample for the enumeration areas (EAs) of which we could not compute spatial data from Google Maps.[9] That produced a sample of 9,788. We then eliminated from the latter households that did not purchase any of the 16 products we examine in our analysis

[9] We included only households in EAs for which the distance data (distance to cities of different sizes) could be computed. There were some EAs in mountainous areas; Google Maps could not calculate the driving route from the EA to the nearest city, so we dropped those households.

(discussed below). That left a sample of 8,839 households. Of these, 5,078 (57%) are rural, and 3,761 (43%) are urban.

Households were interviewed in each enumeration area throughout the twelve-month period, so there is no bias resulting from a correlation between space and time. The HBS collected information on, *inter alia*, household members' education, farm and nonfarm employment, assets, and food and nonfood consumption. Food consumption includes: own-produced food in the household's farm or garden; food purchased and consumed at home; and food purchased from the food service sector and consumed away from home.

The sampling procedure of the HBS used a stratified two-stage sample design and a sampling frame based on the 2002 Tanzania Population and Housing Census National Master Sample. First-stage sampling was done within three geographic strata: Dar es Salaam, other urban areas, and rural areas. The Tanzania National Bureau of Statistics (NBS) defines the "other urban" domain as cities and towns with high population density, level of economic activity, and level of infrastructure. Conversely, the rural domain includes farms and traditional areas with low population density, level of economic activity, and level of infrastructure. In the first stage, sample EAs were chosen within each stratum using probability proportional to size. A listing of households was then conducted within each sample EA. In the second stage, 26 households were randomly selected within each EA.

The household-level food consumption data are drawn from a 28-day diary. Respondents were instructed to record all food consumed (in unit and value terms) and the source of the food (own-produced, purchased, or received as a gift, payment in kind, or food aid) by all members of the household during the day, for 28 consecutive days. In total, consumption data on 183 unique products was collected; this is about twice the number of types of food collected by the Tanzania

LSMS surveys. For illiterate households, enumerators visited daily to record consumption; for other households, enumerators checked in every few days. For food that was not purchased, the household was asked to estimate the monetary value of the food in Tanzanian shillings. Diary transactions in which the product was produced by the household or received as a gift, payment in kind, or as food aid are dropped from the analysis.

We augmented the HBS survey data set with data on spatial variables that we calculated as follows. First, we used 2012 census data to obtain population numbers for, and the administrative classification of, all 33 urban areas. (Tanzania considers three categories of urban area: cities (population above 500,000, from here on referred to as "primary cities"), municipalities (population between 100,000 and 500,000, from here on referred to as "secondary cities"), and towns (population between 20,000 and 100,000) (United Republic of Tanzania 2016). The census numbers indicate that 46% of the urban population in Tanzania resides in primary cities, 41% in secondary cities, and 13% in towns. Using GPS coordinates for each of the surveyed EAs, we then calculated the distances (in kilometers) and travel time (in hours) from each of the 395 EAs to each of the 33 urban areas with the HERE API (application programming interface).[10] Thus, for each EA, we know the distance and actual travel time (via road) to the nearest town, secondary city, and primary city. We then categorized rural EAs based on this travel time: *peri-urban* EAs are less than one hour away from the nearest urban center; *intermediate* EAs, between 1 and 3 hours from the nearest urban center; and *hinterland* EAs, 3 or more hours away from the nearest urban center.[11] Additionally, for urban EAs, we know whether they are located in a town, a secondary city, or a

[10] developer.here.com

[11] We adapt for Tanzania the rural categories of World Bank (2009), which defines peri-urban rural areas as being within one hour of travel time to the nearest city of at least 50,000 people; rural hinterland as being between one hour and 6 hours of the nearest city of >50,000 people; and remote rural areas as being 6+ hours away. In our analysis data, just 11% of rural households are 3 or more hours away from the nearest urban area (in what we call the rural hinterland).

primary city. Together, these three categories each of rural and urban EA are referred to in the remainder of the paper as the geographic zone of the household.

2.3 Descriptive Results

Tables 2.1-2.3 and Figures 2.1 and 2.2 present descriptive results. Table 2.1 shows the value shares of each of the 10 most important (in expenditure terms) products in the Tanzanian purchased-diet across geographic zones and total expenditure terciles.[12] The products include: (1) staples including rice, white maize flour[13], cooking oil, fresh tomatoes[14], and onions; (2) non-staples including spinach/lettuce[15], other leafy vegetables such as amaranth, dried fish, beef with bones, and fresh cow milk. We include tomatoes and onions in the "staples" group because these vegetables are consumed (generally in sauce or stew form) at most meals. Tables 2.2 and 2.3 show, by zone, tercile, and product, the average number of purchases made, the average quantity purchased in each transaction, the average price (in USD) per kilogram or per liter, and the ratio of prices paid by the lowest tercile to the prices paid by the highest tercile. In this paper, we use "unit size" and "quantity" interchangeably because key informants told us (and our own research confirms – see also the third essay in this dissertation) that the majority of maize flour sold in 2011-2012, when this data set was collected, was sold unpackaged. Hence, if a household reports buying 5 kilograms of maize flour, we assume they bought 5 kg in one transaction (rather than buying five 1 kilogram scoops of maize flour).

Six findings stand out.

[12] The terciles are calculated by taking the sample only for that zone and ranking the total expenditure per adult equivalent (AE) of each household and then dividing that sample into terciles. Thus, a lowest tercile household in an urban area may be in the second or highest tercile in a rural zone so the terciles are not comparable across zones and are thus just relative measures in a given zone.
[13] Brown maize flour (*dona*) is not reported in the HBS.
[14] Processed tomatoes are not in our list because they constitute only a tiny share of purchases.
[15] Spinach and lettuce are aggregated in the survey questionnaire and so cannot be examined separately.

First, in contrast to the common image that urban and rural consumers have widely differing diets, implying different purchased-food baskets, we found that the purchased-diet composition of rural and urban households is similar. Most products have roughly the same share of total food purchases in rural and urban areas, with several exceptions. Rice is more consumed in urban than rural areas, although in urban areas rice purchase is not correlated with income. Maize flour has a higher budget share in rural than urban areas and among the lowest tercile in all zones but the hinterland zone. Beef's share is only a bit higher for urban households (10%) compared to rural (8%) and is only weakly correlated with incomes in urban areas.

Second, traditional retail far dominates in Tanzania food retail; supermarkets have only a tiny share. Small shops tend to sell non-perishables like staple grains and oil, and wet markets and street vendors sell perishables like fruits and vegetables and milk; they also sell non-perishable goods, but consumers tend to buy the latter mainly at small shops. In all zones, between 50% and 55% of purchase transactions were made at shops, with the rest mainly at markets. Few transactions occurred at street vendors (around 1% in urban and 3% in rural areas) or at supermarkets (0.02% in both urban and rural). Hence, we ignore these categories in the descriptive tables, and in the regressions, merely lump the few observations on supermarkets with shops.

Tables 2.12 and 2.13 in Appendix B show where households buy different products. Regardless of the tercile, over 50% of rice, maize flour, and cooking oil is purchased at shops (*duka*). By contrast, most purchases of vegetables, dried fish, and beef were made at markets. Fresh milk purchase is an exception with more diversified retail sources; for rural households, 43% of milk purchases were from street vendors (reaching 50% in the rural hinterland who may not be located near markets and may be serviced by street vendors who come to them). But in cities 82% of milk was sold in shops (often with refrigerators), with only a small share for street vendors.

Third, contrary to the common image that the poor buy frequently and the richer households less frequently, a crucial finding from our data is that purchase frequency does not differ much over terciles on average. Households buy the main items in the diet (maize flour, tomatoes, onions, and greens) with similar frequency over items: every 3-4 days in rural areas, and 2-3 days in urban areas. In rural areas, the lowest tercile households buy maize flour, tomatoes, onions, and greens (plus oil and dried fish) with the same frequency as do the upper tercile households. The latter even buy some products more frequently than do the poorer households, including rice, tomatoes, onions, beef, and milk (Tables 2.2 and 2.3).

Fourth, contrary to the common idea that richer households buy larger unit sizes than poorer households, a crucial finding here is that lower and upper tercile households tend to purchase similar unit sizes, proxied here by similar quantities per transaction. That is, while the literature tends to posit that the rich buy large units and the poor small, we find they both tend toward smaller units bought frequently. The main exceptions include cooking oil, which is purchased in smaller quantities by both rural and urban poor households, and tomatoes and onions, which the poor tend to buy in smaller quantities on average. Specifically, between 82% and 89% of cooking oil transactions in the upper rural terciles are for quantities less than or equal to a quarter of a liter and around 95% of the poor tercile's transactions are for this small quantity; these shares are similar for urban households. Rural poor households tend to purchase around 27% of their tomatoes in quantities of 250 grams or less, whereas this share for rich households is between 14% and 20%; urban poor buy around 20% in such small quantities while the rich buy at most 10% in small quantities. Lastly, the same shares for onions are between 43% and 60% for rural poor and between 26% and 50% for rural rich, and between 62% and 69% for urban poor and between 37% and 52% for urban rich (Tables 2.2 and 2.3).

Importantly, our results show that very few purchases are made in very small or very large units. Very little is purchased in very small units (less than 0.25kg or liters), with the main exceptions of cooking oil, tomatoes, and onions (Tables 2.2 and 2.3). Interestingly for rice and maize flour, it is surprisingly the rich who are more likely to buy in smaller units than do the poor. Across terciles in each of the rural zones and in primary cities, households in the upper tercile conduct a greater share of transactions than the poor in which they purchase 1 kilogram or less of rice or maize flour. For upper tercile households in primary cities, this share is 100% (Table 2.2). Additionally, in most zones, the richer households are buying more maize flour in quantities of 0.25kg or less; as we will show below, this implies that the rich in these zones are paying a higher price for maize flour.

Moreover, even rice and maize flour purchases are very seldom made in large quantities (greater than or equal to 5 kilograms) (Tables 2.2 and 2.3). Just 3% (2%) of rice transactions in rural (urban) areas are for 5kg or more, and just 6% (2%) of maize flour transactions in rural (urban) areas are for 5kg or more. These shares are not correlated with income tercile, which challenges the conventional wisdom that rich households stock up on large quantities of (non-perishable) grains.

Fifth, a central finding of the paper is that the average price paid by the lower and upper terciles is very similar. The price paid by the poorest tercile (T1) is hardly distinguishable from the price paid by the richest tercile (T3) for all the non-perishable staples (rice, maize flour, and cooking oil) in all zones. Only for onions and dried fish (in rural areas) and onions, spinach/lettuce, other leafy vegetables (in secondary cities) and beef (in towns) do the poor pay more than the richer households, but generally not more than 10% more, so the poor's prices are still very close to the rich's. For tomatoes, onions, spinach/lettuce, other leafy vegetables, beef, and milk, the rural poor actually pay less (at most about 25% less but generally up to 20% less), and for milk, the urban poor pay less (about 5%) than the richer households (Tables 2.2 and 2.3).

Sixth, the above results show that the poor in fact do not pay more for the main foods than do the richer households (the opposite of the common image that the poor pay more); our results make sense in light of the above findings that the poor were not in general buying in smaller units than the rich and thus appear to not be missing out on a bulk discount. We test the latter point specifically. We find that there is very little bulk discounting for most products except for leafy green vegetables after the initial step from the very small unit to a small unit; after that step the buyer generally does not get a discount. Recall that above we showed that there are very few purchases of the very small unit (0.25 kg and less). Figures 2.1 (rural) and 2.2 (urban) show the price per kg for different unit size categories, relative to the less than 0.25 kg or liter category, for each product. (Table 2.14 contains the same information.) In both rural and urban areas, all products except for rice, beef, and milk have a "threshold" discount where the price per kg drops quickly as quantity increases from less than 0.25 to between 0.25 and 0.5 kilograms or liters. For staple foods except rice (which does not exhibit any bulk discounting), there is very little "post-threshold" (0.25 kg) discount, although there is a large reduction in price moving from a unit size of less than 0.25 kg to more than 0.25 kg. The latter would be important, and enough to ensure that the poor pay more, were either (1) the poor more apt to buy the unit of less than 0.25 kg and (2) it common for that small size to be bought. However, we showed that the former is true for only tomatoes and onions, and the latter is true only for tomatoes and onions (the poor tend to buy these vegetables in small quantities) and for cooking oil (most households, poor and rich, buy small quantities of oil). Only for the non-staple foods of leafy green vegetables and animal products (dried fish, beef, and milk) is there is a steep "post-threshold" discount whereby the price per kg continues to fall steeply as the quantity purchased increases. However, again, for non-staples the poor are not in general more apt than the rich to buy small unit sizes and thus less able to enjoy those discounts.

2.4 Econometric Methods

Our econometric strategy aims at answering two key questions. The first is whether bulk discounts exist for our chosen set of products. It is generally assumed, and confirmed in empirical research, that purchasing a greater quantity leads to a decrease in the unit price paid, all else equal (e.g., Beatty 2010, Dillon et al. 2021). Hence, for each product, we regress the natural log of the price per kg or per liter on quantity and other control variables including total expenditure and its square, the household's total food purchased share, and a vector of supply shifters (for more information, see section 4.1). Knowing whether bulk discounts exist is an important piece of the puzzle because it can help explain, if the poor do indeed pay more, why this might be the case. The second key question is whether the poor pay more than average for their specific food basket. The average cost of the food basket is calculated by month and by geographic zone. We then define an expensiveness index which is the ratio of the household's actual cost of the basket to the average cost, and this is regressed against total expenditure and supply shifters (the full process to compute the expensiveness index is described in section 4.2). These estimates help us answer the question if the poor pay more for a given quantity of a given set of products, conditional on spatial and temporal variables.

2.4.1 Bulk Discount Equation

We estimate the following equation to test, separately for each product, for bulk discounts. Note that this equation is at the transaction level:

$$\ln(v_{it}) = \alpha + \theta \ln(q_{it}) + \rho \ln(x_i) + \delta(\ln(x_i))^2 + \mu share_i + \boldsymbol{Z_{it}^s \beta} + e_{it} \qquad (1)$$

The logged price per kilogram or liter v for household i, in transaction t, is a function of: (a) the natural log of quantity purchased (which is used in this paper as the unit of the transaction) q[16]; (b) total household consumption (purchases of food and nonfood items plus the value of own-produced food) per adult equivalent x, as well as its square; (c) the share (in value terms) of purchased food in total household consumption; and (d) a vector of supply shifters $\boldsymbol{Z^s}$. The key parameter of interest is θ. Our null hypothesis is that $\theta = 0$, and our alternative hypothesis is that $\theta < 0$ (i.e., the larger the quantity purchased, the lower the price per kg). Quantity enters the equation in logged form to account for diminishing marginal discounts. We include total purchased share because we expect a lower purchase share is the same as saying the household depends less on the market for food and thus may be more willing to bargain for a lower price as it has a "fallback" position in general (although not necessarily for the product in question).

As supply shifters, we use dummy variables for: (1) the type of retail in the transaction (shop (traditional shop (*duka*) and supermarket – we lump supermarkets into shop because there are very few supermarket observations), street vendor; and market stall (as the base category); zone dummies (for city size category and rural areas including peri-urban, intermediate, and hinterland); region dummies for the 21 administrative regions in Tanzania; month dummies; and region by month dummies.

The data have a panel structure (where the cross-sectional unit is a unique combination of household, product type and transaction date). We used a pooled Two Stage Least Squares (2SLS)

[16] As mentioned above, key informants told us that the majority of maize flour sold in Tanzania in 2011-2012 was sold unpackaged. Hence, we test whether an increase in the quantity purchased (which is analogous to an increase in the unit size), all else equal, is correlated with a decrease in the price per kilogram/liter.

estimator.[17] Because we suspected that quantity purchased is endogenous to price per kg, we instrumented for quantity by using the following as instrumental variables.

The first IVs are the following set of demographic variables: (a) the size of the household in adult equivalents; (b) the dependency ratio (i.e., the ratio of the number of children under age 15 and age 65+ adults to the number of working-age adults in the household).

The second IVs are the set of variables capturing shares in total household FTEs (full time equivalents of employment) of: (a) daily/hourly wage work and self-employment (which we liken to daily wage work because most self-employment is small scale informal firms with a small "take home" profit each day); (b) monthly salary work; and (c) own-farming (which we assume to pay out in the most 'lumpy' way when the harvest is sold per year or per season). Figure 2.3 shows the average household shares of each of six FTE categories (and Table 2.15 shows FTE shares by total expenditure tercile). We find that own-farm FTEs as a share dominates in all rural zones, but with a jump in non-farm self-employment in the peri-urban zone. In urban zones, non-farm self-employment work is on average between 43% and 50% of total household FTEs, with a steady rise in monthly wage (salary) work from towns to primary cities. Table 2.15 shows that the poor in every rural zone rely much more heavily on own-farm work, whereas non-farm self-employment and monthly wage work become more important for the upper two terciles in all rural zones. Hence there does not appear to be a clear correlation between income and more liquid cash flows at least in rural areas, as is commonly assumed. In contrast, we observe a stronger correlation

[17] Because of the panel nature of the data, we also tried to estimate the equation with Mundlak-Chamberlin correlated random effects (CRE), whereby the time averages of all time-varying variables are included as independent variables. However, we only had two time-varying explanatory variables (quantity purchased in the transaction and retail outlet), and these are highly correlated with the transaction (time)-varying variables, resulting in multicollinearity. Recall the descriptive results showed that households, when buying a certain product, tend to buy the same quantity every time they make a purchase, and to buy it at the same type of retail outlet, thus creating a lack of variation for these variables for a given household.

between monthly salaried work and income in the urban areas. This correlation is sharpest in the towns, where on average 48% of all household FTEs in the highest tercile come from monthly wage work, and just 17% of all FTEs in the lowest tercile come from salaried work. In primary cities we even see that self-employment even falls from the poorest to the richest tercile.

The third IV is used only in the urban regressions: we use a dummy variable equal to 1 if the household owns a refrigerator or a freezer; in the data, no rural households owned a refrigerator/freezer. Having a refrigerator grants the possibility of buying larger quantities of perishable items like vegetables.

We justify the instruments as follows. First, the size of the household likely influences how much is bought in a given visit to the shop or market. Second, a higher dependency ratio is correlated with less purchase of meals away from home so that more would be bought in a shop or market transaction for home cooking.

Third, we use the FTE variables to reflect the periodicity of cash income and thus a proxy for a liquidity constraint on how much the household might buy in a given transaction. The "poor pay more" literature has traditionally posited that the poor would typically only earn cash daily in small amounts (rather than in a large monthly salary) and buy much smaller units than the rich and those units would have higher price per kg and less discount (Attanasio and Frayne 2006, Mussa 2015, Rao 2000). The FTE shares test this directly. The traditional hypothesis might be proved wrong simply because it might be found that the poor save and then buy a larger unit. However, our main reason for hypothesizing that the FTE shares will not affect discounts is that we showed in the descriptive section that there is actually little variation in unit size (and frequency of transactions) over income strata.

The results of a weak IV test, a test for endogeneity, and a test for overidentification are found at the bottom of Tables 2.4 (for rural households) and 2.5 (for urban households). These tables also present the first-stage regression results, which are discussed in detail in section 5.1.1. The weak IV test was developed by Montiel Olea and Pflueger (2013) and is robust in the case of clustered standard errors. If the effective F statistic does not exceed some fraction τ of a "worst-case" benchmark, then the null hypothesis is not rejected and the instruments are deemed weak (Bennett et al. 2017, Montiel Olea and Pflueger 2013). In all regressions except the urban milk model, the effective F statistic exceeds the τ = 30% critical level (in the rural milk model, the effective F exceeds the τ = 20% critical value). Hence, except for the milk regressions, the IVs appear to be sufficiently strong. The endogeneity test was performed using the estat(endog) post-estimation command in Stata. In most models except for the milk models, the rural beef, and the urban dried fish, the null hypothesis that the quantity variable can be considered exogenous is roundly rejected. Finally, the results of the overidentification test (performed using estat(overid) in Stata) suggest that the oil and milk regressions (as well as the rural tomato and urban beef) regressions may be overidentified. Nonetheless, to be consistent, we report the 2SLS results for all food products in the main text. OLS results are reported in tables 2.18 and 2.19.

Moreover, the exclusion restriction holds for the chosen IVs, that is, the IVs should not be included in the price equation, (1); we explain as follows by IV variable. First, the demographic IVs (household size and dependency ratio) are not included in the price equation because there is no quality differentiation per product in the survey data for the set of products we chose to examine using the HBS data.[18] We exclude demographic variables from the price equation because they

[18] The main quality differentiator with maize flour is whether it is white (*sembe*) or brown (*dona*). The HBS specifically asked only about white maize flour. Likewise, the HBS differentiates between sunflower oil and palm oil (the latter of which is much more widely consumed in Tanzania, and which we use in our analysis), so we know there are not hidden quality differences in maize flour and cooking oil. Additionally, items like tomatoes, onions, leafy

would have had an effect on price only through affecting demand for differentiated quality, but the latter is ruled out by the nature of the products analyzed. Second, the employment composition (FTE) variables should affect the price paid (after controlling for income) only via the size of the transacted unit, and in no other way would employment directly affect price. Third, as argued above, possessing a refrigerator allows households to buy perishable goods in bulk and hence is likely to affect the quantity purchased, which we argue is the only way in which owning a refrigerator should affect the price paid (again, after controlling for income).

2.4.2 Expensiveness Index Equation

We estimated an equation to test if more is paid for a food basket comprised of 16 items as a correlate of income proxied as above by total expenditure per AE, and as a correlate of various measures of poverty versus non-poverty of households. The dependent variable is an "expensiveness index," calculated as the ratio of the cost of the household's food basket of 16 items to the average cost of the same food basket, where average costs were defined using prices of the 16 products per zone (as defined above) and month. This is very similar to Beatty's (2010) methodology except here zones are added to the equation; Beatty only calculates averages by month. We add zones to account for spatial heterogeneity in prices.

The expensiveness index is calculated as follows. Because the average cost of the 16-item food basket is calculated for each month and zone combination, only the most common food products that are consumed around the country were included in the calculation; this is because it is necessary to have product prices in each zone and month and those were not observed for less

green vegetables, and beef are mostly cooked into a stew, so aesthetic differences are irrelevant. The main exception is with rice, which is quality-differentiated (you can buy broken rice, fragrant rice, long-grain rice, regional varieties, etc.) – this could explain why we find no evidence for bulk discounts for rice, either in the descriptives or in the econometrics. Quantity of rice purchased could be correlated with quality (e.g., the higher-quality rice may only be sold in a 5kg package).

common foods in the data. In order of decreasing frequency in the average household's consumption basket, these include fresh tomatoes, onions, maize flour, cooking oil, rice, tea, brown sugar, spinach and lettuce (which are lumped together into one category), other leafy vegetables, dried fish, fried buns, beef, coconut, salt, fresh cow milk, and plantains. Together, these products account for 60% of total food purchases summed across all households. Table 2.16 in Appendix B shows that most households purchase the 16 products.[19]

First, we compute total purchase expenditure χ on the 16 food products for household *j*, located in zone *g* for the month recalled by the household. This is the sum, over food products and transactions, of the product of the price per kg *v* and quantity purchased *q* of product *i* in transaction *t* (which takes place in month *m*):

$$\chi_j = \sum_{i\epsilon I, t\epsilon m} v_{it}^j q_{it}^j = \sum_{i\epsilon I, t\epsilon m} \chi_{it}^j \qquad (2)$$

where iϵI=(tomatoes, onions,…, plantains) and mϵM=(January, February,…, December).

Next, for each food product *i*, we constructed the average price per kg over all (N) transactions *t* for all households in zone *g*, in month *m*:

$$\overline{v_{\iota,m,g}} = \frac{1}{N} \sum_{j\epsilon J, t\epsilon m, g\epsilon G} \chi_{it}^j \qquad (3)$$

Next, we constructed the average cost of purchasing household *j*'s food basket:

$$\widetilde{\chi_J} = \sum_{i\epsilon I} \overline{v_{\iota,m,g}}\, q_{it}^j \qquad (4)$$

Finally, the expensiveness index is the ratio of the cost of the food basket paid by the household, χ_j, relative to the cost of the basket at the average prices, $\widetilde{\chi_J}$:

$$I^j = \frac{\chi_j}{\widetilde{\chi_J}} \qquad (5)$$

[19] The main exception is fresh cow milk. Only 25% of households in most zones reported consuming fresh milk (though those who do consume it tend to purchase it).

Because the average prices are calculated by month and by zone, for ease of interpretation, we ran these regressions separately by each of the six zones (i.e., we ran them six times for each of the three key explanatory variables (discussed below), resulting in 18 sets of results) using OLS. OLS is sufficient because the dependent variable is not limited.

The estimated equation is hence

$$\ln(I^j) = \gamma + \eta key_j + \pi share_i + \boldsymbol{Z_j^s}\boldsymbol{\varphi} + e_{it} \qquad (6)$$

where key_j is one of three key righthand side variables, of each of which represents a different measure of poverty. That is, equation (6) is run three times, once for each measure of household poverty. The first is the natural log of total expenditure per adult equivalent, and its square. The second is a 0/1 poverty indicator. We run two versions of the regressions with this indicator: one version in which one "basic needs" poverty line was used for all households in the sample (about $24 USD per adult equivalent for a month), and one version where two poverty lines were used: $24 for rural households, and about $32 (i.e. a 30% increase from $24) for urban households, in a crude attempt to adjust for higher urban cost of living.[20] The $24 poverty line is taken from 2011/2012 HBS Key Findings document and accounts for both food and non-food goods such as clothing needed for long-term physical well-being (United Republic of Tanzania 2013, p. 3).

The third set of regressions uses poverty gap and poverty severity. The poverty gap for household i with total expenditure per adult equivalent y_i is defined as $\frac{z-y_i}{z}$, where z is the basic needs poverty line defined above. Poverty severity is the square of the poverty gap and puts more weight on households that are further below the poverty line (i.e., the poorest households). These

[20] We use individual wages as a rough proxy for the cost of living. OECD (2020) reports that, in a survey of six African countries, wages were about 30% higher in towns and cities compared to rural areas. We acknowledge that this is a rough estimate because some products and services may be cheaper in the rural areas (e.g. agricultural products produced locally), but some may be more expensive (e.g. imported fuel or food products not manufactured locally).

regressions were run twice, once with one poverty line for all households, and once with two poverty lines.

Figure 2.4 displays descriptive statistics on average monthly total expenditure per adult equivalent and the share of households that are poor in each. (This information plus poverty gap and poverty severity, all by terciles within each zone, can be found in Table 2.17).

The only other explanatory variables we include in this equation are the purchased share variable defined above and used in the bulk discount equations, and the same set of supply shifters included in equation (1), that is, region and month dummies and region by month interactions (but excluding retail type, because equation (6) is estimated at the household level). For the same reason we did not use household demographics as covariates in the bulk discount equations, we excluded demographic variables in the expensiveness index regressions.

2.5 Econometric Results

2.5.1 Bulk Discount

To keep the discussion manageable, we present and discuss 2SLS results for the five staple foods (rice, maize flour, cooking oil, tomatoes, and onions) and staple proteins (dried fish, beef, and fresh cow milk). (The results of OLS regressions are presented in Tables 2.18 and 2.19 for rural and urban households, respectively.)

2.5.1.1 First stage regressions

The first stage regression results are presented in tables 2.4 (rural) and 2.5 (urban). Recall in the first stage we regress the natural log of quantity purchased on the set of instrumental variables (adult equivalents, dependency ratio, the set of FTE share variables, and in urban regressions only, refrigerator ownership), the natural log of total expenditure and its square, the share of purchases in the overall food budget, a set of retail dummies, and geographic zone dummies.

There is generally strong evidence that the size of the household as measured by the number of adult equivalents is positively associated with quantity purchased, holding all else equal. In contrast, the only strong evidence that dependency ratio is significantly (and positively) correlated with quantity is in the grain (rice and maize flour) regressions for both rural and urban households, and in the rural beef and milk regressions. This makes sense, as it may be easier to cook a big pot of gruel (e.g., ugali, which is the staple maize flour porridge) or rice, accompanied with a staple protein like beef or milk, when there are more children and elderly in the home.

In regard to the FTE share variables (where the share of weekly wage FTEs is the base category), we find that the hourly/daily and monthly shares have positive and significant correlations with quantity purchased in the rural tomato regressions (Table 2.4), and the monthly share is positively correlated with quantity in the urban onion regression (Table 2.5). Own-farm work is positively associated with the quantity of rice and maize flour in the urban areas and beef in both urban and rural. This suggests that urban households with a greater share of wages that come more irregularly tend to purchase larger bags of grains, perhaps when they are paid, and this is feasible because rice and maize flour are not perishable.

For urban households, the correlation of refrigerator ownership and quantity purchased is generally as expected (i.e., positive and significant, holding all else constant) in the oil, tomatoes, onions, and milk regressions. We posit that it is positive even for non-perishable oil because households can cook larger meals and save the leftovers in the fridge.

The average marginal effect of total expenditure is positive and significant in most regressions, both rural and urban – i.e., the richer a household is, the more likely it is to purchase greater quantities, holding the other variables constant. However, though the estimates are statistically significant, they are extremely small. For example, a 10% increase in total expenditure per adult equivalent is associated with a 1% increase in rice quantity purchased on average, *ceteris paribus*, in the rural regression, and a 1.5% average increase in the urban regression.

Compared to the wet market, households tend to buy smaller quantities of staple grains (rice and maize flour) and oil at the shops. This complements our finding that the staples rice, maize flour, and oil tend to be purchased mainly at the shops – it appears that households make smaller and more frequent purchases at the shops. Finally, turning to the geographic zone variables, the evidence suggests that the rural intermediate and hinterland zones tend to be positively correlated with quantity compared to the peri-urban zone (e.g., for rice, maize flour, oil, and onions), suggesting perhaps that more rural households are more likely to do larger stock-up trips. There is some evidence that households in bigger cities may buy more than those in the towns (for example, secondary city is positively correlated with quantity of rice and tomatoes purchased, and primary city is positively correlated with quantity of oil and dried fish purchased).

2.5.1.2 Second stage regressions

The results of the second-stage regressions are in Tables 2.6 and 2.4.7. There are two surprising key results. First, contrary to our original hypothesis, we find little evidence for bulk discounts, with some exceptions in the rural regressions (the two important staple foods maize flour and cooking oil, and milk) and in the urban regressions (oil and beef). The urban results even suggest a positive correlation between quantity and the average unit value in several regressions (tomatoes, onions, and dried fish, the latter of which also has a positive estimated coefficient in the rural results).

The second key result is that the average marginal effect of natural log of total expenditure is generally positive and significant in the three rural zone regressions, suggesting that the poor actually pay less, holding all else constant, though the magnitude of the coefficients is quite small in most cases. Hence, the econometric evidence, at least for the rural sample, supports the descriptive results in which we find that the unit value paid by the poorest tercile is often less than that paid by the richest tercile, but in most cases is within 10%. Specifically, a 10% increase in total expenditure per adult equivalent is associated with between a 0.1% and 0.9% increase in price paid. The poor do not appear to be paying more than the rich and are, practically speaking, paying about the same as the rich, holding constant other variables such as retail outlet and geographic zone.

Rural shops tend to have higher prices compared to wet markets for the staples of rice, maize flour, cooking oil, and onions, though urban shops sell maize flour at a lower price on average than the wet markets do. Street vendors also have a price advantage compared to markets on beef, a result which holds in both the rural and urban regressions, and on rice (in the rural areas) and maize flour (in the urban areas).

Finally, regarding the correlations between geographic zone and average price, there are mixed results in the rural staple grain results. Rice is more expensive in the intermediate zone than in the peri-urban zone, but maize flour and oil are cheaper. This could be because rice is only grown in certain regions and hence goes through secondary cities and towns before it reaches many rural areas. The results suggest that tomatoes are more expensive in the hinterland, which could be due to higher transaction costs. In the urban regressions, secondary cities are less expensive compared to towns for oil and onions but have more expensive beef and milk. Primary cities have cheaper staple grains, oil, tomatoes, and dried fish.

2.5.2 Expensiveness Index

In the discussion that follows, we will focus on the results that use two basic needs poverty lines – one for rural households, and one for urban households that is 30% higher (Tables 2.9 and 2.10). Of course, only the urban results change between this set of regressions and the set that only uses one poverty line, and even then, the results between the two are nearly identical. And, no matter what variable is used to measure poverty, the basic story is the same as what we have found in the descriptive statistics, and in the bulk discount results: the poor pay do not pay more than the rich, holding constant their geographic location and the month of interview.

2.5.2.1 Poverty Variables

The results are generally stronger in the rural geographic zone regressions (Tables 2.8, 2.9, and 2.10). Results using the continuous variable suggest that the average marginal effect of natural log of total expenditure per adult equivalent ranges from 0.039 in the peri-urban zone to 0.07 in the hinterland (Table 2.8). This means that a 10% increase in total expenditure increases the expensiveness index on average by between 0.3% and 0.7%. These are not practically meaningful magnitudes.

In the results that use a 0/1 poverty indicator (Table 2.9), the estimated coefficient on this dummy variable is positive and significant in all the rural geographic zones. This means that, *ceteris paribus*, a household that is deemed poor by our definition (recall, in the rural sample this is a household whose monthly total expenditure per adult equivalent is below about $24 USD, and in the urban sample, about $32 USD) pays *less* than average for their food basket than a household who is not identified as poor. (This coefficient estimate is not statistically significant in the urban regressions).

Finally, the regressions that use poverty gap and poverty severity variables suggest that the poorer a household is (i.e., the further below the poverty line they are), the less they pay compared to the average cost of their food basket (in other words, the smaller is their expensiveness index, on average) in the intermediate and peri-urban regressions (Table 2.10). Hence, the results are consistent and clear: statistically, the poor pay less than average for their food baskets than the rich. One possible explanation for this counterintuitive result is that food retailers in richer neighborhoods engage in monopsonistic behavior. In other words, such retailers could engage in price discrimination by raising their prices, with the knowledge that rich households can afford to pay more.

2.5.2.2 Share of Purchases in Total Food Budget

For rural households, the share of the household's food budget that comes from purchases (vs. own-production) has a positive and significant estimated coefficient – meaning that the more food a household purchases, the more they pay for their food basket relative to the average cost. This could be due to a loss in bargaining power for households who rely more on the market. (This coefficient is generally not significant in the urban regressions, probably because urban households, with few exceptions, buy most of their food).

2.6 Conclusion

In this paper, we revisit the question of whether the poor pay more. We use a data set from Tanzania that includes a detailed 28-day food diary and spatial data that allow us to segment rural households by distance from the nearest town, and urban households by the size of the city. This is a valuable contribution to a literature that has tended to categorize households by only urban/rural. We also test the correlation between measures of a household's liquidity constraints (using data on full time equivalents (FTEs) spent in work with different pay frequencies) and the quantity purchased of several food items. Our analysis yielded several insights.

First is the result from the title of the paper: the poor do not pay more. After controlling for quantity purchased, spatiality, time of the survey, and several other demographic variables, we find that the poor do not pay more than the rich. The regressions suggest there is even a positive relationship between income and amount paid, but the magnitudes are very small, leading us to conclude that the poor and the rich pay essentially the same.

Second, and a possible explanation for the first result, is that the poor and the rich are buying roughly the same products, in the same quantity for each transaction, and the same number of times per month, and this is true across geographic zones. This challenges the very commonly held notion that the poor are forced to buy in very small quantities, while the rich can afford to stock up with a few weeks' worth of goods (at least, non-perishable goods like rice and flour).

Third, based on our urban first stage 2SLS results, a household's level of liquidity as measured by share of total work time spent in different types of employment is not significantly correlated with the amount they purchase in any given transaction. This is some evidence that, contrary to popular belief, liquidity constraints are not a driver of small purchases.

Fourth, bulk discounts do exist for the products we study in this paper, but only up to a certain threshold of unit size – beyond this threshold, there is very little additional discounting (the exception being with leafy green vegetables and animal-source proteins like beef and milk, where there continues to be substantial post-threshold discounting, especially for rural households).

In conclusion, our results present intriguing evidence that, at least for our sample of Tanzanian households and for the chosen food products, the poor do not pay more. Future research should replicate this analysis, where food expenditure and spatial data allow, to test the external validity of these results. That is, additional analysis should test whether our results are robust to changes in factors such as the national food environment and employment opportunities. If it turns out that the results are unique to Tanzania, qualitative research (e.g. with focus groups) could dig deeper into the purchasing behavior of poor and rich households to understand what drives their purchasing decisions.

APPENDICES

APPENDIX A: Tables and Figures

Table 2.1 Average household share of each product in total food expenditure

	Rural										Urban									
	Hinterland			Intermediate			Peri-Urban			Rural Overall	Towns			Secondary Cities			Primary Cities			Urban Overall
	T1	T2	T3	T1	T2	T3	T1	T2	T3		T1	T2	T3	T1	T2	T3	T1	T2	T3	
Rice	10	11	13	10	11	12	11	13	15	12	13	16	16	15	18	17	17	17	16	17
White Maize Flour	9	10	9	19	17	13	20	14	11	15	19	14	9	16	12	9	13	9	7	11
Cooking Oil	6	4	5	5	5	5	6	5	7	5	7	6	8	5	6	6	6	6	7	6
Tomatoes	3	2	3	3	3	3	4	3	4	3	4	4	4	4	4	4	4	4	4	4
Onions	1	1	2	1	1	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Spinach/Lettuce	2	1	1	2	1	1	2	1	1	1	1	1	1	2	2	2	2	2	2	2
Other Leafy Veg.	1	2	1	2	1	1	2	1	1	1	1	1	1	1	1	1	2	2	1	2
Dried Fish	3	3	2	5	4	3	5	3	3	4	3	2	2	4	3	2	2	2	2	2
Beef with Bones	5	8	9	7	7	8	9	8	9	8	6	7	9	8	10	11	8	10	11	10
Fresh Cow Milk	2	2	2	2	1	4	6	3	4	3	1	2	3	2	3	4	4	3	4	4

Table 2.2 Average number of purchases over 28 days (for household), average per-transaction quantity purchased (in kg unless otherwise noted), and average per-transaction price (USD per kg or per liter), by rural geographic zone and total expenditure per AE terciles (calculated within each zone)

	Hinterland				Intermediate				Peri-Urban				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Rural**
Number of households	190	191	184	565	1008	1013	993	3014	509	512	478	1499	5078
Rice													
Average num. purchases	3	5	8	6	3	5	7	5	4	6	8	6	5
Average quantity purchased	1.7	1.9	2.1	1.9	2.0	1.7	2.0	1.9	1.6	1.8	2.3	1.9	1.9
*Share (of total **quantity** purchased) purchased <0.25kg*	0.08%	0.00%	0.00%	0.01%	0.02%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%
Share purchased <=1kg	30%	32%	45%	38%	25%	32%	33%	32%	32%	35%	39%	36%	34%
Share purchased >=5kg	8%	12%	20%	15%	11%	9%	17%	13%	8%	16%	23%	18%	15%
*Share of **transactions** in quantity <=1kg*	52%	55%	73%	64%	51%	57%	60%	58%	56%	60%	69%	64%	61%
*Share of **transactions** in quantity <0.25kg*	1%	1%	2%	1%	2%	1%	1%	2%	1%	0%	2%	1%	1%
*Share of **transactions** in quantity >=5kg*	1%	2%	3%	3%	3%	2%	3%	3%	2%	3%	4%	3%	3%
Average price (USD)	1.17	1.21	1.18	1.18	1.16	1.20	1.19	1.18	1.19	1.18	1.22	1.20	1.19
Ratio of T1 price to T3 price	0.99				0.97				0.98				
Maize Flour													
Purchases	7	7	7	7	9	10	9	9	10	10	10	10	9
Quantity	3.2	3.8	4.5	3.8	4.1	3.9	4.2	4.1	3.5	3.2	3.2	3.3	3.7

Table 2.2, cont.

	Hinterland				Intermediate				Peri-Urban				Overall Rural
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	
*Share (of total **quantity** purchased) purchased <0.25kg*	0.07%	0.06%	0.00%	0.03%	0.02%	0.01%	0.01%	0.01%	0.02%	0.01 %	0.01%	0.01%	0.01%
Share purchased <=1kg	23%	24%	29%	26%	15%	19%	26%	21%	19%	26%	36%	27%	23%
Share purchased >=5kg	24%	43%	59%	47%	37%	36%	41%	38%	27%	31%	37%	32%	37%
*Share of **transactions** in quantity <=1kg*	51%	60%	80%	67%	44%	49%	64%	54%	47%	60%	77%	63%	58%
*Share of **transactions** in quantity <0.25kg*	5%	4%	8%	6%	3%	2%	3%	3%	3%	3%	8%	5%	4%
*Share of **transactions** in quantity >=5kg*	5%	7%	9%	8%	7%	7%	5%	6%	4%	5%	5%	5%	6%
Price	0.59	0.64	0.68	0.64	0.61	0.62	0.65	0.63	0.60	0.63	0.64	0.62	0.63
Ratio of T1 to T3 price	0.88				0.95				0.94				
Cooking oil (liters)													
Purchases	8	8	9	8	9	12	11	11	11	10	10	10	10
Quantity	0.3	0.4	0.9	0.5	0.2	0.2	0.5	0.3	0.2	0.5	0.9	0.5	0.5
*Share (of total **quantity** purchased) purchased <0.25*	57%	33%	26%	34%	61%	54%	34%	46%	57%	30%	19%	30%	40%
*Share of **transactions** in quantity <=0.25*	94%	90%	82%	88%	94%	93%	89%	92%	95%	89%	82%	89%	91%

Table 2.2, cont.

	Hinterland				Intermediate				Peri-Urban				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Rural**
Price	2.54	2.65	2.13	2.44	2.78	2.89	2.81	2.82	3.05	2.75	2.80	2.87	2.71
Ratio of T1 to T3 price	1.19				0.99				1.09				
Tomatoes													
Purchases	7	9	10	9	9	11	12	11	9	11	11	10	10
Quantity	0.3	0.4	0.4	0.4	0.3	0.3	0.4	0.4	0.3	0.4	0.5	0.4	0.4
Share (of total ***quantity*** *purchased) purchased <0.25*	19%	26%	20%	22%	27%	25%	18%	22%	27%	22%	14%	20%	21%
Price	0.74	0.96	0.92	0.88	0.76	0.83	0.79	0.79	0.74	0.79	0.91	0.81	0.83
Ratio of T1 to T3 price	0.81				0.97				0.81				
Onions													
Purchases	6	6	7	6	7	8	9	8	8	10	10	9	8
Quantity	0.2	0.4	0.5	0.4	0.1	0.2	0.2	0.2	0.2	0.2	0.3	0.2	0.3
Share purchased <0.25	43%	38%	26%	33%	61%	57%	47%	53%	60%	58%	50%	55%	51%
Price	1.33	1.32	1.20	1.28	1.36	1.49	1.51	1.45	1.68	1.55	1.52	1.59	1.44
Ratio of T1 to T3 price	1.11				0.9				1.11				
Spinach and Lettuce													
Purchases	3	3	4	3	3	3	4	4	4	4	4	4	4
Quantity	0.6	0.5	0.5	0.5	0.5	0.4	0.4	0.4	0.5	0.5	0.4	0.5	0.5
Share purchased <0.25	8%	10%	9%	9%	8%	10%	12%	10%	9%	11%	13%	11%	11%
Price	0.69	0.97	0.81	0.82	0.78	0.86	0.85	0.83	0.69	0.79	0.93	0.80	0.82
Ratio of T1 to T3 price	0.86				0.92				0.74				

Table 2.2, cont.

	Hinterland				Intermediate				Peri-Urban				Overall
	T1*	T2	T3	Overall	T1	T2	T3	Overall	T1	T2	T3	Overall	Rural
Other Leafy Veg.													
Purchases	2	4	4	3	3	3	4	3	4	4	3	4	3
Quantity	0.6	0.5	0.5	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.4	0.5	0.5
Share purchased <0.25	9%	7%	6%	7%	8%	8%	10%	9%	11%	10%	11%	11%	9%
Price	0.91	1.06	0.92	0.96	0.81	0.90	0.92	0.88	0.78	0.81	0.95	0.85	0.90
Ratio of T1 to T3 price	0.99				0.88				0.82				
Dried Fish													
Purchases	4	4	3	4	5	6	5	5	5	4	4	4	4
Quantity	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.3	0.3	0.3	0.3
Share purchased <0.25	33%	28%	26%	29%	28%	26%	20%	25%	50%	42%	35%	42%	28%
Price	2.64	3.07	3.20	2.97	2.54	2.68	2.60	2.61	3.09	2.96	2.81	2.95	2.84
Ratio of T1 to T3 price	0.83				0.98				1.10				
Beef													
Purchases	2	2	4	3	2	2	3	3	2	3	4	3	3
Quantity	1.0	1.1	0.9	1.0	0.9	0.9	0.8	0.9	0.8	0.8	0.7	0.8	0.9
Share purchased <0.25	0.00%	0.00%	0.12%	0.08%	0.05%	0.08%	0.03%	0.05%	0.47%	0.58 %	0.11%	0.33%	0.15%
Price	2.22	2.78	3.06	2.68	2.63	2.72	2.75	2.70	2.78	3.03	3.31	3.04	2.81
Ratio of T1 to T3 price	0.73				0.96				0.84				
Fresh Cow Milk (liters)													
Purchases	2	3	4	3	3	3	6	4	7	6	6	6	4
Quantity	0.9	1.4	1.1	1.1	1.0	1.2	1.1	1.1	1.0	1.0	0.9	1.0	1.1
Share purchased <0.25	0.00%	0.00%	0.00%	0.00%	1.43%	0.20%	0.09%	0.26%	0.00%	0.00 %	0.04%	0.02%	0.15%

Table 2.2, cont.

	Hinterland				Intermediate				Peri-Urban				Overall Rural
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	
Price	0.51	0.52	0.62	0.55	0.50	0.52	0.58	0.53	0.68	0.67	0.73	0.69	0.59
Ratio of T1 to T3 price	0.83				0.88				0.94				

Notes: * terciles are calculated within each zone.

Table 2.3 Average number of purchases over 28 days (for household), average per-transaction quantity purchased (in kg unless otherwise noted), and average per-transaction price (USD per kg or per liter), by urban geographic zone and total expenditure per AE terciles (calculated within each zone)

	Towns				Secondary Cities				Primary Cities				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Number of households	103	104	99	306	337	330	302	969	914	854	718	2486	3761
Rice													
Average num. purchases	6	10	9	8	7	11	11	10	10	12	9	10	9
Average quantity purchased	1.6	1.8	3.3	2.3	1.4	1.9	2.3	1.9	1.5	1.9	3.2	2.2	1.6
Share (of total ***quantity*** *purchased) purchased <0.25kg*	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%	0.01%	0.02%	0.01%	0.00%	0.01%	0.01%	0.01%
Share purchased <=1kg	44%	38%	29%	35%	48%	41%	43%	43%	42%	48%	44%	45%	44%
Share purchased >=5kg	14%	15%	35%	23%	6%	14%	24%	16%	7%	17%	33%	18%	18%
Share of ***transactions*** *in quantity <=1kg*	70%	62%	65%	65%	70%	67%	72%	69%	63%	74%	100%	75%	71%
Share of ***transactions*** *in quantity <0.25kg*	2%	1%	1%	1%	2%	2%	5%	3%	1%	1%	3%	2%	2%
Share of ***transactions*** *in quantity >=5kg*	1%	2%	4%	3%	1%	2%	3%	2%	1%	2%	0%	1%	2%
Average price (USD)	1.23	1.23	1.24	1.23	1.17	1.17	1.18	1.17	1.24	1.26	1.28	1.26	1.22
Ratio of T1 price to T3 price	0.99				0.99				0.97				
Maize Flour													
Purchases	13	12	8	11	14	14	12	13	15	13	9	12	12
Quantity	3.0	3.9	3.6	3.5	1.9	1.9	1.7	1.8	1.7	1.9	2.4	2.0	1.8

Table 2.3, cont.

	Towns				Secondary Cities				Primary Cities				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Share (of total ***quantity*** *purchased) purchased <0.25kg*	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.04%	0.06%	0.03%	0.03%
Share purchased <=1kg	29%	29%	29%	29%	43%	47%	58%	48%	49%	58%	58%	54%	50%
Share purchased >=5kg	34%	47%	56%	45%	12%	14%	21%	15%	12%	21%	33%	19%	20%
Share of ***transactions*** *in quantity <=1kg*	61%	68%	78%	68%	69%	75%	87%	76%	75%	86%	100%	84%	80%
Share of ***transactions*** *in quantity <0.25kg*	6%	6%	4%	6%	3%	5%	12%	6%	5%	8%	15%	8%	7%
Share of ***transactions*** *in quantity >=5kg*	3%	6%	10%	6%	1%	2%	2%	2%	1%	2%	0%	1%	2%
Price	0.62	0.65	0.65	0.64	0.61	0.62	0.62	0.62	0.63	0.64	0.64	0.64	0.63
Ratio of T1 to T3 price	0.94				0.98				0.98				
Cooking oil (liters)													
Purchases	18	15	12	15	15	15	12	14	14	14	10	13	14
Quantity	0.2	0.4	1.0	0.6	0.2	0.7	1.2	0.7	0.3	0.5	1.2	0.7	0.5
Share (of total ***quantity*** *purchased) purchased <0.25*	53%	41%	30%	42%	65%	46%	29%	45%	46%	38%	20%	35%	37%

Table 2.3, cont.

	Towns				Secondary Cities				Primary Cities				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
*Share of **transactions** in quantity <0.25*	92%	90%	84%	90%	95%	91%	88%	92%	93%	91%	100%	94%	91%
Price	2.29	2.17	2.44	2.30	2.19	2.23	2.21	2.21	2.24	2.25	2.28	2.26	2.26
Ratio of T1 to T3 price	0.94				0.99				0.98				
Tomatoes													
Purchases	14	13	14	14	13	15	14	14	13	14	12	13	14
Quantity	0.3	0.4	0.6	0.4	0.3	0.4	0.5	0.4	0.3	0.4	0.6	0.5	0.3
Share purchased <0.25	23%	14%	4%	12%	19%	14%	10%	14%	19%	14%	10%	15%	14%
Price	0.73	0.70	0.69	0.71	0.81	0.76	0.77	0.78	0.83	0.84	0.87	0.85	0.78
Ratio of T1 to T3 price	1.06				1.05				0.95				
Onions													
Purchases	10	8	10	10	10	12	10	11	15	15	13	14	11
Quantity	0.2	0.3	0.4	0.3	0.1	0.2	0.3	0.2	0.2	0.2	0.3	0.2	0.2
Share purchased <0.25	62%	46%	37%	47%	62%	58%	52%	57%	69%	68%	49%	62%	60%
Price	1.04	1.17	1.23	1.14	1.36	1.17	1.29	1.28	1.23	1.27	1.39	1.30	1.24
Ratio of T1 to T3 price	0.85				1.05				0.89				
Spinach and Lettuce													
Purchases	4	4	6	5	5	6	6	6	4	5	5	5	5
Quantity	0.4	0.4	0.4	0.4	0.4	0.5	0.4	0.5	0.5	0.4	0.4	0.4	0.3
Share purchased <0.25	13%	7%	10%	10%	8%	7%	7%	7%	10%	9%	10%	10%	9%
Price	0.74	0.69	0.77	0.74	0.86	0.71	0.76	0.78	0.91	0.94	0.99	0.95	0.82

Table 2.3, cont.

	Towns				Secondary Cities				Primary Cities				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Ratio of T1 to T3 price	0.96				1.14				0.92				
Other Leafy Veg.													
Purchases	3	3	4	3	3	4	4	3	4	4	4	4	4
Quantity	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.4	0.5	0.3
Share purchased <0.25	10%	6%	6%	7%	7%	7%	9%	8%	9%	7%	12%	9%	8%
Price	0.73	0.76	0.94	0.81	0.88	0.81	0.84	0.84	0.92	0.96	1.00	0.96	0.87
Ratio of T1 to T3 price	0.78				1.05				0.93				
Dried Fish													
Purchases	4	3	3	4	5	5	4	5	3	3	3	3	4
Quantity	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2
Share purchased <0.25	46%	26%	26%	33%	26%	29%	30%	28%	24%	23%	23%	23%	26%
Price	2.60	2.57	2.61	2.59	2.34	2.41	2.33	2.36	2.44	2.49	2.47	2.47	2.47
Ratio of T1 to T3 price	1				1				0.99				
Beef													
Purchases	3	3	5	3	4	5	5	4	4	4	5	4	4
Quantity	0.6	0.8	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.7	0.8	0.7	0.5
Share purchased <0.25	0.29%	0.12%	0.06%	0.12%	0.22%	0.13%	0.20%	0.18%	0.07%	0.10%	0.10%	0.09%	0.12%
Price	3.12	3.16	3.05	3.11	2.98	2.99	2.97	2.98	3.44	3.48	3.49	3.47	3.19
Ratio of T1 to T3 price	1.02				1				0.98				
Fresh Cow Milk (liters)													
Purchases	3	3	5	4	4	5	5	5	7	5	5	6	5
Quantity	1.3	0.9	1.0	1.1	0.8	0.8	1.0	0.9	0.8	0.9	1.4	1.1	0.8

Table 2.3, cont.

	Towns				Secondary Cities				Primary Cities				Overall
	T1*	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Share purchased <0.25	0.00%	0.24%	0.00%	0.08%	0.18%	0.00%	0.09%	0.08%	0.00%	0.04%	0.04%	0.03%	0.04%
Price	0.64	0.68	0.69	0.67	0.60	0.65	0.68	0.64	0.74	0.75	0.77	0.75	0.69
Ratio of T1 to T3 price	0.93				0.89				0.96				

Notes: * terciles are calculated within each zone.

Figure 2.1 Average per-transaction price (shillings per kg or per liter) by size (transaction quantity) category ("<0.25 category is the numeraire, not shown on the graph), rural households

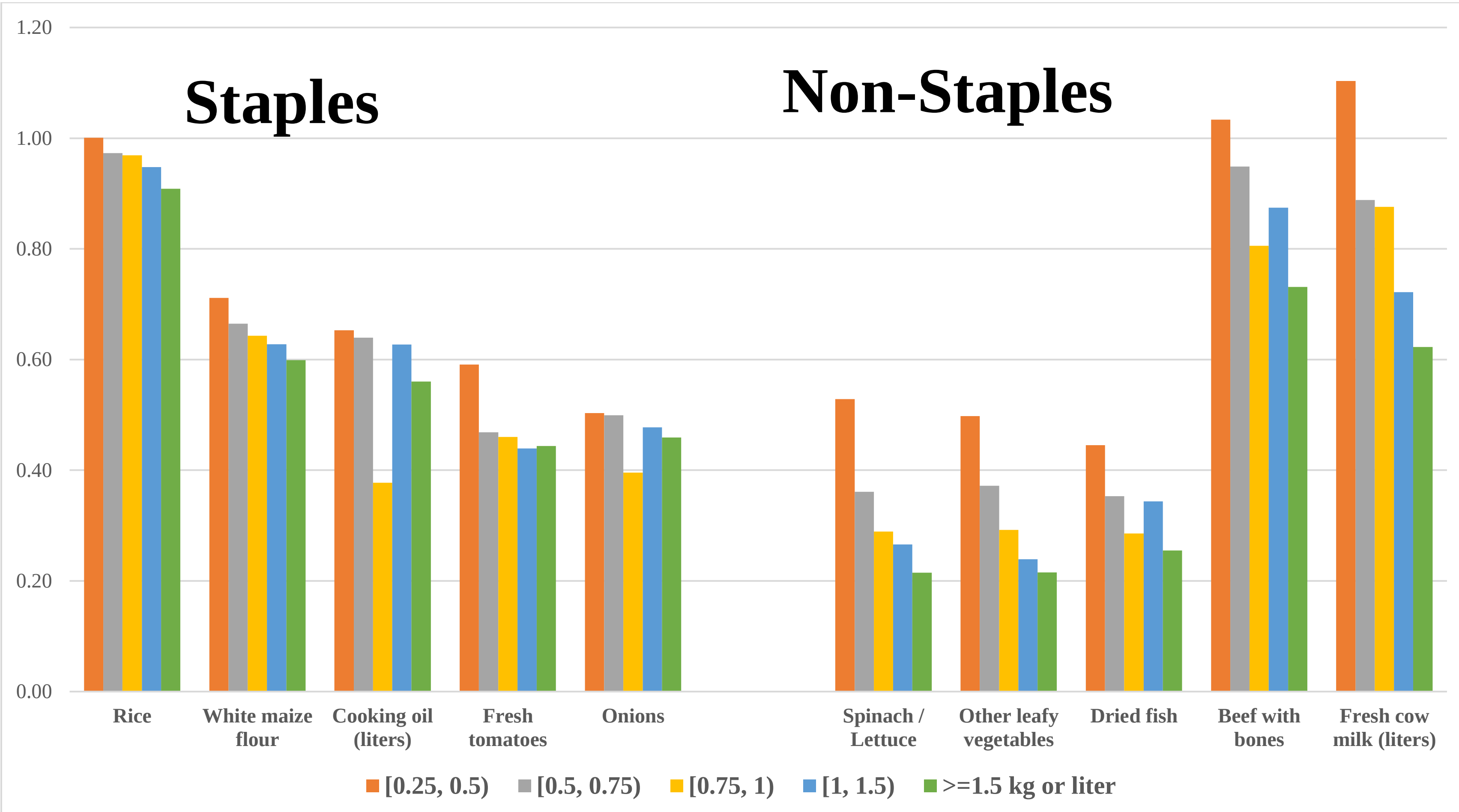


Figure 2.2 Average per-transaction price (shillings per kg or per liter) by size (transaction quantity) category ("<0.25" category is the numeraire, not shown on the graph), urban households

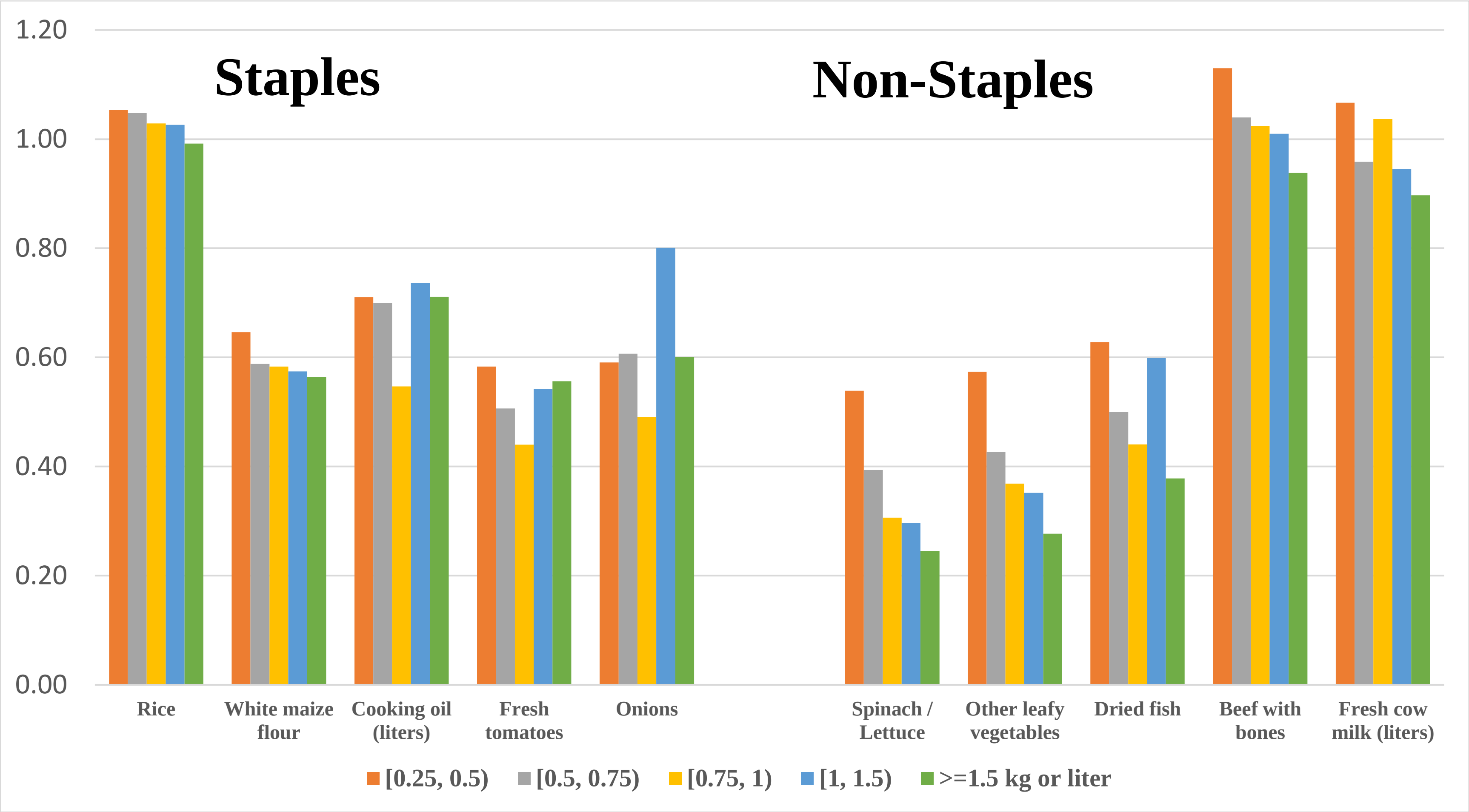


Figure 2.3 Average household shares of each FTE category over geographic zones

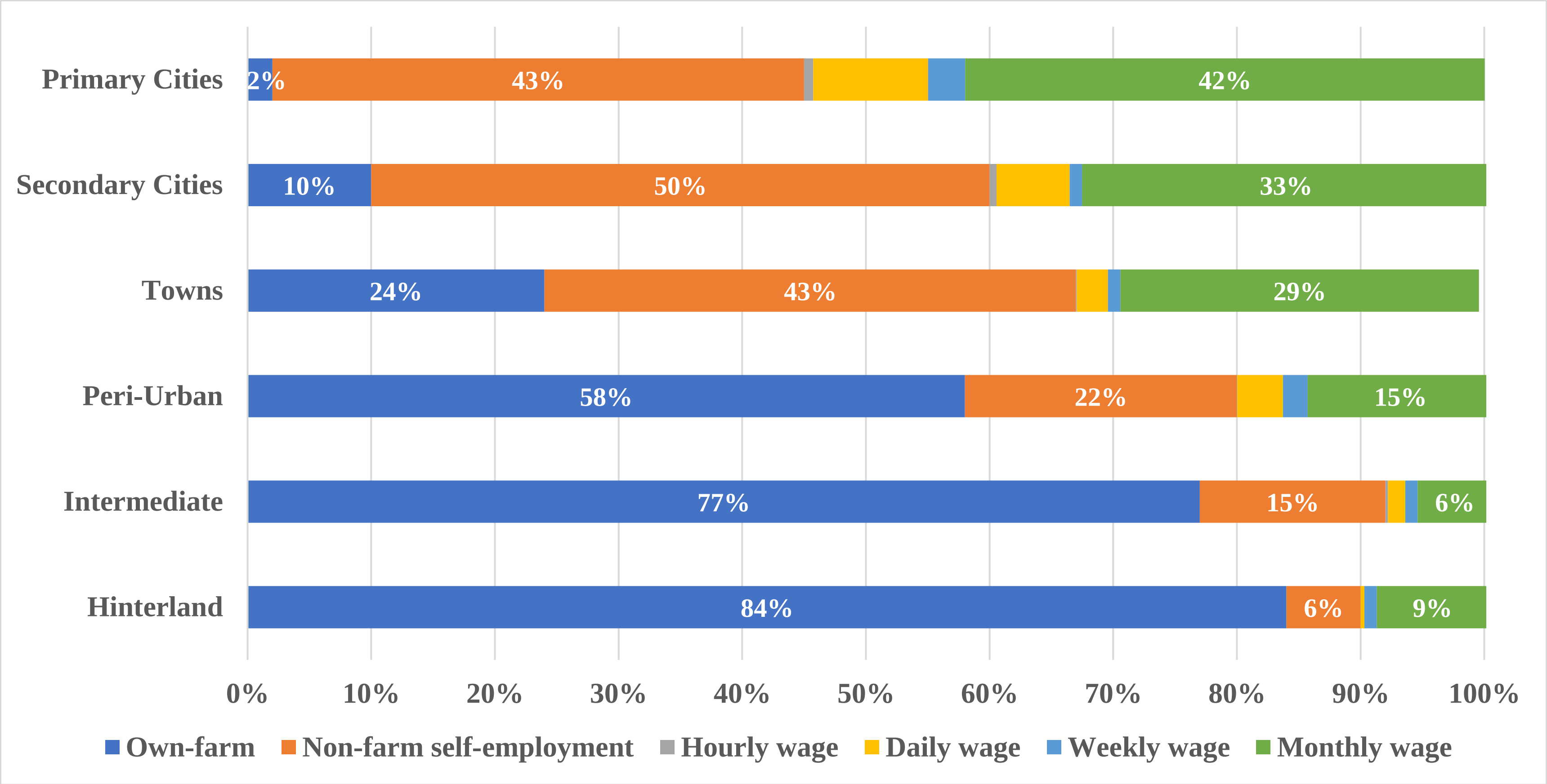


Figure 2.4 Average total expenditure per adult equivalent and poverty rate by geographic zone

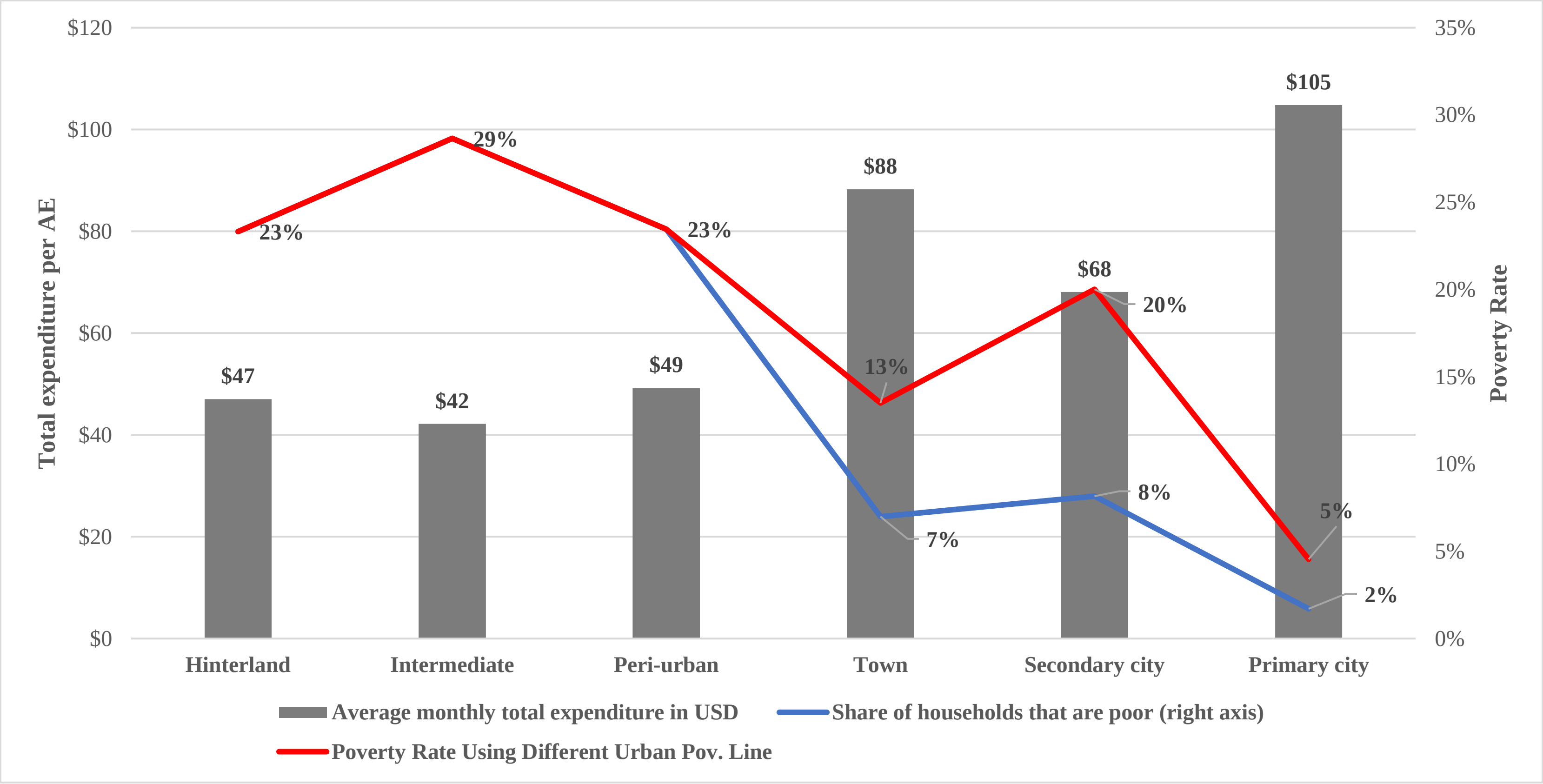


Table 2.4 Bulk discount first stage regression results: rural households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Instrumental Variables								
Adult equivalents	0.107***	0.126***	0.080***	0.046***	0.045***	0.056***	0.068***	0.078***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.013)
Dependency ratio	0.076***	0.054***	0.009	-0.015	-0.001	0.009	0.031*	0.075*
	(0.013)	(0.020)	(0.019)	(0.015)	(0.019)	(0.018)	(0.017)	(0.044)
Share of total FTEs where paid hourly or daily (includes self-employment)	-0.065	-0.278**	-0.308	0.193**	0.229	0.122	0.310*	0.028
	(0.095)	(0.119)	(0.256)	(0.090)	(0.168)	(0.140)	(0.164)	(0.247)
Share of salary work (monthly wage) in total FTEs	0.018	-0.163	0.200	0.351***	0.386**	0.066	0.457***	-0.078
	(0.099)	(0.130)	(0.265)	(0.096)	(0.175)	(0.155)	(0.166)	(0.253)
Share of own-farm in total FTEs	0.152	-0.015	-0.132	0.126	0.228	0.107	0.533***	0.090
	(0.095)	(0.117)	(0.254)	(0.089)	(0.169)	(0.134)	(0.162)	(0.241)
Other Exogenous Variables								
Natural log of total expenditure per AEQ	0.093	3.149***	-0.319	-0.396	-0.550	-0.398	0.795	1.632
	(0.344)	(0.558)	(0.473)	(0.351)	(0.516)	(0.578)	(0.537)	(1.085)
Natural log of total expenditure per AEQ, squared	0.000	-0.138***	0.030	0.023	0.030	0.023	-0.031	-0.062
	(0.015)	(0.025)	(0.021)	(0.016)	(0.023)	(0.026)	(0.024)	(0.047)
Average marginal effect of natural log of total expenditure per AEQ	0.102***	0.108***	0.328***	0.108***	0.111***	0.099***	0.100***	0.251***
	(0.019)	(0.030)	(0.028)	(0.019)	(0.029)	(0.031)	(0.027)	(0.060)
Household purchased share	-0.006	-0.019	0.203***	0.052	-0.119*	0.271***	-0.085	-0.089
	(0.052)	(0.070)	(0.058)	(0.051)	(0.067)	(0.065)	(0.067)	(0.147)
Shop (incl. supermarket)	-0.039**	-0.149***	-0.248***	-0.049**	-0.419***	-0.223***	0.020	-0.157**
	(0.020)	(0.032)	(0.032)	(0.024)	(0.033)	(0.032)	(0.032)	(0.066)
Street Vendor	0.061	0.009	0.082	-0.024	-0.088	0.161*	-0.041	-0.079
	(0.110)	(0.141)	(0.135)	(0.046)	(0.106)	(0.083)	(0.109)	(0.074)
Intermediate	0.042*	0.175***	0.280***	-0.022	0.015	0.107***	-0.015	-0.035

Table 2.4, cont.

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
	(0.023)	(0.042)	(0.032)	(0.025)	(0.038)	(0.035)	(0.037)	(0.080)
Hinterland	0.041	-0.084	0.500***	-0.022	0.391***	0.112*	0.130**	0.126
	(0.039)	(0.058)	(0.060)	(0.045)	(0.062)	(0.058)	(0.052)	(0.101)
Constant	-1.438	-17.872***	-2.469	-0.534	-0.560	-0.206	-6.284**	-11.116*
	(1.960)	(3.157)	(2.647)	(1.948)	(2.877)	(3.195)	(3.019)	(6.170)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Test for endogeneity (*p*-value)	0.028	0.019	0.000	0.000	0.000	0.000	0.992	1.00
Test for overidentification (*p*-value)	0.430	0.804	0.001	0.000	0.477	0.660	1.000	0.000
Montiel Olea and Pflueger Effective F statistic	146.456	121.849	57.607	30.277	16.450	20.381	47.080	8.391±
N	19,151	24,837	35,902	43,618	31,559	17,480	4,352	3,543
R^2	0.269	0.302	0.201	0.183	0.219	0.263	0.295	0.471

Notes: Dependent variable: natural log of quantity purchased in transaction (kg or liters). Clustered (at household level) standard errors in parentheses. The null hypothesis for the endogeneity test is that the suspected endogenous variables can be treated as exogenous. The null hypothesis for the overidentification test indicates that the instruments are valid. *** p<0.01, ** p<0.05, * p<0.10 ± Does not exceed τ=10% critical level.

Table 2.5 Bulk discount first stage regression results: urban households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Instrumental Variables								
=1 if household has a refrigerator or freezer	-0.005	-0.050**	0.104***	0.024	0.030	0.099**	0.020	0.156***
	(0.017)	(0.022)	(0.027)	(0.021)	(0.029)	(0.041)	(0.020)	(0.048)
Adult equivalents	0.143***	0.142***	0.113***	0.058***	0.070***	0.068***	0.108***	0.050***
	(0.005)	(0.006)	(0.006)	(0.005)	(0.007)	(0.010)	(0.006)	(0.011)
Dependency ratio	0.041***	0.034**	0.006	0.005	-0.028	-0.003	0.016	0.001
	(0.011)	(0.014)	(0.017)	(0.015)	(0.019)	(0.024)	(0.015)	(0.045)
Share of total FTEs where paid hourly or daily (includes self-employment)	-0.027	0.013	-0.104	-0.038	-0.048	-0.130	0.004	-0.104
	(0.049)	(0.055)	(0.079)	(0.056)	(0.077)	(0.109)	(0.055)	(0.167)
Share of salary work (monthly wage) in total FTEs	0.006	0.034	-0.056	-0.015	-0.019	-0.036	0.064	-0.180
	(0.050)	(0.056)	(0.081)	(0.057)	(0.079)	(0.110)	(0.055)	(0.161)
Share of own-farm in total FTEs	0.147**	0.131*	-0.001	0.004	-0.140	0.017	0.180**	0.142
	(0.063)	(0.076)	(0.091)	(0.069)	(0.093)	(0.125)	(0.073)	(0.198)
Other Exogenous Variables								
Natural log of total expenditure per AEQ	0.866***	0.429	1.588***	1.309***	0.685*	1.159*	1.222***	0.854
	(0.263)	(0.313)	(0.493)	(0.297)	(0.394)	(0.698)	(0.286)	(0.614)
Natural log of total expenditure per AEQ, squared	-0.031***	-0.015	-0.050**	-0.047***	-0.018	-0.045	-0.042***	-0.028
	(0.011)	(0.013)	(0.021)	(0.013)	(0.017)	(0.030)	(0.012)	(0.025)
Average marginal effect of natural log of total expenditure per AEQ	0.151***	0.078***	0.445***	0.229***	0.257***	0.141***	0.227***	0.184***
	(0.017)	(0.020)	(0.026)	(0.019)	(0.023)	(0.036)	(0.022)	(0.044)
Household purchased share	0.204***	0.306***	0.427***	0.121**	0.047	0.495***	0.259***	0.196
	(0.042)	(0.051)	(0.061)	(0.051)	(0.062)	(0.099)	(0.054)	(0.120)
Shop (incl. supermarket)	-0.100***	-0.067**	-0.062**	0.042	-0.008	-0.152***	0.028	-0.051
	(0.019)	(0.029)	(0.028)	(0.029)	(0.031)	(0.036)	(0.021)	(0.043)
Street Vendor	0.041	0.092	0.376	0.002	0.265	-0.031	-0.062	-0.025
	(0.292)	(0.201)	(0.246)	(0.075)	(0.204)	(0.092)	(0.126)	(0.062)
Secondary City	0.131**	-0.012	0.038	0.159***	0.076	0.099	0.046	0.179

Table 2.5, cont.

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
	(0.051)	(0.087)	(0.064)	(0.051)	(0.073)	(0.079)	(0.059)	(0.121)
Primary City	-0.110**	-0.118	0.143*	-0.003	-0.318***	0.298**	0.110	-0.098
	(0.049)	(0.075)	(0.081)	(0.051)	(0.084)	(0.117)	(0.099)	(0.100)
Constant	-6.643***	-3.865**	-14.490***	-10.248***	-8.357***	-9.795**	-9.522***	-7.103*
	(1.580)	(1.870)	(2.845)	(1.771)	(2.312)	(4.007)	(1.748)	(3.787)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Test for endogeneity (*p*-value)	0.016	0.001	0.000	0.000	0.000	1.000	0.000	1.000
Test for overidentification (*p*-value)	0.036	1.000	0.000	0.126	0.606	1.000	0.000	0.000
Montiel Olea and Pflueger Effective F statistic	281.423	186.187	88.235	36.098	31.890	28.840	91.714	6.032±
N	33,606	40,717	41,720	44,073	41,190	8,349	8,587	3,951
R^2	0.298	0.292	0.227	0.278	0.199	0.238	0.302	0.405

Notes: Dependent variable: natural log of quantity purchased in transaction (kg or liters). Clustered (at household level) standard errors in parentheses. The null hypothesis for the endogeneity test is that the suspected endogenous variables can be treated as exogenous. The null hypothesis for the overidentification test indicates that the instruments are valid. *** p<0.01, ** p<0.05, * p<0.10 ± Does not exceed τ=10% critical level.

Table 2.6 Bulk discount second stage regression results: rural households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Natural log of quantity	-0.008	-0.033*	-0.110**	-0.101	-0.081	0.236**	0.001	-0.138*
	(0.010)	(0.017)	(0.047)	(0.073)	(0.094)	(0.118)	(0.038)	(0.071)
Natural log of total expenditure per AEQ	-0.097	0.482***	0.422	0.514*	0.704*	1.229**	0.261	-0.060
	(0.095)	(0.187)	(0.284)	(0.266)	(0.387)	(0.584)	(0.217)	(0.463)
Natural log of total expenditure per AEQ, squared	0.005	-0.019**	-0.015	-0.019	-0.028	-0.054**	-0.009	0.005
	(0.004)	(0.008)	(0.013)	(0.012)	(0.017)	(0.027)	(0.010)	(0.020)
Average marginal effect of natural log of total expenditure per AEQ	0.017***	0.063***	0.093***	0.094***	0.085***	0.063**	0.050***	0.044**
	(0.004)	(0.010)	(0.019)	(0.015)	(0.020)	(0.028)	(0.012)	(0.021)
Household purchased share	-0.008	0.045*	-0.126***	0.088**	0.053	-0.299***	0.183***	0.063
	(0.012)	(0.024)	(0.037)	(0.038)	(0.047)	(0.069)	(0.036)	(0.047)
Shop (incl. supermarket)	0.019***	0.058***	0.089***	-0.017	0.160***	0.129***	0.003	0.067***
	(0.006)	(0.012)	(0.023)	(0.019)	(0.048)	(0.042)	(0.017)	(0.024)
Street Vendor	-0.089*	-0.062	0.058	-0.012	0.108**	0.009	-0.257***	0.077**
	(0.050)	(0.038)	(0.059)	(0.032)	(0.052)	(0.099)	(0.084)	(0.034)
Intermediate	0.021***	-0.044***	-0.177***	0.030	0.030	-0.118***	0.019	-0.006
	(0.006)	(0.014)	(0.025)	(0.022)	(0.026)	(0.037)	(0.022)	(0.028)
Hinterland	0.005	0.013	-0.169***	0.149***	-0.041	-0.102*	-0.027	-0.036
	(0.010)	(0.021)	(0.046)	(0.037)	(0.053)	(0.061)	(0.031)	(0.035)
Constant	8.017***	3.783***	5.063***	4.094***	3.057	1.240	6.541***	6.770**
	(0.540)	(1.061)	(1.544)	(1.455)	(2.113)	(3.195)	(1.240)	(2.632)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19,151	24,837	35,902	43,618	31,559	17,480	4,352	3,543

Notes: Dependent variable: natural log of price per kg or liter. Clustered (at household level) standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 2.7 Bulk discount second stage regression results: urban households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Natural log of quantity	-0.003	0.004	-0.049*	0.268***	0.111*	0.150**	-0.038*	0.028
	(0.007)	(0.009)	(0.026)	(0.085)	(0.066)	(0.076)	(0.020)	(0.035)
Natural log of total expenditure per AEQ	0.066	0.209**	0.083	-0.859***	-0.466	-0.181	-0.177	-0.148
	(0.060)	(0.097)	(0.235)	(0.277)	(0.299)	(0.592)	(0.133)	(0.148)
Natural log of total expenditure per AEQ, squared	-0.002	-0.008*	-0.003	0.036***	0.021*	0.009	0.008	0.007
	(0.002)	(0.004)	(0.010)	(0.012)	(0.013)	(0.026)	(0.005)	(0.006)
Average marginal effect of natural log of total expenditure per AEQ	0.028***	0.030***	0.018	-0.026	0.023	0.023	0.016*	0.013
	(0.004)	(0.004)	(0.014)	(0.021)	(0.023)	(0.028)	(0.009)	(0.010)
Household purchased share	0.033***	0.006	0.013	0.114**	0.212***	-0.228**	0.027	0.055
	(0.012)	(0.014)	(0.037)	(0.057)	(0.056)	(0.104)	(0.027)	(0.036)
Shop (incl. supermarket)	0.004	-0.019***	0.050***	0.065**	0.006	0.039	-0.014	0.004
	(0.004)	(0.007)	(0.013)	(0.028)	(0.023)	(0.033)	(0.011)	(0.018)
Street Vendor	-0.068	-0.055**	-0.094	-0.053	-0.044	0.187***	-0.260**	0.021
	(0.046)	(0.027)	(0.102)	(0.064)	(0.134)	(0.070)	(0.110)	(0.022)
Secondary City	0.008	0.001	-0.104**	-0.044	-0.129**	0.026	0.137***	0.096***
	(0.010)	(0.026)	(0.041)	(0.056)	(0.057)	(0.066)	(0.043)	(0.036)
Primary City	-0.040***	-0.054***	-0.251***	-0.179***	0.082	-0.277***	-0.026	0.054
	(0.014)	(0.020)	(0.045)	(0.050)	(0.069)	(0.091)	(0.025)	(0.083)
Constant	6.977***	5.506***	7.586***	12.521***	10.294***	9.515***	9.300***	7.678***
	(0.360)	(0.567)	(1.385)	(1.691)	(1.802)	(3.393)	(0.813)	(0.880)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	33,606	40,717	41,720	44,073	41,190	8,349	8,587	3,951

Notes: Dependent variable: natural log of price per kg or liter. Clustered (at household level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table 2.8 Expensiveness index regression results, using the continuous variable natural log of total expenditure per adult equivalent

	Rural			Urban		
	Hinterland	**Intermediate**	**Peri-Urban**	**Towns**	**Secondary Cities**	**Primary Cities**
Natural log of total expenditure per AEQ in USD	0.617***	0.132	0.104	-0.042	0.057	-0.013
	(0.139)	(0.106)	(0.063)	(0.078)	(0.055)	(0.036)
Natural log of total exp. per AEQ, squared	-0.075***	-0.011	-0.009	0.006	-0.007	0.002
	(0.018)	(0.013)	(0.008)	(0.009)	(0.007)	(0.004)
Average marginal effect of natural log of total expenditure per AEQ	0.070***	0.057***	0.039***	0.005	0.003	0.007
	(0.023)	(0.014)	(0.011)	(0.013)	(0.009)	(0.004)
Household purchased share	0.212***	0.095***	0.071***	-0.009	-0.028	-0.008
	(0.078)	(0.025)	(0.026)	(0.046)	(0.041)	(0.013)
Constant	-1.475***	-0.292	-0.357**	-0.404**	-0.094	-0.095
	(0.288)	(0.204)	(0.151)	(0.191)	(0.132)	(0.099)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes
N	565	3,014	1,499	306	969	2,486
R^2	0.386	0.217	0.365	0.715	0.426	0.185

Notes: Dependent variable: natural log of the expensiveness index. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table 2.9 Expensiveness index regression results, using 0/1 poverty indicator with two poverty lines

	Rural			Urban		
	Hinterland	**Intermediate**	**Peri-Urban**	**Towns**	**Secondary Cities**	**Primary Cities**
=1 if the household is poor	-0.098***	-0.069***	-0.040**	-0.022	-0.010	-0.010
	(0.037)	(0.015)	(0.017)	(0.023)	(0.013)	(0.013)
Household purchased share	0.252***	0.102***	0.076***	-0.014	-0.026	-0.014
	(0.074)	(0.024)	(0.026)	(0.043)	(0.040)	(0.012)
Constant	-0.270	0.039	-0.100	-0.474***	0.026	-0.102**
	(0.173)	(0.039)	(0.091)	(0.061)	(0.069)	(0.052)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes
N	565	3,014	1,499	306	969	2,486
R^2	0.367	0.215	0.360	0.716	0.425	0.184

Notes: Dependent variable: natural log of the expensiveness index. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table 2.10 Expensiveness index regression results, using poverty gap and poverty severity with two poverty lines

	Rural			Urban		
	Hinterland	**Intermediate**	**Peri-Urban**	**Towns**	**Secondary Cities**	**Primary Cities**
Poverty Gap (in USD)	-0.073	-0.545***	-0.343**	0.012	-0.073	-0.063
	(0.291)	(0.143)	(0.153)	(0.263)	(0.116)	(0.174)
Poverty Severity (Poverty Gap Squared)	-0.711	0.745**	0.396	-0.144	0.066	0.022
	(0.481)	(0.318)	(0.295)	(0.542)	(0.230)	(0.474)
Average marginal effect of poverty gap	-0.171	-0.430***	-0.293**	0.002	-0.067	-0.063
	(0.229)	(0.098)	(0.120)	(0.227)	(0.097)	(0.165)
Household purchased share	0.188**	0.103***	0.074***	-0.019	-0.026	-0.014
	(0.077)	(0.024)	(0.026)	(0.043)	(0.040)	(0.012)
Constant	-0.258	0.043	-0.112	-0.479***	0.026	-0.102**
	(0.177)	(0.035)	(0.084)	(0.063)	(0.069)	(0.052)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes
N	565	3,014	1,499	306	969	2,486
R^2	0.394	0.218	0.364	0.715	0.426	0.184

Notes: Dependent variable: natural log of the expensiveness index. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

APPENDIX B: Supplemental Tables

Table 2.11 Share of transactions at each type of retail outlet within each spatial category, using 16 products of interest

	Rural				Urban			
	Hinterland	Intermediate	Peri-urban	Overall	Town	Secondary City	Primary City	Overall
Market	46%	45%	45%	45%	44%	47%	44%	44%
Shop (duka)	47%	52%	53%	52%	55%	52%	55%	54%
Street vendor	6%	3%	2%	3%	2%	1%	1%	1%
Supermarket	0.01%	0.03%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
TOTAL	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 2.12 Overall share of each retail type in each rural tercile/geographic zone, using quantity purchased (in kg unless otherwise noted)

	Hinterland				Intermediate				Peri-Urban				Overall
	T1	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Rural**
Number of households	190	191	184	565	1008	1013	993	3014	509	512	478	1499	5078
Rice													
Market	30%	30%	39%	34%	36%	26%	29%	29%	29%	26%	25%	26%	29%
Shop	67%	70%	60%	64%	64%	74%	71%	70%	71%	74%	75%	74%	71%
Street Vendor	3%	1%	2%	2%	1%	0%	1%	1%	0%	0%	0%	0%	0%
White Maize Flour													
Market	34%	31%	29%	30%	43%	29%	29%	33%	53%	31%	33%	38%	34%
Shop	65%	69%	71%	70%	52%	69%	69%	65%	46%	68%	67%	61%	64%
Street Vendor	1%	0%	0%	0%	6%	1%	1%	3%	0%	1%	0%	0%	2%
Cooking Oil (liters)													
Market	17%	20%	20%	20%	22%	18%	10%	15%	30%	18%	16%	20%	17%
Shop	83%	80%	80%	80%	78%	82%	90%	85%	70%	82%	84%	80%	83%
Street Vendor	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Tomatoes													
Market	87%	72%	85%	81%	81%	79%	83%	81%	80%	83%	91%	85%	82%
Shop	7%	4%	4%	5%	15%	19%	15%	16%	17%	14%	8%	12%	14%
Street Vendor	6%	24%	11%	14%	4%	2%	2%	3%	3%	3%	0%	2%	4%
Onions													
Market	83%	76%	85%	82%	81%	83%	81%	81%	69%	80%	91%	82%	82%
Shop	10%	8%	10%	9%	18%	17%	18%	17%	29%	19%	8%	17%	16%
Street Vendor	6%	16%	5%	9%	1%	1%	2%	1%	1%	0%	1%	1%	2%
Spinach / Lettuce													
Market	81%	70%	65%	70%	79%	75%	79%	78%	82%	85%	88%	85%	79%
Shop	2%	2%	15%	8%	5%	8%	6%	6%	4%	4%	3%	4%	6%
Street Vendor	17%	28%	20%	21%	17%	17%	15%	16%	14%	11%	9%	11%	15%
Other leafy vegetables													

Table 2.12, cont.

	Hinterland				Intermediate				Peri-Urban				Overall
	T1	T2	T3	Overall	T1	T2	T3	Overall	T1	T2	T3	Overall	Rural
Market	78%	86%	69%	76%	75%	73%	81%	77%	82%	84%	80%	82%	78%
Shop	11%	3%	21%	13%	8%	7%	4%	6%	9%	7%	6%	7%	7%
Street Vendor	12%	12%	11%	11%	17%	19%	16%	17%	9%	9%	14%	11%	14%
Dried fish													
Market	61%	57%	80%	67%	81%	77%	76%	78%	52%	66%	70%	63%	74%
Shop	36%	39%	19%	31%	18%	22%	23%	21%	46%	31%	29%	35%	25%
Street Vendor	3%	4%	1%	3%	1%	1%	1%	1%	3%	2%	1%	2%	1%
Beef with bones													
Market	97%	75%	45%	59%	80%	65%	69%	70%	71%	76%	81%	77%	71%
Shop	2%	16%	54%	38%	18%	32%	30%	29%	28%	23%	19%	22%	27%
Street Vendor	1%	9%	2%	3%	2%	2%	1%	2%	1%	1%	0%	1%	1%
Fresh cow milk (liters)													
Market	29%	44%	31%	34%	48%	31%	11%	19%	25%	27%	7%	18%	21%
Shop	2%	10%	19%	15%	22%	23%	41%	35%	53%	53%	37%	47%	37%
Street Vendor	69%	46%	50%	50%	30%	47%	48%	46%	21%	20%	55%	35%	43%

Table 2.13 Overall share of each retail type in each urban tercile/geographic zone, using quantity purchased (in kg unless otherwise noted)

	Towns				Secondary Cities				Primary Cities				Overall
	T1	T2	T3	Overall	T1	T2	T3	Overall	T1	T2	T3	Overall	Urban
Number of households	103	104	99	306	337	330	302	969	914	854	718	2486	3761
Rice													
Market	44%	23%	57%	43%	14%	24%	30%	24%	10%	12%	11%	11%	17%
Shop	56%	77%	43%	57%	85%	76%	69%	76%	90%	88%	89%	89%	83%
Street Vendor	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
White Maize Flour													
Market	35%	39%	35%	36%	13%	14%	19%	15%	7%	7%	5%	6%	11%
Shop	65%	61%	65%	64%	87%	86%	81%	85%	93%	93%	95%	94%	89%
Street Vendor	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Cooking Oil (liters)													
Market	15%	9%	12%	12%	17%	13%	15%	15%	5%	6%	7%	6%	8%
Shop	85%	91%	87%	88%	82%	85%	85%	84%	95%	94%	93%	94%	92%
Street Vendor	0%	0%	1%	0%	1%	2%	0%	1%	0%	0%	0%	0%	0%
Tomatoes													
Market	96%	93%	95%	95%	89%	94%	95%	93%	97%	97%	96%	97%	96%
Shop	4%	6%	4%	5%	9%	6%	5%	6%	3%	3%	4%	3%	4%
Street Vendor	0%	1%	1%	1%	2%	0%	0%	1%	0%	0%	0%	0%	0%
Onions													
Market	82%	89%	90%	88%	87%	91%	92%	90%	97%	97%	96%	96%	94%
Shop	12%	11%	8%	10%	12%	9%	8%	9%	3%	3%	4%	3%	5%
Street Vendor	6%	0%	2%	3%	1%	1%	0%	1%	0%	0%	0%	0%	0%
Spinach / Lettuce													
Market	92%	90%	86%	89%	96%	94%	96%	95%	89%	88%	88%	88%	90%
Shop	3%	2%	2%	2%	2%	2%	2%	2%	3%	2%	3%	3%	2%
Street Vendor	4%	9%	12%	9%	2%	3%	2%	2%	9%	10%	9%	9%	7%
Other leafy vegetables													

Table 2.13, cont.

	Towns				Secondary Cities				Primary Cities				Overall Urban
	T1	T2	T3	Overall	T1	T2	T3	Overall	T1	T2	T3	Overall	
Market	89%	79%	85%	85%	86%	89%	88%	88%	86%	85%	87%	86%	86%
Shop	2%	4%	2%	3%	5%	4%	3%	4%	3%	3%	3%	3%	3%
Street Vendor	8%	17%	13%	13%	9%	7%	9%	8%	12%	12%	10%	11%	11%
Dried fish													
Market	69%	64%	85%	71%	76%	82%	87%	81%	85%	94%	96%	91%	85%
Shop	30%	36%	15%	28%	16%	14%	13%	15%	14%	6%	4%	9%	13%
Street Vendor	1%	0%	0%	0%	8%	3%	0%	5%	1%	1%	1%	1%	2%
Beef with bones													
Market	80%	84%	84%	83%	80%	76%	91%	83%	65%	68%	54%	62%	69%
Shop	20%	16%	16%	17%	17%	24%	9%	17%	35%	32%	46%	38%	31%
Street Vendor	0%	0%	0%	0%	3%	0%	0%	1%	0%	0%	0%	0%	0%
Fresh cow milk (liters)													
Market	2%	28%	20%	21%	17%	29%	12%	19%	5%	6%	4%	5%	9%
Shop	65%	23%	44%	38%	23%	43%	61%	49%	89%	79%	79%	82%	71%
Street Vendor	33%	49%	37%	40%	60%	28%	27%	32%	6%	15%	17%	13%	19%

Table 2.14 Average price per kg or liter, by urban and rural

	URBAN						RURAL					
	<0.25 kg or liter	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1)	[1, 1.5)	>=1.5 kg or liter	<0.25 kg or liter	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1)	[1, 1.5)	>=1.5 kg or liter
Staples												
Rice	1	1.05	1.05	1.03	1.03	0.99	1	1.00	0.97	0.97	0.95	0.91
White maize flour	1	0.65	0.59	0.58	0.57	0.56	1	0.71	0.66	0.64	0.63	0.60
Cooking oil (liters)	1	0.71	0.70	0.55	0.74	0.71	1	0.65	0.64	0.38	0.63	0.56
Fresh tomatoes	1	0.58	0.51	0.44	0.54	0.56	1	0.59	0.47	0.46	0.44	0.44
Onions	1	0.59	0.61	0.49	0.80	0.60	1	0.50	0.50	0.40	0.48	0.46
Non-Staples												
Spinach / Lettuce	1	0.54	0.39	0.31	0.30	0.25	1	0.53	0.36	0.29	0.27	0.22
Other leafy vegetables	1	0.57	0.43	0.37	0.35	0.28	1	0.50	0.37	0.29	0.24	0.22
Dried fish	1	0.63	0.50	0.44	0.60	0.38	1	0.45	0.35	0.29	0.34	0.26
Beef with bones	1	1.13	1.04	1.02	1.01	0.94	1	1.03	0.95	0.81	0.87	0.73
Fresh cow milk (liters)	1	1.07	0.96	1.04	0.95	0.90	1	1.10	0.89	0.88	0.72	0.62

Table 2.15 Average household shares of each of six categories (own-farm, other self-employment, and four wage: hourly, daily, weekly, and monthly) in total FTEs

RURAL	Hinterland				Intermediate				Peri-Urban				Overall
	T1	T2	T3	Overall	T1	T2	T3	Overall	T1	T2	T3	Overall	Rural
Number of households	190	191	184	565	1008	1013	993	3014	509	512	478	1499	5078
Share of … in total FTEs													
Own-farm work	**92%**	**89%**	**70%**	**84%**	**90%**	**80%**	**60%**	**77%**	**81%**	**60%**	**33%**	**58%**	**73%**
Hourly/Daily/Self-Employment	**3%**	**5%**	**12%**	**6%**	**8%**	**16%**	**27%**	**17%**	**14%**	**26%**	**36%**	**26%**	**16%**
Hourly wage work	0.00%	0.00%	0.00%	0.00%	0.05%	0.18%	0.39%	0.21%	0.01%	0.04%	0.00%	0.02%	0.1%
Daily wage work	0.4%	0.5%	0.0%	0.3%	1.0%	1.6%	1.5%	1.4%	1.9%	3.4%	5.7%	3.7%	2%
Self-employment	3%	4%	12%	6%	7%	14%	25%	15%	12%	23%	30%	22%	14%
Weekly wage work	**1%**	**1%**	**1%**	**1%**	**0%**	**1%**	**1%**	**1%**	**1%**	**2%**	**3%**	**2%**	**1%**
Monthly wage work	**4%**	**6%**	**16%**	**9%**	**2%**	**4%**	**11%**	**6%**	**4%**	**11%**	**28%**	**15%**	**10%**
URBAN	**Towns**				**Secondary Cities**				**Primary Cities**				**Overall**
	T1	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Number of households	103	104	99	306	337	330	302	969	914	854	718	2486	3761
Share of … in total FTEs													
Own-farm work	**41%**	**25%**	**6%**	**24%**	**14%**	**10%**	**5%**	**10%**	**4%**	**1%**	**1%**	**2%**	**12%**
Hourly/Daily/Self-Employment	**42%**	**50%**	**46%**	**46%**	**57%**	**58%**	**54%**	**56%**	**61%**	**55%**	**43%**	**53%**	**52%**
Hourly wage work	0.00%	0.22%	0.00%	0.07%	0.20%	0.48%	1.02%	0.57%	0.57%	0.51%	1.12%	0.73%	0.5%
Daily wage work	3.1%	2.6%	1.7%	2.5%	7.2%	6.2%	4.3%	5.9%	9.5%	11.2%	7.4%	9.3%	6%
Self-employment	39%	47%	44%	43%	50%	51%	49%	50%	51%	43%	34%	43%	45%
Weekly wage work	**0%**	**2%**	**1%**	**1%**	**1%**	**0%**	**2%**	**1%**	**3%**	**3%**	**2%**	**3%**	**2%**
Monthly wage work	**17%**	**23%**	**48%**	**29%**	**27%**	**32%**	**39%**	**33%**	**32%**	**42%**	**54%**	**42%**	**35%**

Table 2.16 Share of households consuming each product and conditional purchase shares by geographic zone

	Rural			Urban		
	Hinterland	Intermediate	Peri-urban	Towns	Secondary cities	Primary cities
Rice						
Consumed	77%	72%	80%	83%	94%	94%
Purchased (share of those who consumed)	85%	93%	96%	100%	100%	100%
White Maize Flour						
Consumed	74%	77%	76%	69%	89%	93%
Purchased	76%	72%	74%	100%	99%	99%
Cooking Oil						
Consumed	74%	71%	68%	75%	78%	89%
Purchased	98%	99%	99%	100%	100%	100%
Tomatoes						
Consumed	89%	88%	93%	97%	97%	97%
Purchased	98%	97%	99%	100%	100%	100%
Onions						
Consumed	82%	80%	88%	93%	93%	96%
Purchased	99%	99%	100%	100%	100%	100%
Spinach/Lettuce						
Consumed	78%	68%	78%	89%	84%	84%
Purchased	82%	77%	86%	98%	98%	100%
Other Leafy						
Consumed	88%	77%	79%	80%	66%	76%
Purchased	64%	58%	71%	92%	97%	99%
Dried Fish						
Consumed	80%	77%	75%	75%	84%	69%
Purchased	99%	100%	100%	100%	100%	100%
Beef with Bones						
Consumed	49%	51%	62%	84%	78%	86%
Purchased	97%	98%	99%	100%	100%	100%

Table 2.16, cont.

		Rural			Urban		
		Hinterland	Intermediate	Peri-urban	Towns	Secondary cities	Primary cities
Fresh cow milk							
	Consumed	26%	24%	25%	36%	23%	25%
	Purchased	91%	72%	83%	96%	99%	99%

Table 2.17 Poverty statistics - 0/1 poverty indicator, poverty gap, and poverty severity

RURAL		**Hinterland**				**Intermediate**				**Peri-Urban**			**Overall**
	T1	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Rural**
Number of households	190	191	184	565	1008	1013	993	3014	509	512	478	1499	5078
Average monthly total expenditure (USD)	$21	$41	$89	$47	$20	$35	$78	$42	$22	$43	$99	$49	$46
Share of households that are poor (total expenditure per AE less than $24.32)	64%	0%	0%	23%	78%	0%	0%	29%	58%	0%	0%	23%	25%
Average poverty gap (USD) (=0 if not poor)	0.18	0	0	0.07	0.21	0	0	0.08	0.15	0	0	0.06	0.07
Average poverty severity (=0 if not poor)	0.07	0	0	0.03	0.08	0	0	0.03	0.06	0	0	0.02	0.03
URBAN		**Towns**				**Secondary Cities**				**Primary Cities**			**Overall**
	T1	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**T1**	**T2**	**T3**	**Overall**	**Urban**
Number of households	103	104	99	306	337	330	302	969	914	854	718	2486	3761
Average monthly total expenditure (USD)	$33	$65	$171	$88	$31	$58	$130	$68	$48	$89	$196	$105	$87
Share of households that are poor (using same poverty line as for rural households)	21%	0%	0%	7%	21%	0%	0%	8%	5%	0%	0%	2%	6%
Average poverty gap (USD)	0.05	0	0	0.02	0.04	0	0	0.02	0.01	0	0	0	0.01
Average poverty severity	0.02	0	0	0.01	0.01	0	0	0.01	0	0	0	0	0
Using a 30% higher poverty line for urban households													
Share of households that are poor	41%	0%	0%	13%	51%	0%	0%	20%	12%	0%	0%	5%	13%
Average poverty gap (USD)	0.11	0	0	0.04	0.12	0	0	0.04	0.03	0	0	0.01	0
Average poverty severity	0.04	0	0	0.01	0.04	0	0	0.02	0.01	0	0	0	0.01

Note: total expenditure variable has been winsorized

Table 2.18 OLS results for bulk discount equation: rural households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Natural log of quantity	-0.027***	-0.070***	-0.341***	-0.504***	-0.493***	-0.541***	-0.097***	-0.150***
	(0.004)	(0.005)	(0.009)	(0.010)	(0.008)	(0.009)	(0.020)	(0.016)
Natural log of total expenditure per AEQ	-0.118	0.536***	0.229	0.125	0.226	0.642**	0.359*	0.008
	(0.092)	(0.179)	(0.237)	(0.204)	(0.296)	(0.298)	(0.217)	(0.433)
Natural log of total expenditure per AEQ, squared	0.006	-0.022***	-0.004	0.000	-0.005	-0.026*	-0.014	0.002
	(0.004)	(0.008)	(0.011)	(0.009)	(0.013)	(0.013)	(0.009)	(0.019)
Average marginal effect of natural log of total expenditure per AEQ	0.015***	0.057***	0.147***	0.126***	0.110***	0.069***	0.052***	0.044**
	(0.004)	(0.009)	(0.013)	(0.011)	(0.013)	(0.015)	(0.012)	(0.019)
Household purchased share	-0.012	0.038	-0.100***	0.111***	0.004	-0.133***	0.151***	0.065
	(0.012)	(0.024)	(0.032)	(0.027)	(0.030)	(0.032)	(0.035)	(0.048)
Shop	0.017***	0.047***	0.026	-0.047***	-0.015	-0.070***	0.002	0.066***
	(0.006)	(0.012)	(0.017)	(0.014)	(0.016)	(0.016)	(0.017)	(0.022)
Street Vendor	-0.091*	-0.048	0.074	-0.027	0.079**	0.131**	-0.249***	0.078**
	(0.051)	(0.037)	(0.058)	(0.023)	(0.038)	(0.053)	(0.086)	(0.034)
Intermediate	0.022***	-0.039***	-0.111***	0.008	0.032*	-0.038**	0.014	-0.008
	(0.006)	(0.014)	(0.019)	(0.016)	(0.017)	(0.018)	(0.022)	(0.029)
Hinterland	0.003	0.014	-0.034	0.141***	0.127***	0.003	-0.021	-0.033
	(0.010)	(0.020)	(0.034)	(0.026)	(0.026)	(0.030)	(0.032)	(0.034)
Constant	8.153***	3.550***	5.321***	5.334***	4.491***	3.232*	5.946***	6.390***
	(0.521)	(1.023)	(1.306)	(1.131)	(1.635)	(1.651)	(1.238)	(2.449)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19,533	25,472	36,540	44,407	32,186	17,710	4,437	3,568
R^2	0.478	0.333	0.482	0.548	0.565	0.615	0.298	0.774

Notes: Clustered (at household level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table 2.19 OLS results for bulk discount equation: urban households

	Rice	Maize Flour	Oil	Tomatoes	Onions	Dried Fish	Beef	Fresh Cow Milk
Natural log of quantity	-0.020***	-0.038***	-0.227***	-0.426***	-0.371***	-0.470***	-0.070***	-0.062***
	(0.003)	(0.003)	(0.007)	(0.010)	(0.010)	(0.014)	(0.013)	(0.010)
Natural log of total expenditure per AEQ	0.033	0.155*	0.160	-0.421**	-0.572**	-0.112	-0.209	-0.165
	(0.059)	(0.094)	(0.192)	(0.181)	(0.240)	(0.382)	(0.134)	(0.142)
Natural log of total expenditure per AEQ, squared	-0.000	-0.006	-0.004	0.021***	0.029***	0.008	0.009*	0.008
	(0.002)	(0.004)	(0.008)	(0.008)	(0.010)	(0.017)	(0.005)	(0.006)
Average marginal effect of natural log of total expenditure per AEQ	0.027***	0.019***	0.070***	0.075***	0.097***	0.062***	0.010	0.021**
	(0.004)	(0.004)	(0.010)	(0.011)	(0.012)	(0.018)	(0.009)	(0.010)
Household purchased share	0.033***	0.018	0.079**	0.178***	0.214***	0.057	0.020	0.079**
	(0.011)	(0.013)	(0.031)	(0.034)	(0.037)	(0.058)	(0.029)	(0.033)
Shop	0.003	-0.024***	0.037***	0.098***	0.006	-0.045**	-0.006	-0.008
	(0.004)	(0.007)	(0.012)	(0.020)	(0.018)	(0.022)	(0.011)	(0.017)
Street Vendor	-0.068	-0.050*	-0.023	-0.064*	0.095	0.149***	-0.266**	0.013
	(0.048)	(0.028)	(0.080)	(0.037)	(0.075)	(0.052)	(0.111)	(0.022)
Secondary City	0.011	-0.015	-0.098***	0.041	-0.111***	0.041	0.133***	0.114***
	(0.011)	(0.025)	(0.035)	(0.034)	(0.040)	(0.043)	(0.043)	(0.030)
Primary City	-0.041***	-0.048**	-0.221***	-0.169***	-0.006	-0.054	-0.030	0.040
	(0.014)	(0.021)	(0.035)	(0.036)	(0.051)	(0.051)	(0.024)	(0.085)
Constant	7.177***	5.884***	6.434***	8.678***	9.243***	7.660***	9.525***	7.661***
	(0.356)	(0.544)	(1.123)	(1.070)	(1.393)	(2.205)	(0.824)	(0.843)
Includes region, month, and region*month interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	35,171	42,988	43,891	46,234	43,291	8,798	9,021	4,181
R^2	0.446	0.341	0.356	0.531	0.428	0.483	0.148	0.582

Notes: Clustered (at household level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

REFERENCES

REFERENCES

Attanasio, O. & Frayne, C. (2006). Do the poor pay more? Working paper.

Bai, Y., Alemu, R., Block, S.A., Headey, D., & Masters, W.A. (2021). Cost and affordability of nutritious diets at retail prices: Evidence from 177 countries. *Food Policy*, *99,* 1-17.

Beatty, T.K.M. (2010). Do the poor pay more for food? Evidence from the United Kingdom. *American Journal of Agricultural Economics*, *92*(3), 608-621.

Bennett, D.L., Faria, H.J., Gwartney, J.D., & Morales, D.R. (2017). Economic institutions and comparative economic development: A post-colonial perspective. *World Development, 96*, 503-519.

Christiaensen, L., De Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: The role of rural diversification and secondary towns. *Agricultural Economics*, *44*, 435-447.

Christiaensen, L., & Todo, Y. (2014). Poverty reduction during the rural-urban transformation - The role of the missing middle. *World Development*, *63,* 43-58.

Christiaensen, L. & Kanbur, R. (2017). Secondary towns and poverty reduction: Refocusing the urbanization agenda. *Annual Review of Resource Economics*, *9*, 405-419.

Chung, C. & Myers, S.L. (1999). Do the Poor Pay More for Food? An Analysis of Grocery Store Availability and Food Price Disparities. *The Journal of Consumer Affairs*, *33*(2), 276-296.

Demmler, K. M., Ecker, O., & Qaim, M. (2018). Supermarket shopping and nutritional outcomes: A panel data analysis for urban Kenya. *World Development, 102,* 292-303.

Dantzig, G. (1991). Stigler's nutrition model: An example of formulation and solution. In Linear Programming and Extensions (pp. 551-567). Princeton University Press: Princeton, New Jersey.

Dillon, B., De Weerdt, J., & O'Donoghue, T. (2021). Paying more for less: Why don't households in Tanzania take advantage of bulk discounts? *The World Bank Economic Review, 35*(1), 148-179.

Frank, R., Douglas, S., & Polli, R. (1967). Household correlates of package-size proneness for grocery products. *Journal of Marketing Research, IV,* 381-384.

Kimenju, S.C., Rischke, R., Klasen, S., & Qaim, M. (2015). Do supermarkets contribute to the obesity pandemic in developing countries? *Public Health Nutrition, 18,* 3224-3233.

Kunreuther, H. (1973). Why the poor pay more for food: Theoretical and empirical evidence. *The Journal of Business, 46*(3), 368-383.

MacDonald, J. & Nelson, P. (1991). Do the poor still pay more? Food price variations in large metropolitan areas. *Journal of Urban Economics, 30*(3), 344-359.

Mendoza, R.U. (2011). Why do the poor pay more? Exploring the poverty penalty concept. *Journal of International Development*, *23*, 1-28.

Montiel Olea, J.L. & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business Economics and Statistics, 31*(3), 358-369.

OECD. (2020). Life in cities, towns, and semi-dense areas, and rural areas. Obtained February 17, 2021, from https://www.oecd-ilibrary.org/sites/cd35184c-en/index.html?itemId=/content/component/cd35184c-en

Orhun, A.Y. & Palazzolo, M. (2018). Frugality is hard to afford. Ross School of Business Paper No. 1309.

Mussa, R. (2015). Do the poor pay more for maize in Malawi? *Journal of International Development*, *27*, 546-563.

Rao, V. (2000). Price heterogeneity and “real” inequality: A case study of prices and poverty in rural south India. *Review of Income and Wealth*,*46*(2), 201-211.

Rischke, R., Kimenju, S., Klasen, S., & Qaim, M. (2015). Supermarkets and food consumption patterns: The case of small towns in Kenya. *Food Policy*, *52*, 9-21.

United Republic of Tanzania. (2013). Key Findings: 2011/12 Household Budget Survey Tanzania Mainland. National Bureau of Statistics, Ministry of Finance: Dar es Salaam.

United Republic of Tanzania. (2016). Criteria and procedures for establishing/promoting local government authorities. Accessed July 26, 2016 from http://www.nachingweadc.go.tz/storage/app/uploads/public/5a1/7e0/343/5a17e034314be844048999.pdf

World Bank. (2009). Africa’s Infrastructure: A Time for Transformation. World Bank Publications.

ESSAY 3: THE MODERNIZATION OF THE MAIZE FLOUR VALUE CHAIN IN TANZANIA

3.1 Introduction

The agrifood system in Tanzania, like in many other countries across the developing world, is rapidly transforming. Annual economic growth of 6-7% is intimately linked with, and is a driver of, both demand-side and supply-side changes in the food system (World Bank 2019). The increase in income drives diversification in food consumption: namely, as incomes rise, we tend to see a greater demand for fruits and vegetables, animal protein, and convenient foods such as highly processed food and food away from home, which is coupled with a decrease in demand for staple grains (a result in agricultural economics known as Bennett's Law) (Bennett 1941). At the same time, changes in food demand encourage changes in the organization of agrifood supply chains. For example, supply chains are lengthening geographically and the flows of food are diversifying; not only is food moving from primarily rural production areas to urban consumption zones, but urban processed food is finding its way to even hinterland rural areas (Reardon and Timmer 2014, Reardon et al. 2019). In the maize flour value chain, recent years have seen an explosion in the number of brands sold in secondary cities and the surrounding rural areas in Tanzania, displacing the traditional custom-milled and unbranded flour.

Despite the food system changes that are apparent on the ground in developing countries like Tanzania, the majority of the literature on transforming food systems has been on niche products like horticulture and niche marketing channels like exports and supermarkets, and sometimes the intersection of the two (e.g., various papers such as Maertens and Swinnen 2009 and Maertens et al. 2012 look at horticulture exports). There has been relatively little focus on staple food value chains in Africa, with the exception of Minten et al.'s 2015 description of the teff value chain in Ethiopia, and Soullier and Moustier's 2020 paper on Senegal's modernizing rice value chain.

Within the literature on supermarketization there has been a debate on how governments can foster "competitiveness with inclusiveness" (for an overview, see Reardon and Gulati 2008); that is, how to craft policies to help traditional retailers modernize so as to be able to compete with supermarkets and other modern retail outlets that threaten to steal business from the traditional retailers. Reardon and Gulati (2008) cite several examples of municipalities helping wetmarkets upgrade their infrastructure so they can better compete with the encroaching supermarkets. Other research has been done on the transformation of traditional supply chain actors. Verhofstadt and Maertens (2013) document four main horticulture supply chains in Rwanda and hence argue that the horticulture marketing system is more complicated than the conventional 'traditional versus modern' dichotomy because traditional markets are beginning to modernize and form a "transitional" value chain. While not on modernizing traditional retailers per se, Qaim et al. (2019) note that, while the conventional view is that supermarkets are selling the bulk of ultra-processed junk food, traditional shops are also selling such food. In general, though, the conventional wisdom we perceive is that the traditional supply chain is not changing, a view echoed very recently in Chaboud and Moustier (2021).

Other sub-strands of literature have looked at different dimensions of changing food systems, but it has not yet been brought together in one package. A strand on procurement change looks at disintermediation (i.e. higher levels of vertical coordination) and geographically lengthening supply chains. An early wave of work in the 1980's, led by the anthropologist Norbert Dannhaeuser, documented changes in the structure of wholesale and retail of consumer goods (both food and non-food) in India and the Philippines. Dannhaeuser (1977, 1980, 1984) detailed the urban-rural supply chains of branded products and how urban company agents were displacing the stockists who traditionally worked on the spot market. Some urban companies in India went further by introducing retail franchises and even fully vertically integrated company stores in the secondary cities and rural towns. However, Dannhaeuser's focus was on Asia, not Africa, and he did not conduct systematic surveys of retailers. More recent work (also in Asia) documents both disintermediation and lengthening supply chains in Asian potato value chains (Reardon et al. 2012).

Other sub-strands focus on the rise of self-service grocery stores (a transitional segment of the retail sector) and the rise of branding. On self-service, Reardon and Gulati (2008) discuss the trend in food retail in the United States from small mom-and-pop stores, to self-service shops, to supermarkets and hypermarkets. A developing country example comes from Ethiopia, where the hierarchy of fruit and vegetable retail shops includes the traditional "microsellers" (small informal vendors) and the transitional grocery shops (Assefa et al. 2016). On branding, the marketing literature has several papers that discuss the value of a brand to a manufacturer (so-called 'brand equity'), but little work has been done to document the sale of branded products by food retailers in developing countries (for examples in the marketing literature, see Goldfarb et al. 2009, Ailawadi et al. 2003, and Erdem and Swait 1998). Minten and colleagues study the emergence of branded rice in Bangladesh (Minten et al. 2013a) and the rapid rise of branded *makhana* in India, finding that the share of branded *makhana* rose from 25% to 50% in just 5 years (Minten et al. 2013b). In Africa, several studies document the range of processed, branded foods on offer using inventories, but the focus tends to be on the dominance (or not) of imported brands compared to domestic brands, and these studies only document what is on the shelf at the time of the survey rather than volumes sold of each respective brand (Liverpool-Tasie et al. 2016, Snyder et al. 2015, Theriault et al. 2018).

The papers cited above are mostly descriptive in nature, and none has brought all of these practices and innovations (procurement system changes, the rise in self-service shops, adoption of branding, and other practices under the domain of value chain finance) together in one package. Moreover, none has systematically tested the effects of different town and city sizes, and of type of retail outlet, on the adoption of the various innovations (our hypotheses on the expected effects of city size and retail type are presented in section 2). Most of the cited studies focus on one city (for example, Dannhaeuser's work in India centered on the secondary city of Nasik; Snyder et al. 2015 focus on the primary city of Dar es Salaam in Tanzania). However, the research community has recently come to realize the importance of accounting for the size of the city (Christiaensen et al. 2013, 2014; Sauer et al. 2021), and as suggested in the second paragraph, it would be a mistake to think that traditional supply chain actors remain stagnant and unchanging.

We attempt to fill these gaps in the literature by examining various facets of modernization in the maize flour value chain (both procurement and marketing) using a primary survey of processed food retailers conducted in Tanzania. This article thus adds to the scarce literature on modernizing staple value chains in Africa. We present detailed descriptive statistics on the adoption of different procurement and marketing institutions and technologies, and test the effects of spatiality (as measured by town size and distance from the main highway) and retail type on the adoption of these practices. Our research questions are thus: what are the patterns and determinants of adoption of various procurement and marketing practices (such as disintermediation, lengthening supply chains, and branding), and how does this vary by town size (small and medium-sized rural towns and secondary cities) and retail type (traditional shops or *dukas*, mini-supermarkets, and supermarkets)?

This essay is organized as follows. Section 2 provides a conceptual framework and outlines our hypotheses and the rationale for them. Section 3 describes how we collected the retailer data in Tanzania. Section 4 presents descriptive statistics on the various procurement and marketing practices and a comparison of margins across retailer type and town size. Sections 5 and 6 outline the econometric strategy and discuss the econometric results. Section 7 concludes.

3.2 Conceptual Framework

Table 3.1 outlines our hypotheses on selected outcomes (various measures of the structure, conduct, and performance of the maize flour value chain), by retailer type and town size. To inform our hypotheses, we assume that retailers of different levels of modernization and in different town/city sizes face different levels of transaction costs. Specifically, the underlying assumptions we make are (a) holding constant town size, more traditional retailers face higher transaction costs than more modern retailers, and (b) holding constant retailer type, retailers in smaller towns face higher transaction costs than those in the medium-sized towns or those in the secondary cities. Regarding the former assumption, we believe that modern retailers (e.g. supermarkets) are more likely to have economies of scale that lower their transaction costs. Greater economies of scale in the larger towns and cities also likely alleviate transaction costs, so that a duka in a secondary city faces lower costs than a duka in a small-sized town, all else equal.

3.2.1 Hypotheses on Supply Chain Structure

First, in general, we hypothesize that the *intermediational* length (the number of "hands" that maize flour passes through before reaching the retailer) is shorter for dukas because we expect them to procure mostly from local mills (and mostly sell their flour loose). Mini-supermarkets and supermarkets may rely on more intermediaries like wholesalers to get flour from farther away (e.g., branded flour from secondary cities or Dar es Salaam), but supermarkets could also be disintermediated if they work directly with processors (e.g. if they have contracts with suppliers of branded flour). (Note the underlying assumption in this paragraph is that dukas tend to sell unbranded flour whereas more modern retailers, and especially supermarkets, are more likely to sell branded flour. See below for further discussion on this point.)

Second, we hypothesize that the *geographical* length of the upstream supply chain (i.e. where the maize flour is milled) will be short for dukas in the small towns because we expect them to source from their local mills; short or medium for both dukas and mini-supers in the medium towns, because such towns may be better served by road infrastructure and hence could have greater access to flour from the secondary cities and beyond; and short, medium, or long for all types of retailers in the secondary cities, because they could source from mills within the cities or from further away because they are on the main highway and have the greatest access to flour from all over the country.

3.2.2 Hypotheses on Supply Chain Conduct

We next discuss our conjectures on outcomes regarding the conduct of maize flour retailers, including credit receipt and extension, adoption of mobile money and branding, and use of other practices like delivery and regular relationships. For many of these variables, the expected outcome is ambiguous and hence there is no definitive hypothesis. Take as an example the receipt of supplier credit. It could be the case that the most traditional actors (the dukas) receive the greatest amount of credit from their maize flour suppliers (and that supermarkets receive the lowest), because the dukas face the greatest transaction costs and hence are in greatest need of technology and institutions (such as value chain finance) that minimize transaction costs. In other words, we can imagine that the dukas have the greatest incentive of all the retail types to adopt technologies and engage in practices that help them run their business more efficiently and at a lower cost. However, we could just as easily envision the opposite scenario where dukas receive the lowest amount of credit (and have the lowest mobile money adoption rate, or the lowest rate of regular relationships and delivery). They may have the incentive, but lack the capacity, to benefit from transaction-cost-minimizing practices. Suppliers may be hesitant to lend to or deliver to smaller retailers, or to start an ongoing relationship with them, because of the greater risk inherent in conducting business with a more traditional shop.

Relatedly, if we assume that developments in the product cycle (the evolution from niche to commoditized to differentiated products) depend on adoption of innovations that minimize transaction costs along the supply chain, then we can make the following hypothesis related to the adoption of branded maize flour (i.e., branding is one way of differentiating a product). If mini-supers and supermarkets are the retailers who benefit from and adopt such institutions as credit provision, delivery, and regular relationships, and technologies like mobile money, then we hypothesize that these more modern retailers are more likely to sell branded flour than traditional retailers. This is because credit, delivery, etc. serve to reduce both risk and transaction costs, and the reduction in both helps to pave the way for more differentiated products to hit the market. For example, regular relationships can help to ensure a certain (and consistent) maize flour supply and level of flour quality. Having a consistent relationship with a supplier can also help to reduce search costs (the costs associated with linking demand and supply), as can delivery to the retail shop. Mobile money can reduce the risk that the retailer will not get paid and can reduce negotiation costs. Likewise, delivery and supplier credit can both facilitate business transactions. An environment in which business is generally easier to conduct can better support the emergence of branded flour. Hence, we hypothesize that mini-supers and supermarkets will sell more branded flour on average than dukas.

Finally, regarding extension of credit to their customers, we hypothesize a negative relationship between retailer size and customer credit. This hypothesis rests on an assumed correlation between consumer income and the type of retail outlet patronized (i.e., wealthier customers on average are more likely to shop at supermarkets). We also assume that this correlation is strongest in the small and medium towns as compared to the secondary cities. That is, secondary city duka customers may be wealthier than small/medium town duka customers and perhaps do not need credit as much as they do.

3.2.3 Hypothesis on Supply Chain Performance (Margins)

Our final hypothesis regards the margins received for branded flour. When they sell branded flour, we posit that small town dukas will receive the highest margins because they have the least amount of competition. Margins may decrease for medium town dukas and mini-supers because they have more competition due to a greater density of retailers in the larger towns, and the same logic goes for secondary city dukas and minis. However, margins could be high for supermarkets because they could charge more to wealthier customers as a result of monopsonistic competition (Neven et al. 2009).

3.3 Data

We use primary data collected in 2017 in Tanzania from a stratified random sample (without weights) of processed food retailers. The towns selected for the survey were stratified according to administrative status, population, and spatial location as follows. First, because we wanted variation in the distance to the primary city Dar es Salaam, we conducted a census of all towns and cities along (and within 50 kilometers of) the main east-west highway that runs from Dar es Salaam through the secondary cities of Morogoro and Dodoma. The census, which was performed using Google Maps, includes towns starting from roughly the midway point between Dar es Salaam and Morogoro through to roughly 100 kilometers west of Dodoma. This census yielded 83 towns. Next, we stratified the towns based on four categories of distance from Dar es Salaam by splitting our census into roughly four groups: towns east of Morogoro; towns west of Morogoro but east of the halfway point between Morogoro and Dodoma; towns east of Dodoma but west of the halfway point; and towns west of Dodoma. We identified which towns were classified as “administrative” by the government and stratified them by population. “Small” towns contained less than 10,000 inhabitants and “medium” towns contained between 10,000 and 20,000 residents. Finally, we stratified by distance from the main highway (within 25km or between 25km and 50km). This gives us a total of 16 groups (4 east-west categories crossed with 2 town sizes and 2 distance-from-highway categories). Within each east-west category, one administrative town was randomly selected. Then, from the 16 groups, one town was randomly selected (without replacement of the selected administrative town), for a total of 20 towns. We also sampled retailers from the two secondary cities.

The random sample of retailers in each town was chosen as follows. Within each town or city, the universe of wards was constituted. With the help of key informants (government officials from the local authority), rural wards were identified and excluded from the sampling because they had few or no shops. Around a third of wards were randomly selected, and within each ward, three streets were randomly selected. (For rural towns that were too small to be divided into wards, streets were selected from the whole town). The universe of retail outlets that carried at least one of our products of interest (maize flour, wheat flour, or wheat products like bread, pasta, and cookies) was listed.[21] The types of retailers in our sample include *dukas* (small traditional shops), mini-supermarkets (self-service stores with a single checkout counter), and supermarkets (self-service stores with multiple checkout counters). We refer to these as traditional, transitional, and

[21] The survey had multiple objectives in addition to learning about the maize value chain.

modern retailers, respectively. From the listed retail universe, 8 (16) retailers from each small (medium) town were selected, and 40 retailers from each secondary city were selected for the survey. The universe of mini-supermarkets in each selected street was listed and surveyed. The universe of supermarkets in each town was listed but not all supermarkets was surveyed due to refusal to participate by some.

Table 3.2 provides a breakdown, by size of the town and distance from the highway, of the sample of retailers we use in our analysis. A total of 234 retailers sold maize flour. Of these, 190 are dukas, 41 are mini-supermarkets, and 3 are supermarkets. The majority (90%) of them were in medium towns or one of the two secondary cities, and most (65%) were located either on the highway (all retailers in the secondary cities were on the highway) or within 25 kilometers of the highway. Interestingly, all the mini-supermarkets in medium towns were greater than 25km from the highway (and, as we will discuss below and as is shown in Table 3.4, all started their business between 2012 and 2017).

Survey respondents were asked detailed, SKU-level questions on buying and selling prices, number of units currently in stock and number of units acquired in the latest stock-up, and average high-season and low-season daily quantities sold.[22] Additionally, respondents were asked brand-level questions on procurement practices: who is the direct supplier of that brand, whether the respondent has a 'regular relationship' with that supplier (and if so, why), how the most recent stock was paid for (cash, check, or mobile money) if the stock was paid for immediately (at the time of the transaction) or at a later date, and how the brand was delivered to the retail or wholesale premises. Further maize flour procurement questions were asked in a "now and 5 years ago" format: the percentage of transactions that occur with each 'agent type' (wholesaler, processor,

[22] An SKU is a shop-keeping unit. An SKU is defined by a unique combination of product size (e.g., 5kg vs. 10kg vs. 25kg bags of maize flour) and type (e.g., refined vs. whole-grain maize flour, chocolate vs. strawberry cookies).

other retailer, custom milled, etc.), where the suppliers are coming from and where the maize flour is believed to be milled, the percentage of transactions with the supplier that are paid for in cash or mobile money, the percentage of transactions sold on commission vs. through taking possession, and the percentage of maize flour sold packaged vs. loose. Finally, demographic data collected included the age, sex, education, and other occupation(s) of the business owner, assets owned that are related to the business, family and hired labor employed both temporarily and permanently, and the percentage of food sales in total sales.

3.4 Descriptive Results

We first start with some basic demographic statistics, presented in Table 3.4. First, mini-supers have emerged on the scene in just the past 5 years. None of the surveyed mini-supers in medium towns were in business in 2012, and 81% of the mini-supers in the two secondary cities started their business between 2012 and 2017.

Second, there appears to be a higher degree of specialization (as measured by the percentage of the family income coming from the retail business) in secondary cities and medium towns compared to the small towns. For example, on average 84% (100%) of the family income comes from the business for secondary city mini-supers (supermarkets); in the medium towns, this share is around 60%, and small-town dukas generate just over half the family income on average. On the other hand, specialization as measured by the variety of food products offered is higher for the traditional and transitional retailers in the small and medium towns and secondary cities, whereas supermarkets (unsurprisingly) offer the greatest food product selection. In particular, the share of four traditional Tanzanian food products (maize and wheat products, rice, and cooking oil) is lowest (58% on average) in the supermarkets, and ranges between 70% and 83% on average for all other retail type and town size categories. Third, ownership of cold storage (refrigerators and freezers) is greatest in secondary cities (where over three quarters of all retailers have at least one refrigerator, and around one quarter have at least one freezer), slightly lower in the medium towns (where around 40% have at least one fridge or freezer), and the lowest in small towns.

Next, we discuss developments in procurement practices over time (Table 3.4). First, we note a trend toward disintermediation. From 2012 to 2017, there was a general decline (greatest in the secondary cities) in the percentage of transactions done with wholesalers, and corresponding increases in the percentages of transactions with mill agents, and directly from the mill. This finding echoes the earlier work of Dannhaeuser, who documented a decreased role for traditional (non-specialized) stockists in 1980's Asia (Dannhaeuser 1977, 1980, 1984). All retailer types in the secondary cities (dukas, mini-supers, and supermarkets) have the highest average shares of transactions directly from the processor, whereas small and medium town retailers still procure mainly from wholesalers. Note this result does not fully support our hypotheses. We speculated that dukas in the smaller towns would conduct business directly with mills, but the results suggest that these dukas do business mostly with wholesalers.

Second, most retailers report engaging in regular relationships with at least one direct supplier of branded flour (with an overall average of 65%) (Table 3.4). The prevalence of regular relationships is highest for secondary city dukas (78%) and secondary city mini-supers (80%). Interestingly, only 67% of secondary city supermarkets had a regular relationship with at least one supplier, suggesting that this long-term practice of supply-side coordination is negatively related to the modernity of the retailer, which supports our hypothesis.

Third, in all town sizes and all retailer types, the majority of maize flour is believed to be milled in the secondary cities (Morogoro and Dodoma). The dukas in the small towns have seen the greatest increase in the share of maize flour coming from the secondary cities (an increase of 8% from 2012). This is evidence of a geographically lengthening supply chain and suggests that (contrary to our hypothesis that dukas would source their flour from local mills) secondary city processors are out-competing the maize mills in the small towns. Virtually no retailers anywhere, even in the secondary cities, are selling maize flour from the primary city Dar es Salaam, which also contradicts our hypothesis that modern retailers (i.e., supermarkets) might source their flour from processors located further away.

Fourth, around half of all retailers who sell branded flour receive delivery, and (with the exception of medium town mini-supers) this share rises from small to medium towns and from medium towns to secondary cities. Within secondary cities, it rises from dukas to mini-supers and supermarkets. This evidence supports our hypothesis that receipt of delivery services would exhibit a positive relationship with the size of the retailer in the secondary cities.

When it comes to patterns of value chain finance (Table 3.5), we see that cash is still the predominant method of payment for procurement transactions, though mobile money is making inroads in the small and medium towns. All three supermarkets in our sample reported using exclusively cash for their procurement transactions. This result implies that the retailers who have the greatest incentive to adopt this transaction-cost saving technology are the ones who are indeed adopting. Receipt of supplier credit is highest for secondary city mini-supers (42% of whom received credit) and supermarkets (a third or whom received credit). Around 20% of dukas in the secondary cities and medium town retailers benefited from supplier credit, whereas only 12% of small town dukas got any credit. Within secondary cities, the relationship is an inverted U-curve: 22% of dukas, 42% of mini-supers, and 33% of supermarkets received credit. This is also contrary to the negative relationship we hypothesized. Finally, with regard to consumer credit, for all traditional and transitional retailers in our sample, between 15% and 20% of their consumers received credit, but no supermarkets extended consumer credit. Within our sample of secondary city retailers, we find that the dukas extend the most credit (20% of their maize flour customers receive credit), followed by mini-supers (15% receive credit), and finally supermarkets (no customers get credit), a result that supports our hypothesis.

The number and type of shop-keeping units (SKUs) carried, by bag size, differs quite predictably by retailer type (Table 3.6). Small and medium town dukas tend to carry just two SKUs: loose flour, and a 25kg bag of flour. Medium town mini-supers and secondary city retailers tend to carry more bag sizes (including 5kg and 10kg bags), with secondary city mini-supers and supermarkets offering the best variety. For example, conditional on carrying 10kg bags of flour, the supermarkets offer on average 2 SKUs, compared to an average of 1.3 offered by mini-supers, and an average of 1.1 offered by dukas in the secondary cities.

We find a surprising penetration of branded flour over time, even in the smallest rural towns, and even in dukas (Table 3.7). In 2012, around 75% of dukas in the small and medium towns sold only loose maize flour. In just five years, this share dropped to 26% of small town dukas and 14% of dukas in the medium towns, a finding contrary to our hypothesis that the dukas would still be selling mostly loose, unbranded flour. (The mini-supers and supermarkets in our data have never exclusively sold loose maize flour.) Contrary to conventional wisdom, it is not only the more modern retailers (mini-supers and supermarkets) that are carrying branded flour – even the traditional retailers are selling brands. Moreover, we find that retailers tend to specialize in one brand. For those retailers who carry branded flour, most carry an average of one brand (with the exception of secondary city mini-supers and supermarkets, which carry 1.6 and 3 brands on average, respectively).

These brands appear to be coming mainly from the secondary cities – not Dar es Salaam and not from the smaller local towns (Table 3.7). Processors in the secondary cities are providing fierce competition to smaller town brands – even in the small and medium towns. The overwhelming majority of retailers (in all locations) carry at least one brand from Morogoro or Dodoma (see the section "Branding trends (2017)" in Table 3.7). On the other hand, just 12% and 10% of small and medium town dukas, respectively, carry a small or medium town brand. And the retailers who do carry a more local small/medium town brand are concentrated in the small and medium towns; almost no retailers in the secondary cities report having a small/medium town brand. Taken together, these results show the marketing integration of branded processed food, and it is not just the retailers in the bigger cities that are aiding this market development.

Finally, we examine differences in maize flour margins across retail types, town sizes, and bag sizes (Table 3.8). First, we find in general (with secondary city mini-supers and supermarkets as the exceptions) that larger bag sizes have larger margins. Overall, a 25kg bag has an average margin of 19%, whereas a 5kg bag has a margin of just over 14%. Second, there is no clear pattern when comparing across retail types. All retailers in the secondary cities receive roughly similar margins (between around 15% and 17.5%) for 5kg bags of flour. However, for 10kg bags, dukas report the lowest average margin, but for 25kg bags, dukas report the highest average margin. We hypothesized that the traditional retailers in the secondary cities would have the lowest margins, and that more modern retailers would enjoy higher margins; our results do not support this hypothesis.

3.5 Econometric Methodology

3.5.1 Model Definition

The econometric strategy will explore the patterns and determinants of adoption of the procurement and marketing practices (as well as maize flour margins) described in the descriptive results. In particular, the left-hand side variables will include (all shares are 2017 values):

- The share of procurement transactions directly with the processor (as opposed to a wholesaler or mill agent). From the descriptive statistics, this change seems to be most apparent in the secondary cities. This share serves as a measure of disintermediation in the maize flour value chain.
- The share of maize flour milled in one of the two secondary cities, Morogoro or Dodoma. This share measures the geographical lengthening of the supply chain.
- The share of maize flour sold packaged/branded. This will provide a sense of movement through the product cycle (from commoditized, unbranded/unpackaged flour to flour differentiated through the use of packaging and brands).
- 0/1 variables indicating engagement in or receipt of: regular relationships with suppliers; delivery services; and both supplier and customer credit.
- Margin as a share of buying price. This gives us a measure of performance in the supply chain.

As controls, each regression will include retail type (one dummy each for supermarket or mini-supermarket, with the traditional store ('duka') as the omitted category) and town size (one dummy each for medium-sized towns and secondary cities, with small towns as the omitted town size). These variables are of key interest to our research questions; in particular, the estimates on the retail dummies will allow us to test the hypotheses listed in Table 3.1 and discussed in the conceptual framework section. (We attempted to include an interaction of town size and retail type, but Stata was only able to estimate the coefficients on one particular interaction, so we decided to drop it from the analysis). Other controls include a dummy for greater distance (i.e., 25 kilometers or more) from the east-west highway and an interaction of town size and distance, to assess any differences between small and medium towns that are located further from the main road. Lastly, we add retailer demographic variables (age, sex, and education of the shop owner, the percentage of household income that comes from the retail business, and, as a measure of specialization, the combined share of maize, rice, wheat, and oil in total food space) and asset ownership dummies (where the assets include truck/bus, motorcycle, bicycle, and wheelbarrow).

The margins regression is done at the retailer-SKU level (and only for branded flour, because we lack procurement data on flour sold loose). This means that, in addition to the independent variables described above, we also include the following SKU-level variables: bag size (10kg or 25kg, with 5kg as the baseline), brand type (regional or national, with local small/medium town as the omitted type), and two dummies equal to 1 if the flour is fortified with micronutrients or if the package boasts a health claim.

3.5.2 Choice of Estimators

The variables that measure shares (share of transactions with a processor; share of flour milled in secondary cities; share of flour sold packaged) are bound between 0 and 1, so we use a fractional probit estimator (Dorta 2016, Papke and Wooldridge 1996). The results are found in Panel A of Table 3.9. We estimate the limited dependent variable models (use of regular relationships with suppliers; receipt of delivery services; and receipt of supplier credit and provision of customer credit) with probit, because each of these dependent variables takes only 0 or 1. The average partial effects from these regressions are reported in panel B of Table 3.9. Lastly, the margins variable, which is measured as the margin (the difference between the price at which the maize flour was bought and the price at which it was sold) divided by the buying price, takes values between 0 and 1. Hence we also use the fractional regression technique to estimate the margins model and the results are in Table 3.10. Finally, OLS regression results for all dependent variables are presented in Table 3.11 in Appendix B.

3.6 Econometric Results

We now discuss the results of our econometric estimations. For each set of results, we organize the discussion by the right-hand side variables of interest (retail type, town size, and distance from the highway).

3.6.1 'Share' Variables (Panel A, Table 3.9)

We first discuss the results of the fractional probit regressions in Panel A of Table 3.9. The results indicate that supermarkets and mini-supers alike sell more branded flour on average than dukas, holding all else constant. These results reinforce the descriptive findings and provide more support of our hypothesis that more modern retailers sell more branded flour. Mini-supermarkets tend to get more flour that was milled in a secondary city compared to dukas, while supermarkets tend to get less from secondary cities, *ceteris paribus*.

Compared to their small-town counterparts, retailers in the secondary cities tend to get a greater share of maize flour directly from the processor, a greater share of maize flour that was milled in one of the two secondary cities in our study, and a greater share of their flour is sold packaged/branded. These results generally support our hypotheses that secondary city retailers would be more disintermediated, and that branded flour would form a larger share of total retailed flour. On the other hand, medium town retailers are less likely to get flour directly from the processor, suggesting that they are more 'intermediated' than small town retailers.

3.6.2 0/1 Variables (Regular Relationship, Delivery, and Credit) (panel B, Table 3.9)

Next, we cover the results for the various variables that take values of either 0 or 1. We first note that no coefficient was significant in the "regular relationship" or "supplier credit" regressions. That is, the results imply that the retailers in our sample are equally likely to engage in regular relationships and receive credit from at least one supplier of branded flour, regardless of whether they are more modern, which size town they are in, or how far from the highway they are.

Thus, our hypothesis that there is a "J curve" relationship between retail type and probability of having a regular relationship or benefitting from supply-side credit finds no support. Similarly, retail type is also not significant in the other two regressions, meaning that transitional and modern retailers (mini-supermarkets and supermarkets, respectively) are not more likely to receive delivery services or to extend credit to their customers than traditional dukas. These findings are contrary to our hypotheses, because we had predicted some kind of relationship between the modernity of the retailer and the likelihood of receiving or giving various services.

When it comes to town size, the probit results suggest that medium town retailers are more likely than small-town stores to give consumer credit (specifically, they are 22 percentage points more likely on average to let their customers buy on credit). Secondary city retailers are 34 percentage points more likely to have flour delivered to their shops than retailers in the small towns, holding constant retail type and retailer demographics.

Finally, the estimated coefficient on the "far from the highway" variable is not statistically significant in any regression, which means that retailers in the small and medium towns far from the main transportation artery are not more or less likely to receive services like delivery and credit from their suppliers, or offer credit to their customers.

3.6.3 Margins (Share of Margin in Buying Price) (Table 3.10)

Lastly, we discuss the results of the margins (more precisely, the share of the margin in buying price) regression (Table 3.10). (Recall that this is an SKU-level regression that only includes branded flour, because we lack procurement data for loose flour.) Interestingly, we find that retail type, town size, and distance are all insignificant, implying that modern retailers, retailers in the bigger towns/cities, and retailers located farther away from the main highway do not sell their branded flour at higher or lower average margins than dukas, small town retailers, or retailers located on or close to the highway, respectively.

The product-level attributes appear to be more important to the margin. 10kg bags of flour have slightly lower margins (on average, 3% lower) than 5kg bags, *ceteris paribus*, whereas 25kg bags do not have statistically different margins than 5kg bags. Regional and national brands have higher margins (and the results are significant at the 5% level) than local small/medium town brands. In particular, regional brands fetch 8% higher margins on average and national brands fetch 12% higher margins, holding the other variables constant. And finally, whether the flour is fortified with micronutrients or labeled with some other health claim is not significantly correlated with the margin.

In summary, several key findings stand out from the econometric analyses. More modern retailers are more likely to sell branded flour, but are *not* more likely to form regular supplier relationships, get delivery from their suppliers, or give or receive credit. Medium town retailers use intermediationally longer supply chains (because they are less likely to procure directly from the processor), which suggests there is a continued role for wholesalers to supply flour to the medium towns. Taken together, these results suggest that penetration of branded flour is happening, but mainly in transitional and modern shops. Disintermediation is most evident in the secondary cities but not medium towns, so wholesalers are still a crucial link in the maize flour value chain.

3.7 Conclusion

In this paper we used primary survey data from processed food retailers in Tanzania to paint a picture of the maize flour value chain in two secondary cities and surrounding small and medium towns. We examined different indices of the structure, conduct, and performance of the value chain, and analyzed how spatiality (particularly the size of the town) and retail type (duka, mini-supermarket, or supermarket) is correlated with the uptake of several procurement and marketing practices and innovations. Our analysis yielded several interesting results.

First, the evidence points to a lengthening maize flour supply chain. The majority of maize flour being sold everywhere (in all town/city sizes) is believed to come from the secondary cities, and this is the same for both modern retailers and traditional dukas. Mini-supermarkets seem especially likely (compared to dukas) to get flour from the secondary cities, according to the econometric results in Table 3.9. This finding of longer supply chains echoes Reardon et al. 2012 who argue that transitional supply chains can be characterized by increasing length.

Second, we find (using our quasi-panel data) a trend toward disintermediation in the supply chain. In small and medium towns, retailers work mostly with wholesalers (though in small towns almost a third of procurement transactions occur directly with the processor); in secondary cities, retailers tend to work mostly with the mill. It is in secondary cities that retailers have transitioned more toward procuring directly from the processor, though in medium towns (especially for mini-supermarkets) there is a role for mill agents. This fascinating result is reminiscent of Dannhaeuser's work on Indian supply chains from urban primary cities to the secondary cities and beyond, where agents for the manufacturers started to take over the role of the traditional (non-dedicated) wholesalers.

Third, regarding value chain finance, we find that modern retailers (particularly in the secondary cities) are most likely to receive credit, and aside from supermarkets, all retailers are about equally as likely to give credit to their customers. Retailers of all types everywhere use mostly cash to buy their maize flour, but mobile money is gaining popularity in the small and medium towns. Interestingly, it appears that the retailers with the highest transaction costs (those in the smaller towns) and with the greatest incentive to adopt this financial technology are actually the ones doing so.

Fourth, most retailers everywhere sell some packed/branded flour. Only 11% of the retailers we surveyed sell exclusively loose flour, and these are all dukas (mostly in small and medium towns). There has been a huge increase from 2012 to 2017 in the number of retailers selling brands, especially in the small and medium towns. This finding is strikingly similar to Minten et al.'s (2013b) that sales of branded *makhana* really took off in the span of just five years, and it points to a strong underlying consumer demand for differentiated maize flour. Moreover, we find that brands from secondary city processors dominate, even in the smaller rural towns, providing competition for local mills.

Fifth, the econometric results suggest that larger brands (i.e., regional and national brands) fetch higher margins than small-town local brands. Thus there appears to be some brand equity inherent in the brands that have a greater geographical reach, and this likely provides an incentive for retailers in the small and medium town towns to carry such brands.

While this paper gives a comprehensive look at procurement and marketing practices and innovations adopted by maize flour retailers in Tanzania, a logical next step in this research would be to conduct similar analyses with other actors in the maize flour value chain, such as wholesalers and processors. In particular, a detailed processor survey and case studies of major maize flour processors in several Tanzanian cities would give researchers and policymakers further insight on the structure, conduct, and performance of this value chain.

APPENDICES

Appendix A: Tables and Figures

Table 3.1 Indices of modernization: hypotheses

	Small Towns	Medium Towns		Secondary Cities		
Behavior/Technology	**Dukas**	**Dukas**	**Mini-Supers**	**Dukas**	**Mini-Supers**	**Supermarkets**
Structure						
Intermediational length	Short	Short	Medium	Short or medium	Short or medium	Short or medium
Geographical length	Short	Short or medium	Short or medium	Short, medium, or long	Short, medium, or long	Short, medium, or long
Conduct						
Value chain (VC) finance – receipt of credit from supplier	High (or low?)	High (or low?)	High or medium	High or medium	High or medium	Low (or highest?)
VC finance – supply of credit to customers	High	High	High or medium	High or medium	High or medium	Low
VC finance – use of mobile money	High (or low?)	High	High or medium	High	High or medium	Low (or highest?)
Other supply-side coordination (delivery services, regular relationships)	High	High	High or medium	High	High or medium	Low (or highest?)
Packaging/branding of maize flour	Low	Low	Medium	Low or medium	Medium	High
When branded – types of brands	Local	Local or regional	Local or regional	Local, regional, or national	Local, regional, or national	Local, regional, or national
Performance						
Margins	High	Medium	Medium	Low or medium	Low or medium	Medium or high?

Table 3.2 Breakdown of maize flour retailer sample over town size, distance from road, and retail type

		Distance from main road		
		<25km from road	**>25km from road**	**Total**
Small town	Dukas	15	8	**23 (10%)**
Medium town	Dukas	54	58	**127 (54%)**
	Mini-Supers	0	15	
	Medium town total	*54*	*73*	
Secondary city	Dukas	55	0	**84 (36%)**
	Mini-Supers	26	0	
	Supermarkets	3	0	
	Secondary city total	*84*	*0*	
	Total	**153 (65%)**	**81 (35%)**	**234**

Table 3.3 Descriptive statistics on demographic variables

	Small Town	Medium Town		Secondary City			
	Dukas	**Dukas**	**Minis**	**Dukas**	**Minis**	**Supers**	**Overall**
# of maize flour retailers not in business in 2012	2	11	15	3	21	1	53
Share not in business in 2012	9%	10%	100%	5%	81%	33%	23%
Age of owner (average)	39	38	35	39	41	55	39
Sex of owner (% male)	91	84	80	71	80	100	81
Education of owner (categorical variable)	2	2	3	3	3	4	3
% of family income from business (average)	55	62	61	68	84	100	66
% who own at least one of the following assets for the business:							
Truck	9	5	13	11	27	0	10
Motorcycle	22	28	13	22	19	33	24
Bicycle	22	34	20	22	23	33	28
Wheelbarrow (*Mkokoteni*)	9	10	0	6	4	0	7
Refrigerator	39	45	47	76	77	100	56
Freezer	35	38	40	27	15	100	34
Average share of … in food space							
Maize products	13	20	23	22	26	12	20
Wheat products	26	23	31	21	26	28	23
Rice	18	14	12	15	19	10	15
Oil	13	14	17	13	13	8	14
Total share of these four products	**70**	**70**	**83**	**70**	**83**	**58**	**72**

Table 3.4 Procurement patterns in 2012 and 2017

	Small Town			Medium Town						Secondary City											
	Dukas			Dukas			Minis			Dukas			Minis			Supers			Overall		
	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change
% of procurement transactions with …																					
Custom Milled	0	0	**0**	2	1	**-1**	5	3	**-2**	4	2	**-2**	0	0	**0**	0	0	**0**	2	1	**-1**
Wholesalers	69	68	**-1**	86	84	**-2**	86	86	**+1**	33	30	**-3**	37	31	**-6**	27	27	**0**	65	63	**-2**
Mill Agents	0	0	**0**	1	3	**+3**	3	7	**+5**	2	4	**+2**	3	5	**+2**	0	0	**0**	1	4	**+2**
Processor (Mill)	31	32	**+1**	12	13	**+1**	7	4	**-4**	61	64	**+3**	60	64	**+4**	73	73	**0**	32	33	**+1**
% of retailers with regular relationship with supplier[a]	n/a	65	**n/a**	n/a	56	**n/a**	n/a	53	**n/a**	n/a	78	**n/a**	n/a	80	**n/a**	n/a	67	**n/a**	n/a	65	**n/a**
% of direct suppliers of maize flour coming from …																					
Local Towns[b]	32	27	**-5**	63	63	**-1**	14	19	**5**	94	94	**0**	91	97	**+6**	67	67	**0**	68	68	**0**
Morogoro	10	9	**-1**	9	7	**-2**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**	5	4	**-1**
Dodoma	38	38	**0**	19	21	**+2**	86	81	**-5**	6	6	**0**	8	2	**-7**	33	33	**0**	21	21	**0**
Dar es Salaam	0	0	**0**	1	1	**0**	0	0	**0**	0	0	**0**	1	1	**0**	0	0	**0**	1	1	**0**
Foreign	20	26	**+6**	7	8	**+1**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**	5	6	**1**
% of maize flour (volume) believed to be milled in …																					
Locally[b]	19	21	**+2**	14	12	**-2**	9	16	**+7**	90	88	**-2**	90	91	**+2**	100	67	**-33**	43	42	**-2**
Morogoro	16	21	**+5**	22	22	**0**	0	0	**0**	2	2	**0**	0	0	**0**	0	0	**0**	12	13	**+1**
Dodoma	47	50	**+3**	32	32	**0**	86	80	**-6**	6	10	**+4**	8	9	**+1**	0	33	**+33**	27	29	**+2**
Dar es Salaam	6	0	**-6**	20	20	**0**	0	0	**0**	2	0	**-2**	3	0	**-3**	0	0	**0**	11	10	**-1**
Elsewhere in Tanzania	9	8	**-1**	10	11	**+1**	6	4	**-2**	0	0	**0**	0	0	**0**	0	0	**0**	6	6	**0**
Foreign	3	0	**-3**	3	3	**0**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**	2	1	**0**
% of retailers who received delivery[a]	n/a	35	**n/a**	n/a	42	**n/a**	n/a	7	**n/a**	n/a	69	**n/a**	n/a	73	**n/a**	n/a	100	**n/a**	n/a	50	**n/a**

[a] Only for retailers selling branded maize flour. [b] “Local Towns” and “Locally” include the secondary cities for the secondary city retailers.

Table 3.5 Value chain finance patterns in 2012 and 2017

	Small Town			Medium Town						Secondary City											
	Dukas			Dukas			Minis			Dukas			Minis			Supers			Overall		
	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change	2012	2017	Change
% of procurement transactions paid with …																					
Cash	95	92	**-3**	97	96	**-1**	96	93	**-3**	100	100	**0**	97	98	**0**	100	100	**0**	98	97	**-1**
Mobile Money	5	8	**+3**	2	3	**+1**	4	7	**+3**	0	0	**0**	2	2	**0**	0	0	**0**	2	3	**+1**
Bank	0	0	**0**	1.1	1.1	**0**	0	0	**0**	0	0	**0**	1	1	**0**	0	0	**0**	1	1	**0**
% of retailers receiving supplier credit[a]	n/a	12	**n/a**	n/a	18	**n/a**	n/a	20	**n/a**	n/a	22	**n/a**	n/a	42	**n/a**	n/a	33	**n/a**	n/a	22	**n/a**
% of maize flour customers who pay on the spot	76	84	**+8**	82	80	**-2**	75	83	**+8**	82	80	**-2**	80	85	**+5**	100	100	**0**	81	81	**0**
% of customers who receive credit	24	16	**-8**	18	20	**+2**	25	17	**-8**	18	20	**+2**	20	15	**-5**	0	0	**0**	19	19	**0**

[a] Only for retailers selling branded maize flour.

Table 3.6 Average number of SKUs carried, by bag size

	Small Town	Medium Town		Secondary City			
	Dukas	**Dukas**	**Minis**	**Dukas**	**Minis**	**Supers**	**Overall**
N	**23**	**112**	**15**	**55**	**29**	**3**	**234**
Average # of loose SKUs	1	1.1	1.1	1.1	1.3	0	1.1
5kg	0	0.1	0.1	0.4	1.2	1	0.3
10kg	0	0	0.1	0.2	0.8	0.7	0.2
25kg	0.8	0.9	1.1	1.1	1.2	1	1
50kg	0	0.009	0	0	0	0	0
Conditional on carrying SKUs of that row:							
Average # of loose SKUs	1	1.1	1.1	1.2	1.4	0	1.1
5kg	n/a	1.1	1	1.1	1.3	1	1.2
10kg	n/a	1	1	1.1	1.3	2	1.2
25kg	1.1	1.1	1.1	1.2	1.3	3	1.1
50kg	n/a	1	n/a	n/a	n/a	n/a	1

Table 3.7 Branding practices

	Small Town	Medium Town		Secondary City			
	Dukas	**Dukas**	**Minis**	**Dukas**	**Minis**	**Supers**	**Overall**
N	**23**	**112**	**15**	**55**	**26**	**3**	**234**
1. Shares (of retailers) by form of maize flour sold in 2012							
Share selling ONLY LOOSE	74%	75%	0%	60%	0%	0%	57%
Share selling ONLY PACKED/BRANDED	13%	1%	0%	2%	15%	67%	5%
Share selling BOTH loose and packed/branded	4%	14%	0%	33%	4%	0%	15%
Not in business in 2012	9%	10%	100%	5%	81%	33%	23%
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
2. Shares (of retailers) by form of maize flour sold in 2017 (change from 2012 in parentheses)							
Share selling ONLY LOOSE	26% **(-48%)**	14% **(-61%)**	0% **(n/a)**	7% **(-53%)**	0% **(0%)**	0% **(0%)**	11% **(-46%)**
Share selling ONLY PACKED/BRANDED	0% **(-13%)**	0% **(-1%)**	0% **(n/a)**	2% **(0%)**	12% **(-3%)**	100% **(+33%)**	2% **(-3%)**
Share selling BOTH loose and packed/branded	74% **(+70%)**	86% **(+72%)**	100% **(n/a)**	91% **(58%)**	88% **(+84%)**	0% **(0%)**	87% **(+72%)**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
Branding trends (2017)							
Average number of brands of maize flour carried (conditional on carrying branded flour)	1.1	1.1	1.1	1.3	1.6	3	1.3
Share carrying at least one Dar es Salaam brand	0%	5%	0%	0%	0%	33%	3%
Share carrying at least one secondary city brand	88%	84%	80%	100%	100%	67%	90%
Share carrying at least one small/medium town brand	12%	10%	27%	2%	4%	0%	9%
Conditional on carrying a brand of that row:							
Average number of Dar brands	n/a	1	n/a	n/a	n/a	1	1
Average number of secondary city brands	1.1	1.1	1	1.3	1.5	3.4	1.3
Average number of small/medium town brands	1	1.2	1	1	1	n/a	1.1

Table 3.8 Margin analysis: average sales price in T.Sh., average margin, and margin's share of buying price (in bold)

	Small Town	Medium Town		Secondary City			
	Dukas	**Dukas**	**Minis**	**Dukas**	**Minis**	**Supers**	**Overall**
5kg bag		*Sales Price:* 5,171	5,995	5,081	5,369	5,593	5,219
	n/a	*Margin:* 643	500	655	788	833	653
		Share of buying price: **14.2%**	**9.1%**	**14.8%**	**17.2%**	**17.5%**	**14.3%**
10kg bag		11,115		9,560	10,446	10,009	10,277
	n/a	1,500	n/a	667	1,433	1,000	1,150
		15.6%		**7.5%**	**15.9%**	**11.1%**	**12.6%**
25kg bag	*Sales Price:* 26,727	26,239	29,288	22,972	23,783	19,401	23,606
	Margin: 4,000	4,500	7,300	3,619	3,048	1,667	3,769
	Share of buying price: **17.6%**	**20.7%**	**33.2%**	**18.7%**	**14.7%**	**9.4%**	**19%**

Note: Margin is the difference between selling and buying prices in T.Sh. Only packed/branded SKUs. 1 USD ~ 2,250 Tanzanian shillings. Source: https://xe.com/currencycharts/?from=USD&to=TZS&view=5Y

Table 3.9 Average partial effects from fractional regression (for share variables) and probit regressions (for 0/1 variables)

	***Panel A*: Fractional Regression Results (average partial effects)**			***Panel B*: Probit Results (average partial effects)**			
Dependent variable:	**Share from processor**	**Share milled in secondary city**	**Share sold packaged/ branded**	**=1 if regular relationship**	**=1 if receives delivery**	**=1 if supplier credit**	**=1 if customer credit**
Retail Type (omitted: duka)							
=1 if supermarket	0.234	-0.195*	0.892***	-0.067	n/a	0.079	n/a
	(0.208)	(0.101)	(0.013)	(0.300)		(0.267)	
=1 if mini-supermarket	-0.061	0.208***	0.907***	-0.039	-0.132	0.126	-0.147
	(0.054)	(0.069)	(0.013)	(0.093)	(0.087)	(0.085)	(0.091)
Town Size (omitted: small town)							
=1 if medium town	-0.236***	0.012	0.023	-0.079	-0.003	n/a	0.220**
	(0.074)	(0.065)	(0.038)	(0.127)	(0.123)		(0.100)
=1 if secondary city	0.331***	0.421***	0.222***	0.169	0.341**	0.056	0.121
	(0.119)	(0.047)	(0.085)	(0.144)	(0.147)	(0.137)	(0.102)
=1 if >25 km from the highway	0.048	0.018	0.001	-0.009	0.048		0.002
	(0.096)	(0.055)	(0.031)	(0.122)	(0.109)		(0.103)
Includes retailer demographics and asset ownership variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	232	232	181	206	204	202	230

Note: standard errors in parentheses. Fractional regression uses probit link. Retailer demographics include age, sex, and education of owner, and the percentage of household income that comes from the business. Asset ownership dummies include truck/bus, motorcycle, bicycle, and wheelbarrow. The combined share of maize, wheat, oil, and rice in total food space is also included but not shown. *** p<0.01, ** p<0.05, * p<0.1

Table 3.10 Margins analysis: average partial effects from fractional regression (only on branded SKUs)

Dependent variable:	**Margin as share of buying price**
Bag Size (omitted: 5kg)	
=1 if 10kg bag	-0.033*
	(0.018)
=1 if 25kg bag	0.024
	(0.023)
Brand Type (omitted: local small/medium town brand)	
=1 if regional brand (from local secondary city)	0.083**
	(0.032)
=1 if national brand (from Dar or other secondary city)	0.124**
	(0.054)
=1 if flour is fortified with micronutrients	0.007
	(0.025)
=1 if health claim on package	0.005
	(0.024)
Retail Type (omitted: duka)	
=1 if supermarket	0.019
	(0.059)
=1 if mini-supermarket	0.026
	(0.022)
Town Size (omitted: small town)	
=1 if medium town	0.083
	(0.088)
=1 if secondary city	-0.053
	(0.050)
=1 if >25 km from the highway	0.069
	(0.053)
Includes retailer demographics and asset ownership variables	Yes
N	159

Note: standard errors in parentheses. Retailer demographics include age, sex, and education of owner, and the percentage of household income that comes from the business. Asset ownership dummies include truck/bus, motorcycle, bicycle, and wheelbarrow. The combined share of maize, wheat, oil, and rice in total food space is also included but not shown. *** p<0.01, ** p<0.05, * p<0.1

Appendix B: Supplemental Tables

Table 3.11 OLS regression results

	"Share" Variables				0/1 Variables			
Dependent variable:	**From processors**	**From secondary city**	**Packed/branded flour**	**Supply transactions using mobile money**	**Regular relationship**	**Receives delivery**	**Supplier credit**	**Customer credit**
Retail Type (omitted: duka)								
=1 if supermarket	0.267	-0.045	0.968***	0.009	-0.064	0.067	0.081	-0.690***
	(0.253)	(0.084)	(0.072)	(0.014)	(0.282)	(0.107)	(0.272)	(0.104)
=1 if mini-supermarket	-0.077	0.121**	0.881***	0.032*	-0.048	-0.135	0.133	-0.147
	(0.073)	(0.056)	(0.049)	(0.017)	(0.090)	(0.085)	(0.089)	(0.090)
Town Size (omitted: small town)								
=1 if medium town	-0.216**	-0.237*	0.015	-0.029	-0.144	0.010	-0.066	0.151
	(0.108)	(0.133)	(0.026)	(0.042)	(0.157)	(0.152)	(0.121)	(0.134)
=1 if secondary city	0.387***	0.329***	0.187***	-0.050	0.157	0.342**	0.064	0.140
	(0.123)	(0.116)	(0.049)	(0.039)	(0.153)	(0.152)	(0.124)	(0.135)
=1 if >25 km from the highway	-0.047	-0.359*	0.005	-0.053	-0.112	0.061	-0.179	-0.122
	(0.172)	(0.190)	(0.027)	(0.040)	(0.274)	(0.265)	(0.119)	(0.217)
Medium town * >25 km interaction term	0.185	0.527**	0.023	0.032	0.178	-0.026	0.263*	0.214
	(0.178)	(0.212)	(0.038)	(0.043)	(0.292)	(0.283)	(0.141)	(0.231)
Constant	0.406*	0.477**	0.251*	0.069	0.584**	-0.100	0.287	0.855***
	(0.239)	(0.230)	(0.138)	(0.051)	(0.290)	(0.262)	(0.253)	(0.250)
Includes retailer demographics and asset ownership variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	232	232	181	232	206	207	207	232

Note: robust standard errors in parentheses. Retailer demographics include age, sex, and education of owner, and the percentage of household income that comes from the business. Asset ownership dummies include truck/bus, motorcycle, bicycle, and wheelbarrow. The combined share of maize, wheat, oil, and rice in total food space is also included but not shown. *** p<0.01, ** p<0.05, * p<0.1

REFERENCES

REFERENCES

Ailawadi, K.L., Lehmann D.R., & Neslin, S.A. (2003). Revenue premium as an outcome measure of brand equity. *Journal of Marketing*, *67*(4), 1-17.

Assefa, T., Abebe, G., Lamoot, I., & Minten, B. (2016). Urban food retailing and food prices in Africa: the case of Addis Ababa, Ethiopia. *Journal of Agribusiness in Developing and Emerging Economies*, *6*(2), 90-109.

Bennett, M.K. (1941). International contrasts in food consumption. *Geographical Review*, *31*(3), 365-376.

Chaboud, G. & Moustier, P. (2021). The role of diverse distribution channels in reducing food loss and waste: The case of the Cali tomato supply chain in Colombia. *Food Policy*, *98*, 1-12.

Christiaensen, L., De Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: The role of rural diversification and secondary towns. *Agricultural Economics*, *44*, 435-447.

Christiaensen, L. & Todo, Y. (2014). Poverty reduction during the rural-urban transformation - The role of the missing middle. *World Development*, *63*, 43-58.

Dannhaeuser, N. (1977). Distribution and the structure of retail trade in a Philippine commercial town setting. *Economic Development and Cultural Change*, *25*(3), 471-503.

Dannhaeuser, N. (1980). The role of the neighborhood store in developing economies: The case of Dagupan City, Philippines. *The Journal of Developing Areas*, *14*(2), 157-174.

Dannhaeuser, N. (1984). Market penetration of upcountry India: Nasik in comparative perspective. *Economic and Political Weekly*, *19*(34), M90-M98.

Dorta, M. (2016). Introduction to fractional outcome regression models using the fracreg and betareg commands. StataCorp LP: Aguas Calientes, Mexico.

Erdem, T., & Swait, J. (1998). Brand equity as a signaling phenomenon. *Journal of Consumer Psychology*, *7*(2), 131-157.

Goldfarb, A., Lu, Q., & Moorthy, S. (2009). Measuring brand value in an equilibrium framework. *Marketing Science*, *28*(1), 69-86.

Liverpool-Tasie, S., Omonona, B., Ogunleye, W., Abagyeh, I., & Reardon, T. (2016). The presence of processed foods in Sub-Saharan Africa – where from and where to? Evidence from Nigeria. Working paper. East Lansing: Michigan State University.

Maertens, M. & Swinnen, J. (2009). Trade, standards, and poverty: Evidence from Senegal. *World Development*, *37*(1), 161-178.

Maertens, M., Minten, B., & Swinnen, J. (2012). Modern food supply chains and development: Evidence from horticulture export sectors in sub-Saharan Africa. *Development Policy Review*, *30*(4), 473-497.

Minten, B., Murshid, K., & Reardon, T. (2013a). Food quality changes and implications: Evidence from the Rice Value Chain of Bangladesh. *World Development*, *42*(1), 100-113.

Minten, B., Singh, K.M., & Sutradhar, R. (2013b). Branding and agricultural value chains in developing countries: Insights from Bihar (India). *Food Policy*, *38*(1), 23-34.

Minten, B., Tamru, S., Engida, E., & Kuma, T. (2015). Transforming staple food value chains in Africa: The case of teff in Ethiopia. *The Journal of Development Studies*, *52*(5), 627-645.

Papke, L.E. & Wooldridge, J.M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*, 619-632.

Reardon, T. & Gulati, A. (2008). The rise of supermarkets and their development implications: International experience relevant for India. IFPRI Discussion Paper No. 00752. International Food Policy Research Institute: New Delhi, India.

Reardon, T., Chen, K., Minten, B., & Adriano, L. (2012). The quiet revolution in staple food value chains: Enter the dragon, the elephant and the tiger. Asian Development Bank and International Food Policy Research Institute: Philippines.

Reardon, T. & Timmer, C. (2014). Five inter-linked transformations in the Asian agrifood economy: Food security implications. *Global Food Security*, *3*, 108-117.

Reardon, T., Echeverria, R., Berdegué, J., Minten, B., Liverpool-Tasie, S., Tschirley, D., & Zilberman, D. (2019). Rapid transformation of food systems in developing regions: Highlighting the role of agricultural research and innovations. *Agricultural Systems*, *172*, 47-59.

Sauer, C., Reardon, T., Tschirley, D., Liverpool-Tasie, S., Awokuse, T., Alphonce, R., Ndyetabula, D.W., & Waized, B. (2021). Consumption of processed food and food away from home in big cities, small towns, and rural areas of Tanzania. *Agricultural Economics*, 1- 22. DOI: 10.1111/agec.12652

Snyder, J., Ijumba, C., Tschirley, D., & Reardon, T. (2015). Local response to the rapid rise in demand for processed and perishable foods: Results of an inventory of processed food products in Dar es Salaam. Innovation Lab for Food Security, Tanzania Policy Research Brief No. 2. East Lansing: Michigan State University.

Soullier, G. & Moustier, P. (2020). The modernization of the rice value chain in Senegal: a move towards the Asian quiet revolution. *Development Policy Review*. https://doi.org/10.1111/dpr.12459

Theriault, V., Vroegindewey, R., Assima, A., & Keita, N. (2018). Retailing of processed dairy and grain products in Mali: Evidence from a city retail outlet inventory. *Urban Science, 2*(1), 24.

Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015). The rise of a middle class in Eastern and Southern Africa: Implications for food system transformation. *Journal of International Development*, *27*, 628-646.

Wooldridge, J.M. (2002). Econometric analysis of cross section and panel data. MIT Press: Cambridge, Massachusetts.

World Bank. (2019). Tanzania Overview. Retrieved from http://www.worldbank.org/en/country/tanzania/overview on May 9, 2019.