

**FOOD CONSUMPTION PATTERNS IN LIGHT OF RISING INCOMES,
URBANIZATION AND FOOD RETAIL MODERNIZATION:
EVIDENCE FROM EASTERN AND SOUTHERN AFRICA**

By

Michael James Dolislager

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

2017

ABSTRACT

FOOD CONSUMPTION PATTERNS IN LIGHT OF RISING INCOMES, URBANIZATION AND FOOD RETAIL MODERNIZATION: EVIDENCE FROM EASTERN AND SOUTHERN AFRICA

By

Michael James Dolislager

Over the past fifteen years, Eastern and Southern Africa (ESA) has experienced significant changes in its agrifood system. Driven by rapid urbanization and growth in per capita incomes, changing food consumption patterns have been at the core of this broader agrifood system change. Responding to the region's diet transformation, firms – including a limited presence of supermarkets – have arisen in food marketing and distribution, food processing, and food services.

Food consumption patterns are changing in three ways: food is becoming more purchased, more perishable, and more processed. These changes drive a multitude of effects on the agrifood system. This dissertation provides new insight into the nature of the diet transformation that is unfolding in the region, and on some of its drivers and effects. The first essay, *“Diversification “Beyond Staple Foods” in the Diets of Poor Rural and Urban Consumers in Developing Eastern and Southern Africa”*, focuses on the effects of income and urbanization on the commodity makeup and source of foods in household diets. Transitioning food consumption patterns are not solely a middle-class story, as conventionally assumed – in fact, poor households are already consuming surprisingly high levels of purchased food and are consuming greater shares of non-grain foods as their incomes rise. Spatial considerations of increasing city size and reduced distance to cities also have significant positive effects on

purchased share, affirming the expectation that households will purchase more food (compared to consuming own production) when they have increased access to markets.

Essay two, *“Consumption of Processed Food in the Diets of Poor Rural and Urban Consumers in Developing Eastern and Southern Africa”*, further analyzes the same drivers of food consumption patterns on the consumption of processed food. Processed foods, specifically highly processed foods, have penetrated the diets of rural and urban households at all levels across the income distribution of ESA. The income-induced diet change towards processed food begins among the poor regardless of household distance from urban areas, albeit at greater shares among households within larger cities. The patterns of increased processed share with income growth and urbanicity signal a strong future demand for increased food market infrastructure.

The final essay, *“City size, supermarkets, and processed foods: Evidence from Zambia”*, incorporates food retail modernization into the analysis of processed food consumption patterns. Utilizing a unique dataset that contains disaggregated household consumption data, including source and distance to source of each consumed food item, this analysis shows the following: First, conditional on the presence of a supermarket, households in smaller cities consume greater shares of food, by value, from supermarkets than households in larger cities. Second, although supermarket use positively affects the consumption of processed food, households that consume some food from supermarkets continue to purchase the majority of their processed food from non-supermarket retailers.

Together, these three essays provide detail to how the macro trends of income growth, rapid urbanization and the beginning of food retail modernization significantly affect the consumption patterns of ESA, which in turn affect ESA’s entire agrifood system.

Copyright by
MICHAEL JAMES DOLISLAGER
2017

ACKNOWLEDGEMENTS

I wish to express my gratitude for the mentorship and guidance that I have received from Dr. David Tschirley, who has been my major advisor and has funded my research at Michigan State University. Thank you for the many hours you spent working with me throughout these past years. This would not have been possible without your support. I would also like to thank the other members of my doctoral committee, Drs. Thomas Reardon, Nicole Mason, Steven Haggblade and Robert Richardson, for their invaluable contributions to this research. Thank you for the walks around the track, for the meetings and calls working through analysis, and for the many comments and emails that have occurred to bring this research to where it is today.

I also want to extend a big thanks to my friends and family. Thank you to the many of you who have been with me through this process. Special thanks to Helder Zavale, for the hours you spent helping me with empirical analysis early in my graduate experience, and Chew Nkonde, for all your encouragement as we made our way through graduate school. I am especially grateful to you both for your friendship. To my parents, James and Carla Dolislager, thank you for your prayers, encouragement and support during this time and throughout my entire life. And last, but not least, a thanks to my wife, Elyse Dolislager: thank you for your love and patience through these past few years, you mean more to me than words can express.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	xiii
KEY TO ABBREVIATIONS	xv
INTRODUCTION	1
REFERENCES	5
ESSAY 1: DIVERSIFICATION “BEYOND STAPLE FOODS” IN THE DIETS OF POOR RURAL AND URBAN CONSUMERS IN DEVELOPING EASTERN AND SOUTHERN AFRICA	7
1.1 Introduction.....	7
1.2 Conceptual Framework	10
1.3 Methods.....	14
1.3.1 Context.....	14
1.3.2 Data	18
1.3.3 Definitions.....	19
1.3.4 General Model	21
1.3.5 Empirical Model	22
1.3.6 Estimation Methods	28
1.4 Results.....	31
1.4.1 Descriptive Results	32
1.4.2 Nonparametric analysis of the relationship between expenditure and food consumption.....	34
1.4.3 Expenditure Threshold Analysis.....	40
1.4.4 Engel’s Curve Model regression analysis.....	44
1.5 Conclusions.....	57
APPENDIX.....	60
REFERENCES	73
ESSAY 2: CONSUMPTION OF PROCESSED FOOD IN THE DIETS OF POOR RURAL AND URBAN CONSUMERS IN DEVELOPING EASTERN AND SOUTHERN AFRICA	77
2.1 Introduction.....	77
2.2 Conceptual Framework	78
2.3 Methods.....	84
2.3.1 Context.....	84
2.3.2 Data	87
2.3.3 Definitions.....	89
2.3.4 General Model	90
2.3.5 Empirical Model	91

2.3.6	Estimation Methods	96
2.4	Results.....	99
2.4.1	Descriptive statistics and nonparametric analysis of the relationship between processed food consumption and total household expenditure	99
2.4.2	Expenditure Threshold Analysis.....	104
2.4.3	Descriptive statistics of processed food consumption by city size	108
2.4.4	Nonparametric analysis of the relationship between distance to cities and processed food consumption.....	110
2.4.5	Engel's Curve Model regression analysis.....	111
2.4.6	Mediation model analysis	118
2.5	Conclusions.....	120
	APPENDIX.....	123
	REFERENCES	138
ESSAY 3: CITY SIZE, SUPERMARKETS, AND PROCESSED FOODS: EVIDENCE FROM ZAMBIA.....		
3.1	Introduction.....	143
3.2	Conceptual Framework.....	146
3.3	Methods.....	154
3.3.1	Data	154
3.3.2	Definitions.....	157
3.3.3	General Model	158
3.3.4	Empirical Model	158
3.3.5	Estimation Methods	167
3.4	Results.....	172
3.4.1	Descriptive Statistics.....	172
3.4.2	Regression Results	177
3.5	Conclusions.....	186
	APPENDIX.....	189
	REFERENCES	205
CONCLUSIONS.....		209

LIST OF TABLES

Table 1-1: National expenditure and population by expenditure strata and city size	17
Table 1-2: Household descriptive characteristics	19
Table 1-3: Dependent Variables	20
Table 1-4: Determinants of demand in the Engel's Curve Model	24
Table 1-5: Food budget shares of key food consumption patterns, aggregated country and settlement	32
Table 1-6: Food budget shares of select commodity aggregates, pooled data by settlement	34
Table 1-7: Expenditure thresholds and slopes prior to and after expenditure thresholds in the consumption patterns of purchased share and Beyond Staple Foods items – rural data	40
Table 1-8: Expenditure thresholds and slopes prior to and after expenditure thresholds in the consumption patterns of purchased share and Beyond Staple Foods items – urban data	41
Table 1-9: Expenditure thresholds and slopes prior to and after expenditure thresholds for the aggregates of select commodity based aggregates – rural data	42
Table 1-10: Expenditure thresholds and slopes prior to and after expenditure thresholds for the aggregates of select commodity based aggregates – urban data	43
Table 1-11: Average partial effects of expenditure for key food consumption patterns – rural data	45
Table 1-12: Average partial effects of expenditure and city size for key food consumption patterns – urban data	46
Table 1-13: Indirect effects of urban household settlement within primary and secondary cities, estimates by country	50
Table 1-14: Average partial effects of expenditure and distances to various city sizes for key food consumption patterns – rural Zambia	55
Table 1-15: Indirect effects of distance from rural household settlement to primary, secondary or tertiary cities, estimates by country	57
Table 1-A-1: Food budget shares of select commodity aggregates, country data by settlement	61

Table 1-A-2: Average partial effects of the household determinants for key food consumption patterns, urban household estimates by country	63
Table 1-A-3: Average partial effects of the household determinants for key food consumption patterns, rural household estimates by country	64
Table 1-A-4: Average partial effects of the household determinants for select commodity aggregates, urban household estimates by country	65
Table 1-A-5: Average partial effects of the household determinants for select commodity aggregates, rural household estimates by country	69
Table 2-1: Impact of city size on food choice: urban environment, food environment, and demand factors	83
Table 2-2: National expenditure and population by expenditure strata and city size	86
Table 2-3: Household descriptive characteristics	87
Table 2-4: Processed food aggregates.....	90
Table 2-5: Determinants of demand in the Engel's Curve Model	93
Table 2-6: Food budget shares of processed food, urban households by country and expenditure	100
Table 2-7: Food budget shares of processed food, rural households by country and expenditure	100
Table 2-8: Expenditure thresholds and slopes prior to and after expenditure thresholds in processed food consumption patterns	105
Table 2-9: Expenditure thresholds and slopes prior to and after expenditure thresholds for processed food aggregates – urban data.....	106
Table 2-10: Expenditure thresholds and slopes prior to and after expenditure thresholds for processed food aggregates – rural data	107
Table 2-11: Food budget shares of processed food, aggregated country and settlement	109
Table 2-12: Food budget shares of processed food aggregates, pooled data by settlement ..	109
Table 2-13: Average partial effects of expenditure and city size on processed share – urban data	112
Table 2-14: Average partial effects of expenditure and distances to various city sizes on processed share – rural data	113

Table 2-15: Indirect effects of urban household settlement within primary and secondary cities on processed share, estimates by country	119
Table 2-16: Indirect effects of distance from rural household settlement to primary, secondary or tertiary cities on processed share – rural Zambia.....	120
Table 2-A-1: Relationship between food classification scheme in this paper and that in Monteiro et al. (2010)	124
Table 2-A-2: Food budget shares of processed food aggregates, by country and expenditure	124
Table 2-A-3: Food budget shares of processed food aggregates, country data by settlement	128
Table 2-A-4: Average partial effects of the household determinants on processed share, urban household estimates by country	129
Table 2-A-5: Average partial effects of the household determinants on processed share, rural household estimates by country	130
Table 2-A-6: Average partial effects of the household determinants for processed food aggregates, urban household estimates by country	131
Table 2-A-7: Average partial effects of the household determinants for processed food aggregates, rural household estimates by country	135
Table 3-1: Impact of city size on food choices: urban environment, food environment, and demand factors	153
Table 3-2: City populations	155
Table 3-3: Total daily household expenditure per adult equivalent	156
Table 3-4: Kilometers to supermarket relative to closest non-supermarket food retailer to the household	156
Table 3-5: Shares of market value by acquisition type.....	157
Table 3-6: Dependent variables	161
Table 3-7: Key for determinants of demand in equation 3.1	161
Table 3-8: Household characteristics of the sample	167
Table 3-9: Household weighted averages of consumption shares	173
Table 3-10: Household weighted averages of sources of food aggregates	175

Table 3-11: Household weighted averages of consumption shares by low and high processing levels	176
Table 3-12: Fractional probit estimates of the average partial effects of the determinants of household share of supermarket purchases in total food consumption.....	178
Table 3-13: Second stage of control function regression; estimating the share of processed food in total food consumption	181
Table 3-14: City indirect effects via specific explanatory variables.....	185
Table 3-A-1: Prices of commonly consumed items.....	190
Table 3-A-2: First stage of control function regression; OLS estimates of the determinants of the share of supermarket purchases in total food consumption	191
Table 3-A-3: Second stage of control function regression; estimating the share of low processed food in total food consumption	192
Table 3-A-4: Second stage of control function regression; estimating the share of high processed food in total food consumption	193
Table 3-A-5: Estimates of the marginal effects of the determinants of processed food in total food consumption (OLS)	194
Table 3-A-6: Estimates of the marginal effects of the determinants of low processed food in total food consumption (OLS)	195
Table 3-A-7: Estimates of the marginal effects of the determinants of high processed food in total food consumption (2SLS).....	196
Table 3-A-8: Second stage of two stage least squares; estimating the share of processed food in total food consumption (2SLS).....	197
Table 3-A-9: Second stage of two stage least squares; estimating the share of low processed food in total food consumption (2SLS).....	198
Table 3-A-10: Second stage of two stage least squares; estimating the share of high processed food in total food consumption (2SLS)	199
Table 3-A-11: Fractional probit estimates of the determinants of household share of supermarket purchases in total food consumption (CRE)	200
Table 3-A-12: First stage of control function regression; OLS estimates of the determinants of the share of supermarket purchases in total food consumption (CRE).....	201
Table 3-A-13: Second stage of control function regression; estimating the share of processed food in total food consumption (CRE).....	202

Table 3-A-14: Second stage of control function regression; estimating the share of low processed food in total food consumption (CRE)..... 203

Table 3-A-15: Second stage of control function regression; estimating the share of high processed food in total food consumption (CRE)..... 204

LIST OF FIGURES

Figure 1-1: Conceptual framework	12
Figure 1-2: Nonparametric (LOWESS curve) analysis of select food consumption patterns – rural data.....	35
Figure 1-3: Nonparametric (LOWESS curve) analysis of select food consumption patterns – urban data.....	36
Figure 1-4: Nonparametric (LOWESS curve) analysis of select commodity aggregates – pooled rural data	38
Figure 1-5: Nonparametric (LOWESS curve) analysis of select commodity aggregates – pooled urban data.....	39
Figure 1-6: Estimated effects of expenditure on purchased share and the consumption of Beyond Staple Foods items – rural data	46
Figure 1-7: Estimated effects of expenditure on purchased share and the consumption of Beyond Staple Foods items – urban data	47
Figure 1-8: Locally weighted scatterplot smoothing curves representing the typical purchased share relative to household distance to city	53
Figure 1-9: Locally weighted scatterplot smoothing curves representing the typical food budget shares of the consumption Beyond Staple Foods items relative to household distance to city	54
Figure 2-1: Conceptual framework	80
Figure 2-2: Nonparametric (LOWESS curve) analysis of consumption patterns of processed food	101
Figure 2-3: Nonparametric (LOWESS curve) analysis of processed food aggregates – pooled data.....	104
Figure 2-4: Locally weighted scatterplot smoothing curves representing the typical processed share relative to household distance to city	111
Figure 2-5: Estimated effects of expenditure on purchased share and the consumption of processed food	114
Figure 3-1: Conceptual approach linking city size to consumer food choices	147

Figure 3-2: Estimation of indirect effects	172
Figure 3-3: Estimated effects of expenditure on supermarket use and processed food consumption.....	179
Figure 3-A-1: Processing level aggregates compared with aggregates from Monteiro et al. 2010	190

KEY TO ABBREVIATIONS

2SLS	Two Stage Least Squares
AE	Adult Equivalent
BSF	Beyond Staple Foods
CRE	Correlated Random Effects
ESA	Eastern & Southern Africa
FDI	Foreign Direct Investment
IV	Instrument Variable
LOWESS	Locally Weighted Scatterplot Smoothing
OLS	Ordinary Least Squares
PPP	Purchasing Power Parity
SF	Staple Foods
UCS	2007/2008 Zambia Urban Consumption Survey
FP	Fractional Probit

INTRODUCTION

Over the past fifteen years Eastern and Southern Africa (ESA)¹ has experienced significant changes in its agrifood system. Driven by rapid urbanization and growth in per capita incomes, changing food consumption patterns – amounting to a diet transformation - have been at the core of this broader agrifood system change. Average daily expenditures per adult equivalent in ESA remain low, near three dollars², yet World Bank data show that countries in this region have achieved real growth of per capita income³ averaging near three and a half percent per annum since 2000. Urban population since this time has risen from 21% of total population to over one quarter currently⁴, and the United Nations Population Division anticipates that forty-four percent of the population will live in urban areas by 2050. Responding to the region's diet transformation, 10s of thousands of micro, small, and medium size firms have arisen in food marketing and distribution, food processing, and food services including prepared food away from home. Lastly, this region is experiencing food retail modernization as part of the final wave of the supermarket revolution. Although supermarket presence in Sub-Saharan Africa remains the lowest of any major region of the world, rapid increases have been seen in recent years in major urban areas of ESA, driven by rising incomes, urbanization and market liberalization that drove an influx of domestic investment and foreign direct investment.

¹ Eastern and Southern Africa is referring to the following developing Eastern and Southern African nations: Ethiopia, Botswana, Kenya, Malawi, Mozambique, Tanzania, Uganda, and Zambia

² Author's calculation from LSMS survey data.

³ Author's calculation using GNI per capita, PPP (constant international dollars).

⁴ Author's calculations using United Nations Population Division data. Urban share of population calculated at 21% in the year 2000.

Diets are changing in three ways: food is becoming more purchased, more perishable, and more processed. These changes drive a multitude of effects on the agrifood system. Consumer demand for food incentivizes the production of food; therefore household food demand affects the production decisions of farmers and potentially affect the policy actions of governments that support certain production practices. Increased purchasing of food through markets drives private investment in those markets, and requires significant public investment in marketing infrastructure (especially wholesale markets but also retail) and in complementary infrastructure such as roads, electrical networks, and plumbing. Consumer demand for more perishable foods – fresh produce, meats, and milk products – drives a need for investment in cold chains and changing marketing practices to handle such foods. Increased demand for processed food presents enormous opportunities for investment by local entrepreneurs to satisfy this demand, at the same time that it drives increased imports of processed foods, directly competing with local businesses.

This dissertation provides new insight into the nature of the diet transformation that is unfolding in the region, and on some of its drivers and effects. The first essay, *“Diversification Beyond Staple Foods” in the Diets of Poor Rural and Urban Consumers in Developing Eastern and Southern Africa*, focuses on the effects of income and urbanization on the commodity makeup and source of foods in household diets. Bennett’s Law (1941) states that, as incomes increase, households will consume greater shares of non-grain foods, establishing a relationship between income and food consumption patterns. The concept of supernumerary income (Stone, 1954) indicates that households allocate initial income toward required expenditures and that income above a certain threshold is allocated at the household’s discretion. Combining Bennett’s Law and the concept of supernumerary income suggests that food consumption patterns will

change as incomes rise above a level required to meet basic household needs. Yet simple calculations on average household food consumption patterns potentially appear to show diet change happening well below what could be considered supernumerary levels of income. Using nonparametric and parametric analyses, essay one seeks to understand these patterns by testing for changes in the direction and magnitude of the relationships between income and food consumption patterns of specific commodities and channels of food acquisition.

The rapid urbanization of ESA highlights the importance of understanding the effects of urban residence, city size and proximity to cities on food consumption patterns. Urbanization has been shown to affect household lifestyle patterns (Huang and Bouis, 2001; Ferré et al., 2012; Fafchamps and Shilpi, 2003), and this essay considers such effects on commodity food budget shares and food purchase shares in ESA.

Essay two, *“Consumption of Processed Food in the Diets of Poor Rural and Urban Consumers in Developing Eastern and Southern Africa”*, further analyzes the same drivers of food consumption patterns on the consumption of processed food. Household consumption patterns of processed food are expected to vary much as the consumption of non-staple food varies across the income distribution and with increased urbanicity (de Haen et al., 2003; Pingali, 2007). Given such expectations, this essay analyzes consumption patterns with respect to processed food across the entire income distribution and explores the impact of city size and proximity to city on these patterns.

The final essay, *“City size, supermarkets, and processed foods: Evidence from Zambia”*, incorporates food retail modernization into the analysis of food consumption patterns. Supermarkets are a key component in food retail modernization, one that provides households with increased access to a greater diversity of types of foods, thereby affecting household food

consumption patterns (Rischke et al., 2015). To this author's knowledge, no previous study has unpacked the impact of urbanization on shopping and consumption patterns by examining how this impact varies by city size and by distance to cities. Utilizing a unique dataset containing disaggregated household consumption data that includes source and distance to source of each consumed food item, this essay tests for the effects of city size on supermarket use and consumption of processed food, recognizing the affect that increasing supermarket use could have on processed food consumption.

REFERENCES

REFERENCES

- Bennett, M. K. (1941). Wheat in national diets. *Wheat Studies*, (02).
- De Haen, H., Stamoulis, K., Shetty, P., & Pingali, P. (2003). The world food economy in the twenty-first century: challenges for international co-operation. *Development Policy Review*, 21(5-6), 683-696.
- Fafchamps, M., & Shilpi, F. (2003). The spatial division of labour in Nepal. *The Journal of Development Studies*, 39(6), 23-66.
- Ferré, C., Ferreira, F. H., & Lanjouw, P. (2012). Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, lhs007.
- Huang, J., & Bouis, H. (2001). Structural changes in the demand for food in Asia: empirical evidence from Taiwan. *Agricultural Economics*, 26(1), 57-69.
- Rischke, R., Kimenju, S. C., Klasen, S., & Qaim, M. (2015). Supermarkets and food consumption patterns: The case of small towns in Kenya. *Food Policy*, 52, 9-21.
- Pingali, P. (2007). Westernization of Asian diets and the transformation of food systems: implications for research and policy. *Food policy*, 32(3), 281-298.
- Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. *The Economic Journal*, 64(255), 511-527.

ESSAY 1: DIVERSIFICATION “BEYOND STAPLE FOODS” IN THE DIETS OF POOR RURAL AND URBAN CONSUMERS IN DEVELOPING EASTERN AND SOUTHERN AFRICA

1.1 Introduction

The past two decades of rapid per capita income growth and urbanization in many countries of developing Eastern & Southern Africa (ESA)⁵ raise a host of questions regarding related impact on consumer behavior. One question follows: has this rapid income growth had any meaningful impact on food consumption patterns of households within ESA? The answer matters, because what people eat and how the food is sourced carries profound implications for the structure of the agrifood system, the level and type of needed public and private investments, the level and distribution of employment, in addition to nutrition and health.

Bennett’s Law indicates that as incomes increase household diets will transition away from grains and towards what we refer to as “Beyond Staple Foods” (BSF), which include meat, dairy, fish, fresh produce, among other non-grain food products (Bennett, 1941). The observed trend of rising household incomes is therefore expected to have increased the consumption of BSF and if rising income continues the consumption of BSF is expected to continue to disproportionately rise. The extent to which income has and will affect food consumption patterns is conditional on the theory of supernumerary income, which suggests that households will alter their consumption patterns only after required expenses have been met as it is the discretionary income households use to acquire other utility generating goods (Stone, 1954). These theories combined indicate that households at income levels below a supernumerary level of income would consume primarily

⁵ Countries include: Botswana, Burundi, Ethiopia, Kenya, Lesotho, Malawi, Mozambique, Namibia, Rwanda, South Sudan, Swaziland, Tanzania, Uganda, Zambia, and Zimbabwe.

cheap grains that enable households to meet their basic food need for sufficient and reliable calories, and that at incomes above this level, households would devote resources to good-tasting food, novel food, and instrumental food⁶ (Satter, 2007). The empirically observed income level that corresponds with a supernumerary income matters due to the expectation that food consumption patterns will differ prior to and beyond attainment of this income level. Following the theory of supernumerary income, conventional wisdom tends to view significant diet change as a predominantly middle class (and higher) phenomenon, as the presence of discretionary income is commonly used to denote the point of delineation between poor and middle class households in ESA.

The food consumption patterns that impact the agrifood system are also affected by non-income determinants, such as opportunity costs of time. Household opportunity cost of time would vary by household income as well as participation in activities such as nonfarm employment, which is also shown to affect food consumption patterns (Senauer et al., 1986; Kennedy & Reardon, 1996; Gómez et al., 2013). Opportunity cost of time is further increased by household settlement in urban areas that offer additional employment or entrepreneurial opportunities and a generally faster pace of life (Bettencourt et al., 2007). The work of Huang & Bouis (2001) use data from Taiwan to highlight these patterns by showing that general changes in “tastes and lifestyles”, specifically in food consumption patterns, result from structural change such as the aforementioned engagement in nonfarm employment and settlement in urban areas. Haggblade et al. (2010) noted that households in Africa increasingly depend on nonfarm employment, additionally the 2014 revision of the World Urbanization Prospects by the United Nations cites significant growth in

⁶ Instrumental food is defined by Satter 2007 as food that contributes to desired physical, cognitive, or spiritual outcomes.

recent and the projected future urbanization rates for ESA, each highlighting growth in the drivers of food consumption patterns in ESA that were found to be significant in Taiwan.

Huang & Bouis (2001) reason that urban settlement would affect food consumption patterns by highlighting multiple differences between urban and rural settlements, including the increased availability of food in markets and the broader exposure to different cultures. These differences also exist to varying degrees across urban areas of varying population sizes, with cities of greater populations providing increased availability of food and broader exposure to different cultures. Thiele & Weiss (2003) have shown that city size (in terms of population) has a significant effect on the diet diversity of German households, although similar analyses have not been completed using data of low-income countries. City size has been shown to affect other lifestyle patterns of households within low-income countries such as income and poverty – where income is greater in larger cities, and medium and small cities lead to greater quantities of migrating households escaping from poverty (Ferré et al., 2012; Christiaensen et al., 2013; Berdegúe et al., 2015). The lifestyle patterns of rural households are also affected by urbanization, as many lifestyle patterns such as employment and agricultural production are affected by a household's distance to urban areas (Fafchamps & Shilpi, 2003; Sharma, 2016).

The effects of city size and distance to cities are evident pertaining to multiple lifestyle patterns, yet to our knowledge, their effects have not been estimated regarding food consumption patterns. We will estimate the effects of city size and distance to cities on the share of purchased food and the share of BSF in the total value of consumed food. We will also empirically test for the level at which these food consumption patterns vary across income strata, according to the joint application of Bennett's Law and the theory of supernumerary income.

The paper proceeds as follows: Section two presents the conceptual model used to articulate the decision process and hypotheses. Section three addresses the methodology, which includes the context of the study, description of data, definitions, general and empirical models, and estimation procedures. Section four reflects on the results and provides discussion. Lastly, section five provides conclusions.

1.2 Conceptual Framework

Food consumption patterns are based on the desire of households to maximize their own utility. Due to the heterogeneity of households, each household possesses a unique utility function that is based on the consumer demand characteristics of individual and household preferences for taste, energy density, and convenience; of social and individual norms and beliefs regarding informal institutions, beauty, and lifestyle; and of health concerns including foods' nutritional content and perceived safety. The maximization of these utility functions is conditional on being able to afford these goods within a household's income (m).

$$\text{Max } u(x), \text{ s.t. } \mathbf{p}x \leq m \quad (1.1)$$

$$\text{optimal bundle: } x(\mathbf{p}, m)$$

Food to be consumed can either be purchased from food markets or sourced from household's own production. The prices included in the household's utility maximization include market prices as well as shadow prices for the production and preparation of all consumed food.

Bennett's Law provides a relationship between household income and food consumption patterns. Stone's theory of supernumerary income further refines the expected relationship between income and consumption patterns, by suggesting that the bulk of income induced diet change would result from purchases by households with increasing levels of discretionary

income. Given that nearly two thirds of the population within ESA⁷ has incomes below the international poverty line⁸, this theory suggests that relatively little diet change would have resulted from the recent income growth. The conceptual framework will relax the assumption that diet change will primarily occur among the non-poor. To facilitate discussion I will coin a term, the subnumerary kink, which will refer to a change in the direction or speed of diet change that would occur at income levels below those that would have normally associated with increasing discretionary income. If I find a subnumerary kink in food consumption patterns, then I would want to know the specific empirical configuration of the pattern: at what income levels does the rise start?; does it build-up incrementally or in quantum steps? I would want also to understand its drivers; what factors create a situation in which poor and rural households are more likely than they previously were to change their consumption behaviors in this way. Put another way, what factors might be driving this changes at lower levels of income (whether a households resides in a rural- or urban area) than in the past? More formally, controlling for income, what non-income factors might be driving these observed patterns?

To incorporate key non-income drivers, Rischke et al. (2010) conceptualized a framework that combines consumer demand, as discussed above, with the food environment, which highlight supply-side factors such as access, availability, price and market-induced desirability of food. Access, not to be confused with economic access, refers to the cost to access foods, which is subject to the density of retailers and the household's time and expense of reaching markets. Availability identifies the types and variety of foods offered by both formal and informal retailers. Prices vary across products and across retailers (Gómez and Ricketts

⁷ Authors' calculation using the World Bank's PovcalNet database.

⁸ The poverty line is defined as \$1.90 daily per adult equivalent expenditure, calculated with purchasing power parity adjusted 2011 constant international dollars (World Bank, 2015)

2013). Market-induced desirability recognizes that the food environment affects the household desire for products via shopping atmosphere and marketing.

Our conceptual approach expands the framework used by Rischke et al. to include the effects that city size (defined by population) has on the urban environment that are exogenous to the food environment and consumer demand. The effects of city size are relevant to urban and rural consumption patterns alike. Urban consumption patterns will be directly affected by the urban environment of the city in which the household is located, and the effect of urban environment on rural households is expected to attenuate with increasing distance from the urban area. The anticipated attenuation of these effects is due to increased cost of accessing cities and the food markets therein as the distance from a household to cities increases (Fafchamps and Shilpi, 2003; Sharma, 2016).

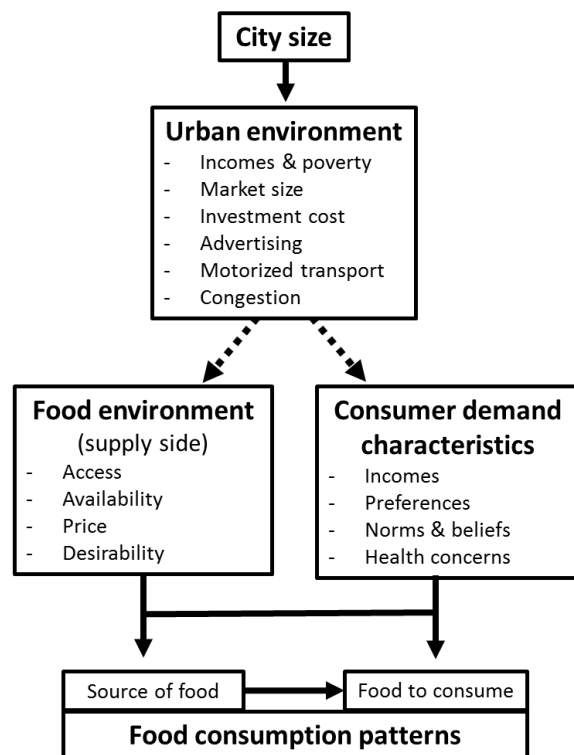


Figure 1-1: Conceptual framework

I suggest the following effects of a rising city size on the factors of urban environment that affect aspects of the food environment and consumer demand as follows. First, based on robust empirical patterns across Africa, incomes rise and poverty rates fall (Ferré et al., 2012, World Bank, 2009; Christiaensen et al., 2013). Both should increase the purchasing of food and the consumption of BSF foods.

Second, the size of the market increases more than proportionally to city size: this follows directly from a positive relationship between city size and mean consumer incomes. Increased size of market attracts investment in retail food markets, thereby improving physical access to them, which favors shopping in them. Increased access to food retail markets expands the variety of foods that is available.

Third, the cost of infrastructural investment rises, given higher costs of land – presuming market mechanisms operate in the land market. Higher investment cost works asymmetrically against investment in formal food retail: open air markets often arise spontaneously (informally) with no explicit cost of land, and street vendors do not typically pay for the space they occupy. These costs raise food prices in the more formal markets, reducing real incomes of consumers, and incentivize the use of informal retail that may not have the same variety of food as formal retailers. The higher land costs within and near larger cities provide a disincentive to urban and peri-urban farming that could be used for own production of food.

Fourth, consumers are more exposed to modern advertising. Advertising increases the likelihood that consumer preferences will align with the advertised foods that are commonly non-staple foods made more available from modern food retailers.

Fifth, access to public and private motorized transport increases. Access to motorized transport has an asymmetric effect on physical access to food: motorized transport improves

physical access for outlets such as supermarkets that are larger and fewer in numbers, but should have little or no effect on access to other outlet types, which have developed historically to serve largely foot-bound consumers. Access to transport thus favors the use of supermarket shopping, which exposes to consumers to greater variety of foods.

Sixth, congestion increases increase the cost of moving a given distance by motorized transport. However, alternative forms of transportation such as by foot or bicycle – both negatively associated with household incomes – are relatively insensitive to city size. Congestion thus asymmetrically affects physical access to food, it diminishes the access advantage conveyed by motorized transport and so reduces the probability of shopping in a formal retail market. Additionally, the increased cost of transportation increases the costs of cross-shopping⁹, reducing the benefits that multiple markets offer in terms of available variety of food.

This framework does not provide consistent expected effects of city size on the key food consumption patterns, but the generally anticipated trend leads me to hypothesize positive cumulative effects of city size and negative cumulative effects of distance to cities on the shares of purchased food and BSF foods in the household's total value of consumed food.

1.3 Methods

1.3.1 Context

Developing Eastern and Southern Africa, which I term “ESA” here, has been rapidly urbanizing. In the year 2000, twenty-one percent of the population in developing Eastern and

⁹ The practice of cross-shopping is where households shop at various locations. Cross-shopping is common in China (Goldman 2000), Israel (Hino 2014) and in South Africa, where households travel to locations away from their area of residence to acquire food (D’Haese and van Huylenbroeck 2005, Strydom 2011).

Southern Africa lived in urban settlements; by 2015, the urbanization level was 26 percent, and the United Nations Population Division¹⁰ projects that by 2050 the urban population share will increase to forty-four percent. I will use the nations of Malawi, Tanzania, Uganda and Zambia to represent this urbanizing region of Eastern Southern Africa. I define three city “types” based on population: primary cities with populations greater than one million, secondary cities with populations between one million and one hundred thousand, and tertiary cities with populations below one hundred thousand. These countries include three primary cities, thirty-eight secondary cities and over three hundred tertiary cities.

Two other characteristics of note regarding this region include the rising average per capita income and a modernizing food system. Data utilized are from 2010, and in the decade prior to 2010, the population weighted annual growth in per capita GNI¹¹ of these four nations was positive each year averaging near four percent since 2004. The relationship between household income and consumption of food, as indicated by Bennett’s Law, requires income also to be accounted for in this analysis. Daily per adult equivalent income is rising from low levels, resulting in an average income of just above three dollars¹² per day (Table 1-1). In spite of the low average incomes, the “supermarket revolution” is occurring in Eastern and Southern Africa (Reardon and Timmer, 2007). This revolution is modernizing the food system with the expansion of supermarket and other formal food retail outlets, it is impacting how households buy and sell

¹⁰ World Urbanization Prospects: The 2014 Revision

¹¹ World Bank data (<http://data.worldbank.org/>): GNI/capita, PPP (constant 2011 international \$)

¹² Expenditure data are converted to daily expenditure levels valued in purchasing power parity (PPP) adjusted constant 2011 international dollars using historic exchange rates obtained from XE.com and the conversion information for constant PPP valuation from worldbank.org. Household expenditures are evaluated at per adult equivalent (AE) values, which are calculated as one AE for each household member fifteen years of age or older, 0.75 AE for each child aged five to fourteen, and 0.50 AE for each child younger than five years old.

food. Formal retailers typically enter markets by establishing their retail outlets near relatively affluent areas within the greater market (Battersby and Peyton, 2014), which has resulted in supermarkets directing significant investment into larger cities and lesser investments into secondary and tertiary cities, thereby altering the food retail options across city size.

Table 1-1: National expenditure and population by expenditure strata and city size

	Total Expenditure (per adult equivalent)					Represented Population: Total ('000s) and Cumulative Density by Expenditure Strata								
	Malawi	Tanzania	Uganda	Zambia	Pooled	Total	< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00
National	2.48 (1.67)	3.03 (2.26)	3.51 (2.54)	3.13 (1.74)	3.11 (2.19)	101,923	2.1	13.8	30.2	45.2	67.4	86.2	96.7	100
Rural	1.97 (1.52)	2.38 (1.97)	2.78 (2.23)	1.78 (1.30)	2.37 (1.91)	77,041	2.6	16.9	36.3	53.2	76.6	93.4	99.1	100
Urban	5.31 (3.09)	4.91 (3.50)	6.10 (4.37)	5.66 (3.74)	5.39 (3.78)	24,882	0.4	4.1	11.2	20.4	38.8	64.0	89.2	100
Primary City		7.85 (5.83)	8.11 (6.39)	7.41 (4.62)	7.82 (5.72)	6,126	0.0	1.5	3.4	6.6	17.0	41.7	77.4	100
Secondary City	5.25 (3.14)	4.35 (3.34)	8.67 (4.56)	5.20 (3.68)	5.06 (3.43)	7,178	0.2	3.7	11.3	22.1	42.7	67.4	90.8	100
Tertiary City	5.52 (2.84)	3.52 (2.76)	5.10 (3.77)	4.48 (2.99)	4.31 (3.13)	11,578	0.8	5.6	15.2	26.6	48.0	73.6	94.4	100

Source: authors' calculations using national household level surveys

Notes: Daily average expenditure levels per adult equivalent are presented in purchasing power parity adjusted constant 2011 international dollars. Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

1.3.2 Data

City level population data are obtained from <http://www.citypopulation.de/>, which assembles the most recent available census data and uses country definitions of urban settlements. City population data is paired with household level data to identify the population size where urban households reside. Distance measurements for this analysis are approximated using distance from center of rural constituency to the nearest primary, secondary and tertiary cities for the Zambia data using the distance calculator at <https://www.daftlogic.com/>.

Household data are acquired from four nationally representative household surveys: 2010/2011 Malawi Integrated Household Survey (sample size of 12,271), 2010/2011 Tanzania National Panel Survey (3,924), 2009/2010 Uganda National Panel Survey (2,975), and 2010 Zambia Living Conditions Monitoring Survey (19,397). These datasets were selected due to the nations' geographic proximity, similarities in national demographics, the timing of the surveys, and the similarity of the data collection instruments. The enumerators of the surveys conducted in-person interviews during 2010 and each interview collected household data on consumption, employment, and other household characteristics.

Table 1-2 shows descriptive statistics of pooled data from the four datasets. The expenditure data are converted into daily expenditure levels valued in purchasing power parity (PPP) constant 2011 international dollars, using historic exchange rates obtained from XE.com and conversion factors for constant PPP valuation from worldbank.org. Expenditures are converted to per adult equivalent (AE) terms, and the daily expenditures per AE are used to calculate total household expenditure, total food consumption, and totals for various food consumption aggregates. Each survey collected data on the value of food consumption with a seven day recall, save for the Zambia survey, which used a two- or four week recall depending

on the food item. The valuation of consumed own production varies by dataset. The Uganda and Zambia surveys collected stated values of food consumed from own production¹³, while Malawi and Tanzania collected quantities; I value these quantities using imputed household and community prices of marketed food (see below for detail).

Table 1-2: Household descriptive characteristics

	Pooled Data					
	National	Rural	Urban	Primary Cities	Secondary Cities	Tertiary Cities
Total household expenditure per AE	3.11 (2.19)	2.37 (1.91)	5.39 (3.78)	7.82 (5.72)	5.06 (3.43)	4.31 (3.13)
Nonfarm employment	27.1 (20.0)	21.4 (0.0)	44.7 (50.0)	52.6 (50.0)	41.9 (40.0)	42.1 (40.0)
Dependency ratio	47.1 (50.0)	49.8 (50.0)	38.9 (40.0)	33.4 (33.3)	38.7 (40.0)	41.9 (42.9)
Household adult equivalents	6.7 (5.6)	6.8 (5.6)	6.4 (5.4)	6.5 (5.5)	6.2 (5.2)	6.4 (5.4)
Hectares of farmed land	3.5 (0.9)	4.1 (1.2)	1.4 (0.0)	0.5 (0.0)	1.2 (0.0)	2.0 (0.2)
Own a gas or electric stove	6.9 (0.0)	1.9 (0.0)	21.5 (0.0)	33.2 (0.0)	23.8 (0.0)	12.0 (0.0)
Own a refrigerator	8.7 (0.0)	2.3 (0.0)	27.7 (0.0)	48.8 (0.0)	29.3 (0.0)	12.9 (0.0)
Maximum education attained within the household	13.5 (14.0)	12.3 (14.0)	17.1 (17.0)	20.7 (17.0)	12.9 (13.0)	17.8 (17.0)
Age of household head	45.6 (43.0)	46.2 (44.0)	43.7 (41.0)	43.9 (41.0)	42.8 (41.0)	44.2 (41.0)
Female headed household	22.1 (0.0)	21.2 (0.0)	24.9 (0.0)	26.1 (0.0)	21.2 (0.0)	26.6 (0.0)
Own a phone	55.4 (100.0)	46.7 (0.0)	82.4 (100.0)	93.3 (100.0)	84.3 (100.0)	75.5 (100.0)
Kilometers to Market	9.5 (4.5)	10.7 (5.7)	5.7 (2.0)	10.3 (2.0)	4.7 (2.0)	3.9 (1.4)
Own a car	2.9 (0.0)	1.2 (0.0)	8.2 (0.0)	12.1 (0.0)	10.7 (0.0)	4.6 (0.0)
Own a motorcycle	4.7 (0.0)	4.6 (0.0)	5.2 (0.0)	3.9 (0.0)	4.3 (0.0)	6.4 (0.0)
Own a bicycle	47.0 (0.0)	51.9 (100.0)	31.9 (0.0)	13.7 (0.0)	35.4 (0.0)	39.3 (0.0)

Source: authors' calculations using national household level surveys

Notes: Daily average expenditure levels per adult equivalent are presented in purchasing power parity adjusted constant 2011 international dollars. Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

1.3.3 Definitions

Shares of aggregated food values are calculated from the household expenditure values to represent *purchased share* and share of *BSF foods*, shares of purchased food value and BSF food

¹³ The observations where Uganda reported farm-gate prices, I imputed market purchase prices to represent the value of own production.

value, respectively, in total food expenditure. Purchased food value does not differentiate between formal or informal food retailers, but represents the total value of household purchases of food. Food items are defined as BSF foods if their primary ingredient is dairy, meat, fruit, vegetables, nuts, vegetable oil, spices or sugar, and otherwise defined as low-income Bennett's Law (SF) food¹⁴.

The BSF aggregation is further disaggregated into 27 commodity based food aggregates based on the food item's primary input (Table 1-3). An example of a food item that is aggregated into a commodity aggregate would be bread being aggregated into the wheat commodity grouping because its primary ingredient is wheat flour.

Table 1-3: Dependent Variables

Table 1.3: Dependent Variables		
Share	Definition	
Purchased Share	Share of purchased food value in total food expenditure	
Beyond Staple Foods items	Foods whose primary ingredient is dairy, meat, fruit, vegetables, nuts, vegetable oil, spices or sugar	
27 commodity based food aggregates		
Rice	Oil Crops	Dairy Products
Maize	Oils & Fats	Poultry
Wheat	Staple Vegetables	Eggs
Other Cereals	Other Vegetables	Sugar and Sweets
Cassava	Fruits	Spices
Potatoes	Plantains	Non-Alcoholic Beverages
Other Tubers	Beef	Alcoholic Beverages
Pulses (with groundnuts)	Other Meat	Food Away From Home
Nuts	Aquatic Products	Other Foods

City size classifies cities by population: primary cities have populations greater than one million, secondary cities have populations below one million and above than 100,000, and tertiary cities have populations below one hundred thousand.

¹⁴ In Uganda, matoke is a low cost staple food and is therefore categorized as a SF food.

1.3.4 General model

To estimate the effects of city size and distance to cities I follow a four part analysis. Parts 1 and 2 establish the empirical configuration of the pattern of diet change across the income distribution, while parts 3 and 4 explore drivers that include the effects of city size on food consumption patterns.

First, I complete nonparametric analysis in the form of LOWESS curves. This analysis imposes no structure on the data, allowing the data to reveal patterns of consumption across the expenditure distribution. This analysis provides a first assessment of where in the expenditure distribution (e.g., above or below the poverty line) changes in consumption patterns begin, and whether the change is sudden and rapid or gradual.

Second, I formally establish breakpoints, and rates of change before and after them, by testing for expenditure thresholds at which a significant change occurs in (a) the rate of change or (b) the direction of the relationship between total expenditure and the consumption shares. This analysis is not as flexible as the nonparametric analysis, but by imposing linear relations on the food consumption patterns before and after an estimated expenditure breakpoint, it formally estimates particular expenditure levels at which patterns change, and establishes rates of change before the breakpoint (among lower income households) and above it (among higher income households).

Third, I examine the drivers of consumption change by using regression analysis of Engel curves to estimate the marginal effects of income and non-income determinants on food consumption shares. The results of the nonparametric and threshold analysis are applied to select an appropriate functional form of the Engel curve model for regression analysis.

Fourth, I use a mediation model to estimate the indirect effects that city size and distance to city have on food consumption patterns that occur via other determinant variables. When combined with the regression analysis of part 3, the indirect analysis provides results of the total effects of city size and distance to cities.

All analyses are completed separately for rural and urban households and by country, with descriptive analysis additionally highlighting city size differences that are analyzed in the parametric analysis.

1.3.5 Empirical model

The empirical model for the nonparametric analysis and the expenditure threshold analysis will estimate the marginal effect of income on the share of the value of a particular food aggregate in the total value of food consumed, without controlling for any other effects. The first two stages estimate Engel's curves for each share of the food aggregates, with adult equivalent total expenditure as the sole independent variable.

$$D_{Food\ Aggregate} = f(Y^*) \quad (1.2)$$

The models estimate the dependent variables: purchased share and the share of BSF foods, along with the twenty-seven commodity food aggregates, which add more detail to patterns of commodity based food consumption.

The third and fourth steps in the analysis use an Engel curve model to estimate the marginal effects of income and non-income household variables, including city size and distance to city, on each of the dependent variables in Table 1-3. The conceptual model posits that city size influences the urban environment, which affects both the food environment and consumer demand characteristics. To empirically specify the model, I include variables that capture as

many of the supply and demand side factors as possible. I include dummy variables for primary and secondary cities that reflect city size differences, and in the rural analysis I include variables for the distances to various city sizes. To the extent that variables are missing, or imperfectly capture the concepts in the conceptual model, I use the conceptual model to interpret the meaning of the city size variables and generate defensible conclusions.

The national datasets capture consumer demand characteristics relatively well, with variables for household per capita expenditure along with demographic variables that proxy for preferences, norms and beliefs, and health concerns. Data are less complete on the food environment. The national datasets provide four variables that are relevant to accessing food: own a car, own a motorcycle, own a bicycle, and distance to nearest market (for rural households). Measures of availability, price and food desirability are not available. The national datasets do not capture household level prices well due to the high prevalence of non-standard units. Systematic differences across cities will be captured by city dummy variables in pooled regressions.

Regarding the six urban environment factors within the conceptual model, I am able to directly control for two of them: income and access to motorized transport. Both are hypothesized to have positive effects on purchased share and the share of BSF food. The four remaining factors are captured by city size dummies – market size, congestion, investment cost, and advertising. Two of these, congestion and investment cost, are hypothesized to have negative or mixed effects on purchased share and the share of BSF food. It would thus not be surprising if I find that increasing city size decreases the shares of purchased or BSF food.

The effects of income, city size, distance to cities and the other selected variables are estimated by the following Engel's Curve Model:

$$D_{Dependent\ Variable} = f(Ac^*, Av^*, P^*, D^*, Y^*, Z^*, W^*, H^*) \quad (1.3)$$

The dependent variables are purchased share and the share of BSF food, with commodity based food consumption further estimated for each of the twenty-seven commodity based aggregates (Table 1-3). The determinants of equation three represent the factors of food environment and consumer demand characteristics, each with household variables that proxy for their effects (Table 1-4). Each of the independent variables are described below.

Table 1-4: Determinants of demand in the Engel's Curve Model

Ac*	Access	Own a car, Own a motorcycle, Own a bicycle, Distance to market (rural), City size (urban), Distance to cities (rural)	Y*	Disposable Income	Daily total expenditure per adult equivalent, City size (urban), Distance to cities (rural)
Av*	Availability	City size (urban), Distance to cities (rural)	Z*	Individual and Household Preferences	Nonfarm employment, Dependency ratio, Adult equivalents, Farmed hectares of land, Own a stove, Own a refrigerator, Maximum level of education, Age of household head, Female head of household, City size (urban), Distance to cities (rural)
P*	Price	City size (urban), Distance to cities (rural)	W*	Social and Individual Norms and Beliefs	Own a stove, Own a refrigerator, Maximum level of education, Age of household head, Female head of household, Own a telephone, City size (urban), Distance to cities (rural)
D*	Desirability	City size (urban), Distance to cities (rural)	H*	Health Concerns	Maximum level of education, Female head of household, Dependency ratio, City size (urban), Distance to cities (rural)

The variables that capture city size effects on urban environment are *dummy variables for residence in primary- and secondary cities (tertiary cities are excluded)*. For rural households, the variables regarding access to markets are *distance to nearest primary city, distance to nearest secondary city, distance to nearest tertiary city*. The distance (kilometer) variables are in logarithmic values to account for diminishing marginal effects of distance. Greater distance from cities reduces the effects of urban environment on food environment, consumer demand characteristics and ultimately food consumption decisions. The attenuation of these effects is

caused by increased costs of accessing urban areas and the diminishing spillover of the effects of the urban environment at greater distances from cities. I expect that the impact of distance will be greater for larger cities, as the density of commercial activity within larger cities is likely to result in greater commercial networks immediately surrounding the cities. The expansive commercial network allows for households near large cities to maintain consumption patterns more similar to those of urban residents, and as distance increases to large cities household food consumption patterns will experience greater marginal change.

The dummy variables *ownership of a car*, *ownership of a motorcycle*, and *ownership of a bicycle* are included to proxy for access to foods. Household ownership of a car would reduce the time required in transportation to markets and would increase household ability to buy and transport food purchased in bulk. Ownership of a motorcycle is expected to have a similar effect as ownership of a car, but offering less benefit on the ability to shop in bulk. Ownership of a bicycle also reduces the transportation costs associated with small purchases, but would likely have little effect on household ability to shop in bulk. *Distance to market* is also included as an independent variable to proxy for access to foods. The distance is measured in kilometers and enters the estimation in a logarithmic form to account for diminishing effects of distance. The greater the distance to the market the greater the shadow price of transportation that is required to purchase food, which is expected to reduce purchased share and BSF food consumption by limiting access to a variety of food. Tanzania and Zambia datasets include variables of household identified distance from household to market, whereas the Malawi and Uganda datasets include variables of distance from community to market that are collected by enumeration teams. Due to certain distances being collected only at the community level, the distance to market variables are only included in the analysis of consumption patterns of rural households.

Total daily household expenditure per AE is included in the model serving as a proxy for income, a key factor of consumer demand. The specific form that expenditure enters the estimation equations is explained in the estimation section.

Variables that proxy for individual and household preferences are included within the model. *Nonfarm employment* is calculated as the percentage of working age¹⁵ adults employed (including self-employment) in nonfarm activities. A household's *dependency ratio* is calculated by dividing the number of dependents by the total members within the household. Nonfarm employment and dependency ratio both increase households' opportunity cost of time, incentivizing households to reduce their time allocated towards producing and preparing food. The additional opportunity cost of time is expected to increase the consumption of purchased food in lieu of consuming food from own production. The number of *household adult equivalents* is the total number of adult equivalents within a household. Based on economies of scale within households, the number of adult equivalents reduces the average household opportunity cost of time. More household adult equivalents are therefore expected to reduce purchased share.

Independent variables that proxy for the shadow prices of household ability to produce, prepare and/or preserve food that contribute to individual and household preferences are *land farmed*, *own a gas/electric stove* and *own a refrigerator*. *Land farmed* indicates the hectares of land that a household farmed during the previous harvest season. This variable is included in logarithmic form to account for diminishing marginal effects of the amount of land farmed on food consumption. Larger areas of farmed land are more likely to be used for commercial purposes, reducing the marginal effect on household consumption as farmed area increases.

¹⁵ Ages 15-64

Farming one's own land is expected to lead to greater consumption of SF food from own production, so reducing demand for purchased and BSF food. Ownership of a *gas/electric stove* will reduce the preparation cost of food and thus reduce purchased food, which on average requires less preparation. Ownership of a *refrigerator* will reduce the cost of food preservation, enabling households to purchase greater quantities of perishable foods. The dummy variables for ownership of a gas/electric stove or a refrigerator take values of one if affirmative and zero if negative.

The final three variables that proxy for individual and household preferences are the following: The *maximum level of education* within a household is an ordinal value based on the level of education acquired that varies by country dataset that represents the household's maximum educational level. Increased education informs households on food safety and cleanliness issues, influencing which food items to consume and from where to acquire their food (Turrell and Kavanagh, 2006). The *age of the head of the household* is included in the empirical model as elderly households are more likely to maintain traditional diets, therefore older households will consume lower shares of purchased and BSF foods than younger households. Finally, the *gender of the head of the household* could influence the consumption patterns as consumption decisions have been shown to vary by gender (Smith, 2003). One highlighted characteristic of female heads of households is a willingness to delay gratification, leading to the expectation of less consumption of the primary food consumption patterns.

Variables included to proxy for social and individual norms include household *ownership of a telephone*, in addition to ownership of a gas/electric stove, ownership of a refrigerator, maximum level of education, age of household head, and female head of household, which have already been described. A dummy variable for household ownership of a phone has the value of

one if the household owns a phone, zero otherwise. Ownership of a phone indicates a step towards modernization and household potential willingness to modernize their food consumption patterns.

The variables included to proxy for health concerns are maximum level of education, female head of household, and dependency ratio, which have already been described

1.3.6 Estimation methods

LOWESS curve regressions (Cleveland, 1979; Cleveland and Devlin, 1988) are first applied to observe the nonparametric relationship between the food budget share and the household per adult equivalent total expenditure. The second application of LOWESS curves is used to map the relationship between distance to various city types and rural food consumption patterns. The nonparametric estimation is conducted via locally weighted scatterplot smoothing (LOWESS), which is a robust form of local polynomial smoothing. The LOWESS procedure does not impose functional form on the relationship between dependent and independent variables. Rather it uses locally weighted least squares estimation on subsets of the data to form a curve that represents the relationship between the variables. Compared with other forms of local polynomial smoothing, this specific form of local polynomial smoothing is more resistant to statistical outliers because it applies smaller local weights to estimates with large residuals. Accounting for statistical outliers is valuable given the nature of household survey data. To appropriately represent the sample population, the LOWESS analysis is conducted on expanded data where duplicate observations are recognized according to population weights.

In the second step of the analysis, I test for expenditure thresholds at which the direction or rate of change in the food consumption patterns could change. I apply a non-linear estimation method to find the best fit of two linear least squared curves that are connected at the point that

optimizes the fit of the two curves. This is known as a piecewise regression (Serber and Wild, 1989) and takes the following form:

$$E[y|x] = \begin{cases} \beta_{10} + \beta_{11}x, & x \leq \alpha \\ \beta_{20} + \beta_{21}x, & x > \alpha \end{cases} \quad (1.4)$$

$$s. t. \quad \beta_{10} + \beta_{11}\alpha = \beta_{20} + \beta_{21}\alpha$$

In this analysis, household food budget shares of food item aggregates (y) are regressed against household total expenditure per adult equivalent (x). This method provides inflection point values and slopes of each curve around the inflection points with robust standard errors. Statistically significant changes in the consumption patterns are identified when the piecewise regression identifies a kink in the consumption curve at the inflection point.

The nonparametric analysis and the threshold analysis lead into the third step in the analysis by identifying the functional form of the Engel curve that would provide the best fit for the patterns of food consumption. In the base functional form of an Engel curve, a share of consumption is the dependent variable and the independent variables include income and other household variables. Multiple forms of this model have been proposed regarding the specific form in which income relates to the dependent variable. Four of the options are; the linear relationship suggested in Engel's original work (1857), the linear relationship between log income and the dependent variable as identified by Working (1943), Working's model with the addition of the inverse of income as proposed by Leser in 1963, or the functional form proposed by Banks, Blundell and Lewbel (1996) that includes both log income and squared log income. The inclusion of non-linear relationships between income and the consumption of goods has been an improvement to the linear functional form as the non-linear relationships enable the analysis of consumption patterns that exhibit differing marginal, potentially inverted, effects of income on the relative shares of consumption of goods over strata of household income.

For the selection of the appropriate functional form of the Engel curve I follow a two-step process. First, the results of the nonparametric and semi-parametric analyses are used to narrow the options of functional form based on the general shapes of the observed patterns between total expenditure and the highlighted expenditure shares and the potential for the functional form to fit the observed patterns. Second, regressions using the reduced list of functional forms are completed with the food expenditure shares serving as the dependent variables, and the estimated fit of the regressions with the various functional forms are considered. Comparisons of adjusted R^2 or Pearson's chi-squared values divided by degrees of freedom are made between like dependent variables, and the functional form that most consistently provides the best fit is selected. I will follow the recommendation of Subramanian and Deaton (1996) to treat total expenditure as exogenous in food demand analysis.

The shares of consumption that serve as the dependent variables in Engel curve regressions are bound between zero and one. The standard estimation method for such a dependent variable is to convert it with a log-odds transformation and then to estimate with ordinary least squares (Wooldridge, 2010). The data used in this analysis are primarily collected in a one-week survey period, which typically results in many households with particular food budget shares of zero or one hundred percent – values that cannot be used in the log-odds transformation. Given the bounded nature of the dependent variable, the likelihood of reported zero or one hundred percent budget shares, and the assumption of asymptotically normal standard errors, the appropriate method of estimation in this specific equation is the fractional probit model (Papke and Wooldridge, 1996 and 2008). This model estimates the marginal effects of the predictors of a dependent variable that takes values of a closed set of zero to one.

Due to the selection of the fractional probit model, the second step of the selection of the functional form of the Engel curve will be completed by comparing Pearson's chi-squared values divided by degrees of freedom from regressions using each functional form. This estimation method, coupled with the above empirical model, enables the estimation of the marginal effects of the independent variables that serve as the determinants for the food consumption patterns.

In the fourth step of the analysis a mediation model is used to test for indirect effects of the variables *city size* and *distance to city* on food consumption patterns. City size and distance to city (predictor variables) could affect other determinant variables (mediator variables) that then affect food consumption patterns; this indirect effect would not be captured without the use of a mediation model. The indirect effects of predictor variables are calculated by separately multiplying the estimated effect of a mediator variable on a dependent variable by an estimated effect of the predictor variable on the mediator. OLS is used to estimate the city size and distance to city effects on the mediator variables, and bootstrapping is used to obtain standard errors for the indirect effects (Wooldridge, 2010). The independent variables used in the regressions that determine the effect of the predictor variables on the mediator variables are all of the independent variables in the empirical model, save for the independent variable being treated as the mediator variable. This estimation of indirect effects follows the methodology described in Hayes and Preacher (2010).

1.4 Results

I first address descriptive statistics, followed by the analysis of LOWESS curves and piecewise regressions to observe varying consumption patterns across income strata. Then I will discuss the results of the full Engel's Curve Model regression and the indirect effects of city size

on urban consumption patterns. I will return to LOWESS curve analysis to show patterns between distance to cities and food consumption patterns, and I will conclude this section with the indirect effects of distance to cities on food consumption patterns.

1.4.1 Descriptive statistics

Table 1-5 highlights patterns of purchased share and consumption of BSF foods, showing average food consumption shares across country by national, rural settlement, and urban settlement by city size. These descriptive statistics do not show causation, but are valuable in observing preliminary patterns between food consumption and city size.

Table 1-5: Food budget shares of key food consumption patterns, aggregated country and settlement

	Purchased Share				
	Pooled	Malawi	Tanzania	Uganda	Zambia
National	54.8 (53.0)	53.9 (52.2)	58.6 (59.0)	46.5 (40.2)	61.8 (66.5)
Rural	45.6 (41.8)	47.9 (45.8)	48.8 (46.1)	39.5 (33.6)	46.6 (43.5)
Urban	83.4 (95.7)	87.3 (95.6)	86.7 (98.3)	71.2 (84.3)	90.5 (97.6)
Primary City	92.9 (100.0)		97.6 (100.0)	82.2 (92.0)	95.1 (100.0)
Secondary City	90.0 (97.6)	90.0 (97.2)	91.5 (99.2)	76.8 (84.3)	90.7 (96.8)
Tertiary City	74.2 (86.1)	77.6 (86.5)	77.6 (87.9)	66.6 (79.1)	85.1 (93.6)

	Beyond Staple Foods items				
	Pooled	Malawi	Tanzania	Uganda	Zambia
National	48.8 (49.8)	50.9 (52.4)	51.6 (52.3)	40.4 (39.8)	56.0 (57.5)
Rural	45.5 (46.1)	49.2 (50.4)	48.2 (48.8)	37.1 (35.7)	52.4 (52.9)
Urban	59.1 (60.0)	60.1 (61.6)	61.4 (62.7)	52.1 (52.6)	62.9 (64.3)
Primary City	65.5 (66.7)		70.7 (71.3)	56.3 (54.7)	64.5 (65.3)
Secondary City	60.3 (61.3)	60.5 (61.7)	59.9 (60.1)	53.6 (52.2)	62.5 (64.1)
Tertiary City	55.0 (55.8)	58.5 (60.8)	56.9 (56.5)	50.4 (51.8)	61.9 (63.5)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

Purchased share is strongly related to settlement type: in every country, the level of purchased share approaches one hundred percent in primary cities and declines monotonically with falling city size, being lowest in rural areas. A key distinction between rural and urban households is that purchased share is almost twice as high among urban households as rural households. However, purchased share in rural areas averages over 45% in every country other than Uganda (40%), higher than what is observed in previous studies on purchased share in rural developing Africa (von Braun et al., 1991; Garrett and Ruel, 1999). The patterns observed for the shares of BSF food are consistent with the ordinal patterns of purchased share, but at a far smaller magnitude in variation, specifically, rural consumption shares of BSF foods average nearly eighty percent of the shares in urban areas.

Table 1-6 further disaggregates commodity based food consumption patterns and displays pooled national data of eight of the twenty-seven commodity-based food aggregates – three low-income Bennett's Law (SF) foods and five BSF foods. Wheat, although a grain, is considered by many households within ESA to be a luxury good and therefore has a consumption pattern similar to that of BSF foods, i.e., monotonic decrease with falling city size. The other highlighted aggregates follow patterns between rural and urban households that are expected based on what is observed in Table 1-5. Monotonic relationships are less prevalent across city size for the commodity aggregates using the pooled data, but there is more consistency with the expected relationships between city size and the commodity aggregates when observing data that remains disaggregated by country (Table 1-A-1). Food away from home¹⁶ represents a significant portion

¹⁶ Food away from is included in the commodity aggregates as food away from home is commonly reported generally as food away from home without indication of specific base commodity.

of urban food share (13.4%), and this share is shown to rise monotonically across city size when considering country specific data (Table 1-A-1).

Table 1-6: Food budget shares of select commodity aggregates, pooled data by settlement

Pooled Data	Staple Foods			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
National	18.1 (13.1)	2.5 (0.0)	6.1 (0.0)	3.2 (2.5)	6.3 (4.9)	3.4 (0.0)	2.8 (0.0)	7.1 (0.0)
Rural	19.9 (15.5)	1.6 (0.0)	7.4 (0.6)	2.9 (2.1)	6.2 (4.5)	2.8 (0.0)	2.8 (0.0)	5.1 (0.0)
Urban	12.5 (8.6)	5.1 (3.6)	2.1 (0.0)	4.2 (3.6)	6.8 (5.7)	5.4 (3.4)	2.8 (0.0)	13.4 (0.0)
Primary City	8.4 (5.7)	6.3 (5.2)	1.0 (0.0)	3.6 (3.2)	6.6 (5.9)	6.1 (5.3)	2.9 (0.0)	21.0 (13.3)
Secondary City	14.6 (11.5)	6.2 (4.3)	1.3 (0.0)	5.3 (4.5)	7.9 (6.5)	5.1 (2.6)	2.4 (0.0)	9.2 (0.0)
Tertiary City	13.3 (9.0)	3.8 (1.4)	3.2 (0.0)	3.9 (3.3)	6.2 (5.1)	5.1 (0.0)	3.1 (0.0)	11.9 (0.0)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

1.4.2 Nonparametric analysis of the relationship between expenditure and food consumption

LOWESS results (Figure 1-2, left panel) show that rural purchased share follows a pattern of initial decline in the share of purchased food followed by a monotonic rise in each of the four countries. This pattern resembles the functional forms of Engel curves for luxury goods as suggested by Leser (1963) or Banks, Blundell and Lewbell (1997). These curves indicate that a typical rural household at any income level does not purchase less than 40% of their consumed food value in any of the observed countries other than Uganda. The initial high levels of purchased share in rural households are likely caused by a lack of physical capital (especially land) required to produce a desired level of own production in addition to need based cash transfers.

Nonparametric analysis of the relationship between income and hectares of previously cultivated land¹⁷ shows that the poorest rural households have the least land, suggesting that poor

¹⁷ Limited sample to households with less than five hectares of land.

households face high shadow prices for the production of food, supporting high shares of purchased share at low income levels. Households with expenditures below one dollar consume more food received as gifts or food in kind compared with the rest of the population (10% vs 6%), therefore it is logical that need based cash transfers that result in food purchases would be the most prevalent at the lowest expenditure levels.

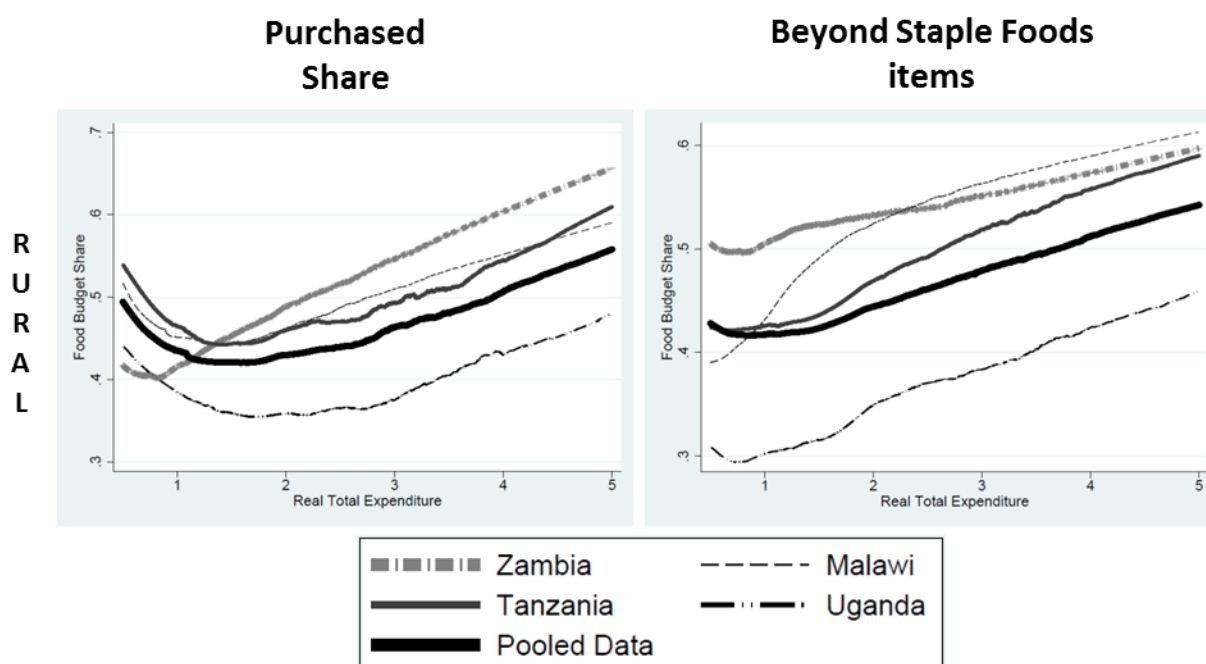


Figure 1-2: Nonparametric (LOWESS curve) analysis of select food consumption patterns – rural data

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

Urban households within the same countries exhibit a similar estimated relationship between income and purchased share with two exceptions (Figure 1-3, left panel). First, urban Zambia does not display an initial decline, instead following an approximate log-linear relationship as suggested by Working (1943) for luxury goods. The lack of an initial decline in purchased share at low income levels suggests that Zambia has relatively less urban agriculture than the other represented nations. Second, urban households exhibit higher levels of purchased

share than rural households. Considering pooled data within the estimated range of total expenditure, the highest estimate of purchased share for rural households is less than the lowest estimate of purchased share for urban households.

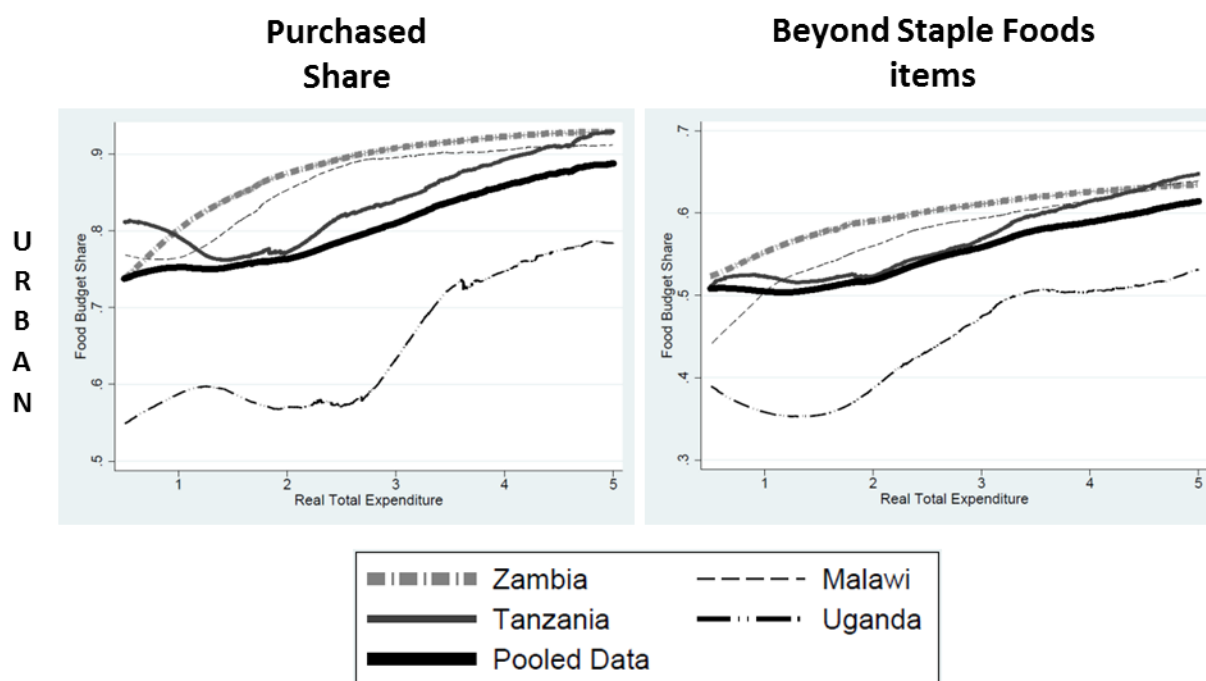


Figure 1-3: Nonparametric (LOWESS curve) analysis of select food consumption patterns – urban data

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

For both rural and urban households the slopes of the fitted estimates of purchased share are positive at all income levels above the poverty line, with the slopes of the pooled data beginning to rise at expenditure levels below the poverty line, indicating a transition towards purchased food at low income levels.

As expected, the LOWESS curves in Figures 1-2 and 1-3 (right panels) are consistent with Bennett's Law as BSF foods have a near linear relationship with increasing slopes starting at low household expenditure levels. The disaggregated commodity based consumption patterns, shown in Figures 1-4 (rural) and 1-5 (urban), show greater variation in the consumption patterns

of individual commodity aggregates than the aggregated pattern of BSF foods in Figures 1-2 and 1-3. Among rural households the food budget shares of the commodity aggregates for maize and cassava (Figure 1-4, top left), and staple vegetables (top right) fall by one-third between daily expenditures of \$0.50 and \$2.00. Over the same income range in urban areas, the estimated food budget share of cassava declines over two thirds (Figure 1-5, top left). Commodity aggregates with rising food budget shares among poor households include wheat, beef, dairy and food away from home. Food away from home exhibits the greatest increase in consumption share across income, again highlighting household demand for attributes such as time saving as expenditures rise. These patterns suggest that households follow the food need hierarchy that indicated that households initially satisfy their need for calories (cheap calories in maize and cassava) and then address higher order needs such as good tasting and novel food, which are satisfied with the consumption of higher valued foods, i.e. meat and food away from home.

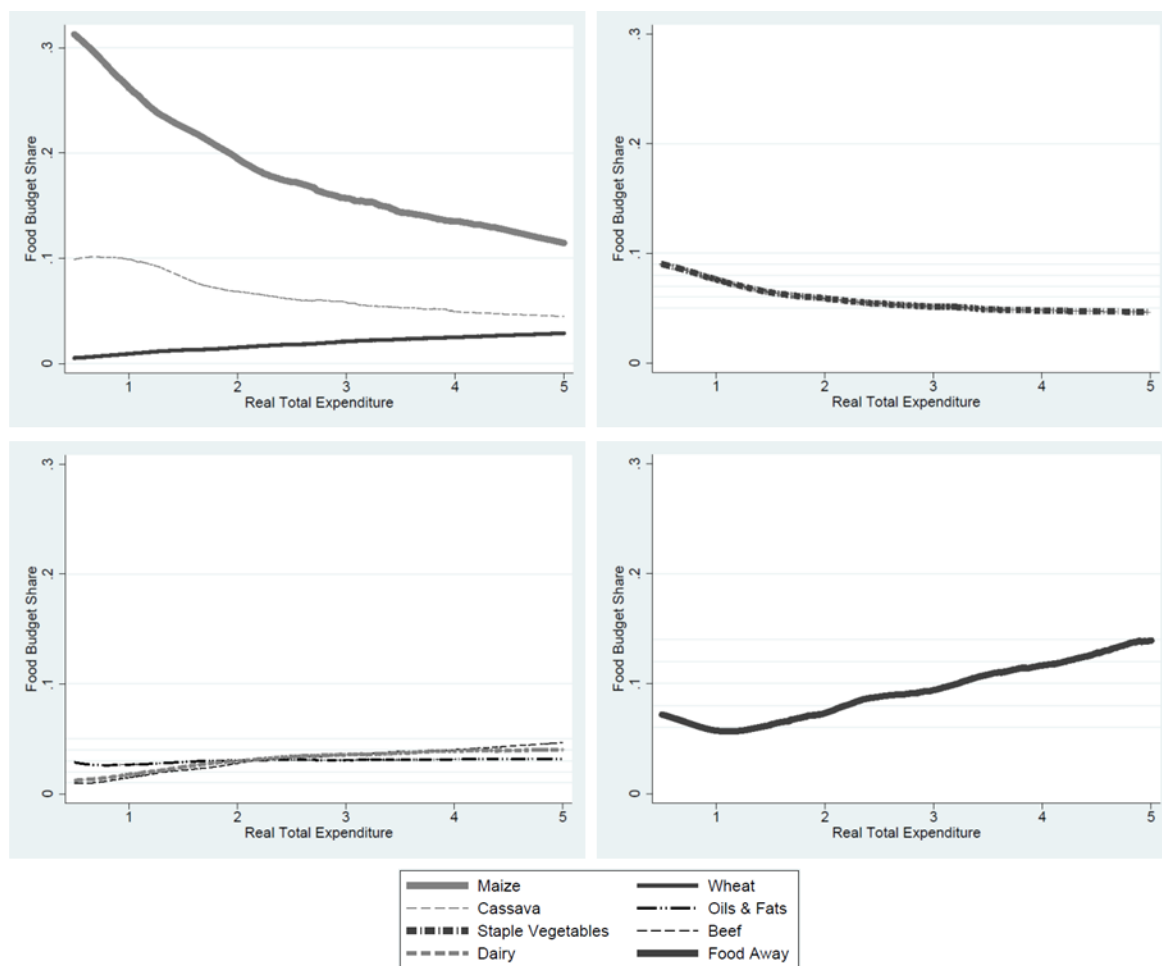


Figure 1-4: Nonparametric (LOWESS curve) analysis of select commodity aggregates – pooled rural data

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

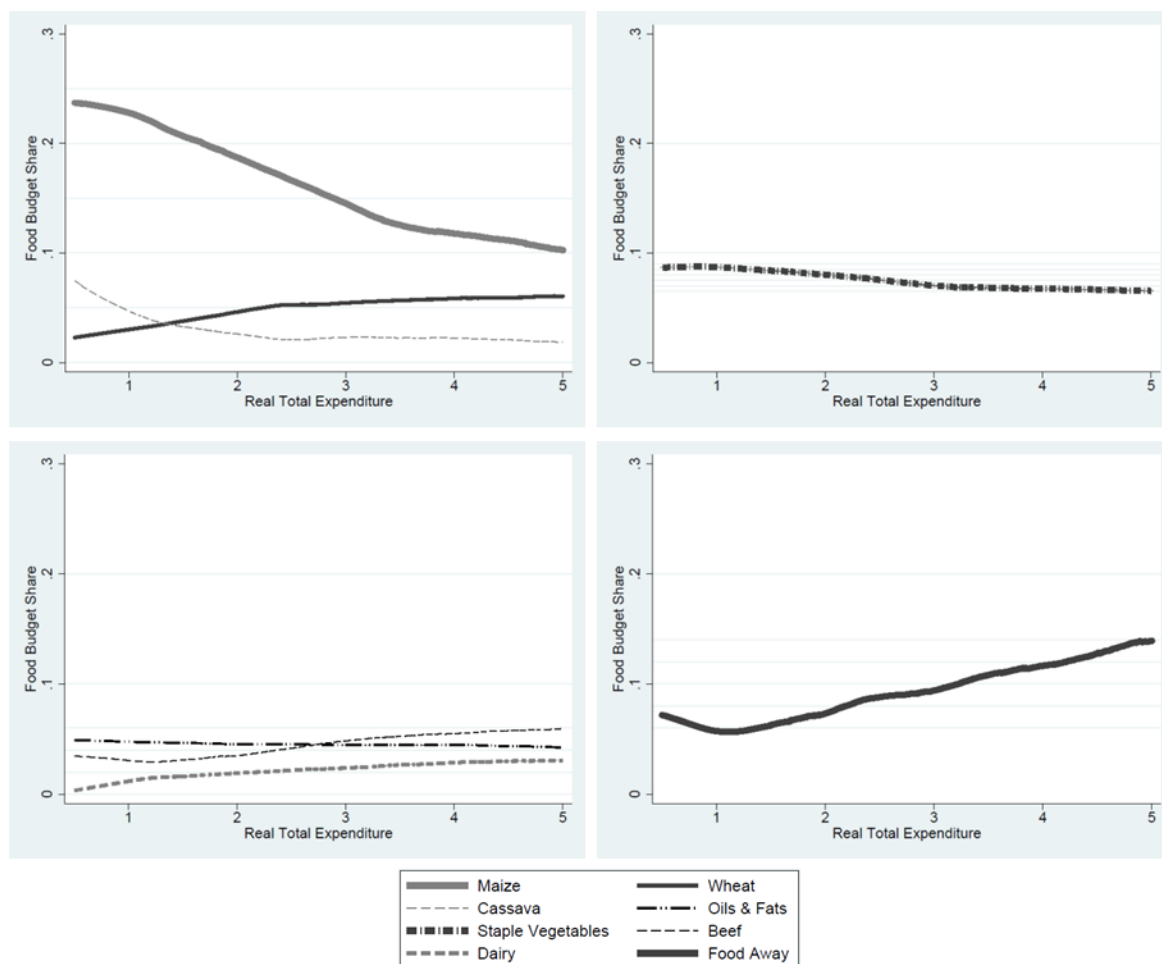


Figure 1-5: Nonparametric (LOWESS curve) analysis of select commodity aggregates – pooled urban data

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

In sum, LOWESS curves provide fitted estimates that indicate changing food consumption patterns across expenditure strata, specifically among the poor. The slopes of the consumption patterns estimated by the LOWESS curves vary across country and settlement, but the general pattern of the primary highlighted food consumption patterns is one of initial decline in food budget share followed by a monotonic rise that begins below the poverty line.

1.4.3 Expenditure Threshold Analysis

The expenditure threshold analysis tests for the presence of statistically significant expenditure thresholds that indicate changing directions or rates of change in the consumption patterns of purchased share or the share of BSF foods. Expenditure threshold analysis indicates that patterns of rural and urban purchased share vary across income levels, and that the steady rise of purchased share begins among poor households.

The results for rural purchased share in Table 1-7 show that in three of four countries there is a statistically significant pattern of declining food budget shares of purchased food followed by a rise in food budget shares, and that three of four of the countries observe the change in direction below or near the poverty line. The pooled data from the four countries show the expenditure threshold of purchased share to occur at \$1.56.

Table 1-7: Expenditure thresholds and slopes prior to and after expenditure thresholds in the consumption patterns of purchased share and Beyond Staple Foods items – rural data

Countries	Purchased Share			Beyond Staple Foods items		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.31 (0.000)	-0.084 (0.000)	0.046 (0.000)	1.950 (0.000)	0.103 (0.000)	0.029 (0.000)
Tanzania	\$1.14 (0.000)	-0.188 (0.011)	0.042 (0.000)	1.312 (0.001)	0.004 (0.904)	0.055 (0.000)
Uganda	\$2.09 (0.000)	-0.046 (0.046)	0.056 (0.000)			0.041 (0.000)
Zambia			0.059 (0.000)	1.568 (0.103)	0.026 (0.035)	0.015 (0.026)
Pooled	\$1.56 (0.000)	-0.069 (0.000)	0.043 (0.000)	1.297 (0.000)	-0.005 (0.789)	0.037 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

In urban households the slopes of the pattern of purchased share are positive both before and after the expenditure breakpoints when they are statistically significant (Table 1-8). This indicates that urban households prefer greater shares of purchased food at all income levels,

which is expected because of high shadow prices of agricultural production for households living in urban settlement due the limited access to farmland.

Table 1-8: Expenditure thresholds and slopes prior to and after expenditure thresholds in the consumption patterns of purchased share and Beyond Staple Foods items – urban data

Countries	Purchased Share			Beyond Staple Foods items		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$2.64 (0.000)	0.081 (0.000)	-0.001 (0.929)	\$2.01 (0.000)	0.069 (0.001)	0.025 (0.000)
Tanzania	\$1.58 (0.001)	-0.075 (0.461)	0.063 (0.000)	\$1.53 (0.008)	-0.022 (0.740)	0.045 (0.000)
Uganda			0.064 (0.003)	\$0.63 (0.000)	-1.050 (0.028)	0.055 (0.000)
Zambia	\$2.02 (0.000)	0.087 (0.000)	0.017 (0.000)	\$1.85 (0.000)	0.047 (0.002)	0.015 (0.000)
Pooled	\$1.67 (0.010)	-0.019 (0.737)	0.048 (0.000)	\$1.42 (0.036)	-0.010 (0.845)	0.035 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

The expenditure threshold analysis for the share of BSF foods leads to two results of note. First, the estimates affirm Bennett's Law and indicate that low incomes households are rapidly moving towards BSF foods, which is shown with greater positive slopes prior to the expenditure breakpoints than beyond them in all but one case where both slopes were statistically significant – the exception has an expenditure breakpoint of \$0.63 (Table 1-8). These results suggest that a supernumerary income is not a necessary condition for household consumption patterns to exhibit Bennett's Law.

Second, the patterns of the disaggregated commodity aggregates further support the non-necessary condition of supernumerary income for a shifting consumption towards BSF foods (Tables 1-9 and 1-10). The select commodity aggregates continue to show that the bulk of the transition within BSF food consumption occurs among poor households. Fifteen of the 18 non-pooled statistically significant pairs of slopes representing commodity food consumption patterns have the same direction and are of greater absolute value before the expenditure breakpoints than

beyond the breakpoints: negative slopes for maize, cassava and staple vegetables, positive for wheat, oil and fats, beef, dairy and food away from home.

Table 1-9: Expenditure thresholds and slopes prior to and after expenditure thresholds for the aggregates of select commodity based aggregates – rural data

Maize				Rice		
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.95 (0.000)	-0.140 (0.000)	-0.038 (0.000)	\$2.35 (0.000)	0.016 (0.000)	0.008 (0.000)
Tanzania	\$1.38 (0.184)	-0.026 (0.404)	-0.043 (0.000)			0.007 (0.000)
Uganda		-0.007 (0.284)	-0.008 (0.088)	\$4.49 (0.000)	0.004 (0.000)	0.061 (0.091)
Zambia	\$1.32 (0.001)	-0.051 (0.013)	-0.022 (0.000)	\$1.46 (0.000)	0.025 (0.000)	0.013 (0.000)
Pooled	\$2.13 (0.000)	-0.074 (0.000)	-0.026 (0.000)	\$1.25 (0.026)	0.008 (0.000)	0.006 (0.000)
Cassava				Oils & Fats		
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.40 (0.000)	0.008 (0.084)	-0.005 (0.000)	\$2.31 (0.000)	0.013 (0.000)	0.001 (0.339)
Tanzania	\$1.64 (0.000)	-0.073 (0.000)	-0.006 (0.016)	\$1.96 (0.044)	0.002 (0.407)	-0.001 (0.670)
Uganda	\$2.38 (0.000)	-0.057 (0.000)	-0.013 (0.121)	\$2.69 (0.000)	0.006 (0.060)	-0.003 (0.071)
Zambia	\$1.94 (0.000)	-0.029 (0.000)	-0.008 (0.385)	\$3.30 (0.008)	0.002 (0.205)	0.005 (0.290)
Pooled	\$1.96 (0.000)	-0.031 (0.000)	-0.009 (0.002)	\$1.98 (0.000)	0.004 (0.014)	0.000 (0.561)
Staple Vegetables				Beef		
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.29 (0.000)	-0.061 (0.000)	-0.012 (0.000)		0.007 (0.000)	
Tanzania	\$1.80 (0.000)	-0.029 (0.000)	-0.009 (0.000)	\$1.70 (0.002)	0.014 (0.016)	0.003 (0.242)
Uganda		0.001 (0.486)	0.001 (0.140)	\$2.42 (0.003)	0.016 (0.000)	0.007 (0.193)
Zambia	\$1.55 (0.003)	0.006 (0.034)	0.001 (0.391)	\$0.86 (0.011)	0.013 (0.073)	0.005 (0.004)
Pooled	\$1.79 (0.000)	-0.024 (0.000)	-0.006 (0.000)	\$2.41 (0.000)	0.013 (0.000)	0.005 (0.091)
Dairy				Food Away From Home		
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.53 (0.001)	0.004 (0.000)	0.007 (0.000)	\$1.78 (0.000)	0.007 (0.000)	0.002 (0.010)
Tanzania	\$1.96 (0.028)	0.010 (0.129)	0.001 (0.787)	\$2.01 (0.000)	0.012 (0.197)	0.037 (0.000)
Uganda	\$2.27 (0.000)	0.020 (0.000)	0.003 (0.461)			0.006 (0.004)
Zambia	\$2.00 (0.414)	0.005 (0.000)	0.004 (0.116)	\$3.09 (0.000)	0.000 (0.791)	0.002 (0.050)
Pooled	\$2.20 (0.000)	0.013 (0.000)	0.003 (0.170)	\$2.86 (0.001)	0.014 (0.000)	0.025 (0.003)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 1-10: Expenditure thresholds and slopes prior to and after expenditure thresholds for the aggregates of select commodity based aggregates – urban data

Maize						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.38 (0.000)	-0.246 (0.000)	-0.033 (0.000)	\$2.57 (0.000)	0.030 (0.000)	0.000 (0.956)
Tanzania	\$1.40 (0.004)	0.031 (0.607)	-0.036 (0.000)	\$2.39 (0.000)	0.014 (0.011)	0.000 (0.947)
Uganda	\$3.12 (0.000)	-0.051 (0.001)	0.029 (0.048)	\$0.90 (0.003)	-0.019 (0.322)	0.010 (0.000)
Zambia	\$2.06 (0.000)	-0.066 (0.000)	-0.028 (0.000)	\$2.49 (0.000)	0.030 (0.000)	0.009 (0.014)
Pooled	\$1.99 (0.062)	-0.048 (0.020)	-0.031 (0.000)	\$2.51 (0.000)	0.015 (0.000)	0.003 (0.328)
Rice						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi		-0.002 (0.034)	-0.002 (0.024)	\$1.61 (0.000)	0.040 (0.000)	0.003 (0.251)
Tanzania	\$1.77 (0.000)	-0.040 (0.246)	-0.002 (0.222)	\$1.76 (0.015)	-0.008 (0.353)	0.000 (0.979)
Uganda	\$1.96 (0.000)	-0.049 (0.099)	0.006 (0.491)	\$2.31 (0.002)	0.009 (0.102)	0.000 (0.934)
Zambia	\$1.59 (0.000)	-0.034 (0.000)	-0.003 (0.000)	\$2.03 (0.000)	-0.007 (0.046)	-0.001 (0.518)
Pooled	\$1.81 (0.000)	-0.034 (0.055)	0.001 (0.636)		-0.001 (0.359)	-0.001 (0.306)
Cassava						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi		-0.002 (0.034)	-0.002 (0.024)	\$1.61 (0.000)	0.040 (0.000)	0.003 (0.251)
Tanzania	\$1.77 (0.000)	-0.040 (0.246)	-0.002 (0.222)	\$1.76 (0.015)	-0.008 (0.353)	0.000 (0.979)
Uganda	\$1.96 (0.000)	-0.049 (0.099)	0.006 (0.491)	\$2.31 (0.002)	0.009 (0.102)	0.000 (0.934)
Zambia	\$1.59 (0.000)	-0.034 (0.000)	-0.003 (0.000)	\$2.03 (0.000)	-0.007 (0.046)	-0.001 (0.518)
Pooled	\$1.81 (0.000)	-0.034 (0.055)	0.001 (0.636)		-0.001 (0.359)	-0.001 (0.306)
Oils & Fats						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.97 (0.000)	-0.041 (0.006)	-0.011 (0.000)	\$1.47 (0.007)	0.004 (0.661)	0.014 (0.000)
Tanzania	\$0.74 (0.007)	0.068 (0.346)	-0.006 (0.007)	\$1.47 (0.002)	-0.026 (0.348)	0.009 (0.014)
Uganda	\$1.80 (0.037)	-0.006 (0.679)	0.005 (0.051)	\$1.99 (0.057)	0.026 (0.231)	0.007 (0.304)
Zambia	\$1.30 (0.028)	0.003 (0.763)	-0.006 (0.000)	\$2.00 (0.000)	0.016 (0.000)	0.004 (0.000)
Pooled	\$0.71 (0.018)	0.040 (0.469)	-0.007 (0.000)	\$1.38 (0.036)	-0.007 (0.728)	0.009 (0.000)
Staple Vegetables						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.97 (0.000)	-0.041 (0.006)	-0.011 (0.000)	\$1.47 (0.007)	0.004 (0.661)	0.014 (0.000)
Tanzania	\$0.74 (0.007)	0.068 (0.346)	-0.006 (0.007)	\$1.47 (0.002)	-0.026 (0.348)	0.009 (0.014)
Uganda	\$1.80 (0.037)	-0.006 (0.679)	0.005 (0.051)	\$1.99 (0.057)	0.026 (0.231)	0.007 (0.304)
Zambia	\$1.30 (0.028)	0.003 (0.763)	-0.006 (0.000)	\$2.00 (0.000)	0.016 (0.000)	0.004 (0.000)
Pooled	\$0.71 (0.018)	0.040 (0.469)	-0.007 (0.000)	\$1.38 (0.036)	-0.007 (0.728)	0.009 (0.000)
Beef						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.62 (0.000)	0.001 (0.521)	0.010 (0.000)		0.001 (0.638)	0.001 (0.519)
Tanzania	\$2.54 (0.052)	0.006 (0.440)	-0.002 (0.557)	\$1.48 (0.018)	-0.008 (0.863)	0.046 (0.000)
Uganda		0.010 (0.171)	0.012 (0.003)	\$0.43 (0.000)	-3.746 (0.000)	0.006 (0.548)
Zambia	\$2.90 (0.001)	0.007 (0.000)	0.004 (0.025)	\$1.35 (0.158)	0.001 (0.312)	0.002 (0.032)
Pooled	\$1.22 (0.003)	0.017 (0.004)	0.005 (0.002)	\$1.41 (0.018)	-0.012 (0.745)	0.024 (0.000)
Dairy						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.62 (0.000)	0.001 (0.521)	0.010 (0.000)		0.001 (0.638)	0.001 (0.519)
Tanzania	\$2.54 (0.052)	0.006 (0.440)	-0.002 (0.557)	\$1.48 (0.018)	-0.008 (0.863)	0.046 (0.000)
Uganda		0.010 (0.171)	0.012 (0.003)	\$0.43 (0.000)	-3.746 (0.000)	0.006 (0.548)
Zambia	\$2.90 (0.001)	0.007 (0.000)	0.004 (0.025)	\$1.35 (0.158)	0.001 (0.312)	0.002 (0.032)
Pooled	\$1.22 (0.003)	0.017 (0.004)	0.005 (0.002)	\$1.41 (0.018)	-0.012 (0.745)	0.024 (0.000)
Food Away From Home						
Countries	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.62 (0.000)	0.001 (0.521)	0.010 (0.000)		0.001 (0.638)	0.001 (0.519)
Tanzania	\$2.54 (0.052)	0.006 (0.440)	-0.002 (0.557)	\$1.48 (0.018)	-0.008 (0.863)	0.046 (0.000)
Uganda		0.010 (0.171)	0.012 (0.003)	\$0.43 (0.000)	-3.746 (0.000)	0.006 (0.548)
Zambia	\$2.90 (0.001)	0.007 (0.000)	0.004 (0.025)	\$1.35 (0.158)	0.001 (0.312)	0.002 (0.032)
Pooled	\$1.22 (0.003)	0.017 (0.004)	0.005 (0.002)	\$1.41 (0.018)	-0.012 (0.745)	0.024 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Overall, the semi-parametric expenditure threshold analysis highlights three key results. First, there is evidence of statistically significant changes in the direction and/or slope of the shares of food aggregates relative to total expenditure levels. Second, if the expenditure

breakpoint signifies a change in the direction of the food consumption pattern, the breakpoint is commonly below the poverty line. Third, if the expenditure breakpoint signifies a change in the slope of the food consumption pattern, the slope is commonly of greater absolute value at expenditures below the expenditure breakpoint than above. These findings indicate that there is significant change in food consumption patterns among poor households in ESA, and that change is often more rapid for the poor than it is for the non-poor.

1.4.4 Engel's Curve Model regression analysis

The fractional probit regressions of the Engel curves build upon the previous analyses by better isolating the impact of income on food consumption patterns and the effects of the various non-income determinants. Following the two-step process for the selection of the functional form of the Engel curve, the functional form suggested by Banks, Blundell and Lewbel (1997) was selected. The nonparametric and semi-parametric analyses highlight the need for the functional form to allow for changing slopes of the food consumption patterns, narrowing the functional form options to those suggested by Banks et al. (1997) and Leser (1963). Fractional probit regressions of the highlighted food consumption patterns with dependent variables specific to these two functional forms resulted in thirteen of the twenty-four regressions having a superior goodness of fit with the Banks et al. suggested functional form. The selected functional form includes non-income independent variables and two independent variables for expenditure: log total household expenditure and the squared log of total household expenditure. As expenditure is included in the model as two separate variables, the average partial effects of expenditure are calculated using the following equation:

$$\frac{\partial}{\partial(\text{expenditure})} = \left[a \frac{1}{\text{expenditure}} + 2b \frac{\ln(\text{expenditure})}{\text{expenditure}} \right] * E\phi[X\beta] \quad (1.5)$$

where a equals the marginal effects of log expenditure and b equals the marginal effects of squared log expenditure, and the scaling factor in the right of the equation is equal to the average of the normal densities of the predicted values of the dependent variable.

The regression analysis portion of the results section will first consider the effects of expenditure. Then I will observe non-expenditure drivers of consumption, including the direct and the indirect effects of city size on urban consumption patterns. Then I will briefly return to LOWESS analysis to graphically represent the relationship between distance to cities and rural food consumption patterns, and conclude by showing the direct and indirect effects of distance to cities on food consumption patterns.

The estimated effects of expenditure provide three results of note. First, the average partial effects of expenditure on purchase share are positive in seven of eight of rural or urban estimates, but in only one of two significant marginal effects (Tables 1-11 and 1-12). It is surprising that this was not more statistically significant given the clear patterns observed in the nonparametric and threshold analyses, but it speaks to the importance of non-income determinants on the share of purchased food. The estimated effects of expenditures on purchased share graphed in Figures 1-6 and 1-7 resemble the patterns of the previous analyses, consistent with the correlation of expenditure and purchase share.

Table 1-11: Average partial effects of expenditure for key food consumption patterns – rural data

Rural	Purchased Share				Beyond Staple Foods items			
	Malawi	Tanzania	Uganda	Zambia	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	-0.015 (0.000)	0.004 (0.525)	0.007 (0.214)	0.007 (0.150)	0.050 (0.000)	0.040 (0.000)	0.023 (0.000)	0.004 (0.269)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

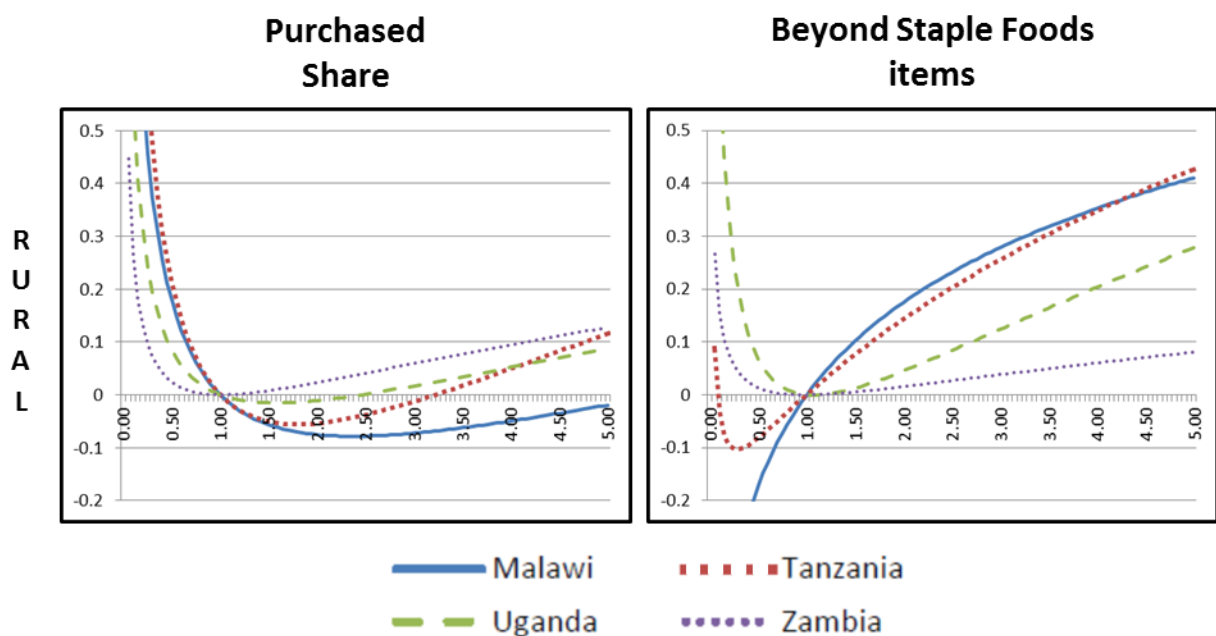


Figure 1-6: Estimated effects of expenditure on purchased share and the consumption of Beyond Staple Foods items – rural data

Table 1-12: Average partial effects of expenditure and city size for key food consumption patterns – urban data

Urban	Purchased Share				Beyond Staple Foods items			
	Malawi	Tanzania	Uganda	Zambia	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.001 (0.756)	0.013 (0.001)	0.004 (0.296)	0.005 (0.296)	0.008 (0.000)	0.025 (0.000)	0.015 (0.000)	0.009 (0.000)
Household located in a primary city		0.090 (0.000)	0.063 (0.048)	0.037 (0.000)		0.017 (0.172)	-0.029 (0.130)	-0.010 (0.058)
Household located in a secondary city	0.090 (0.000)	0.051 (0.000)	0.036 (0.349)	0.022 (0.000)	0.007 (0.445)	-0.017 (0.208)	-0.035 (0.229)	-0.008 (0.087)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

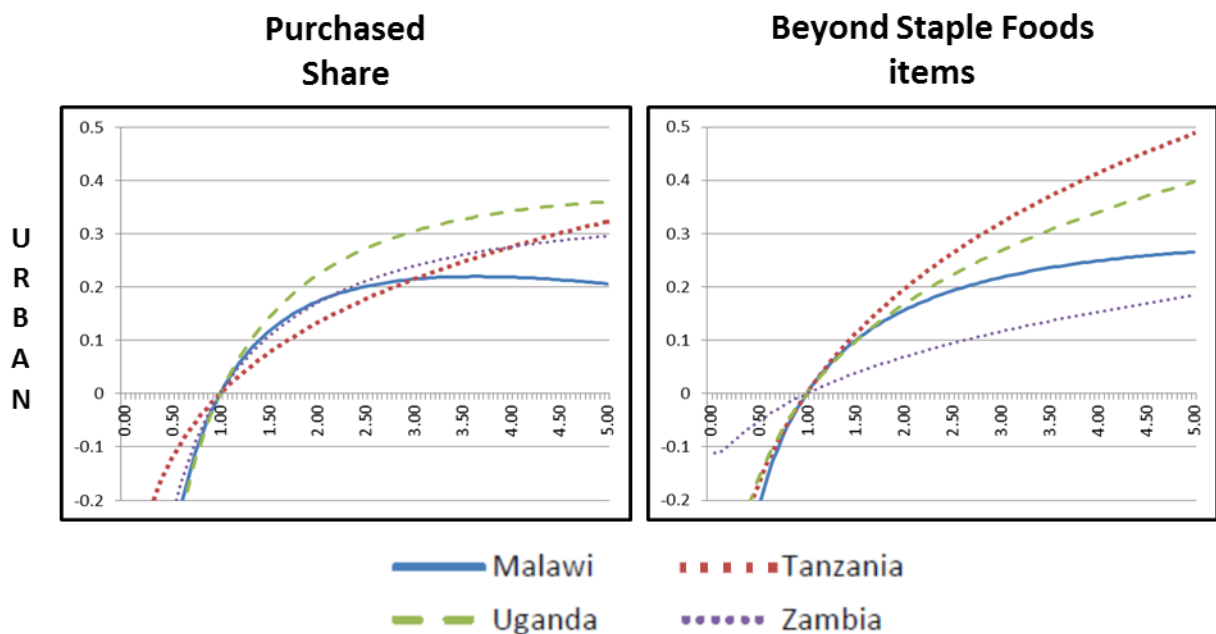


Figure 1-7: Estimated effects of expenditure on purchased share and the consumption of Beyond Staple Foods items – urban data

Second, the average partial effects of expenditure on BSF food are positive and statistically significant in seven of eight estimations, consistent with Bennett's Law (Tables 1-11 and 1-12). When these estimated effects are represented across the expenditure strata their marginal effects are positive at all expenditure levels above one dollar (Figures 1-6 and 1-7). These findings further support that supernumerary income is not a condition for households to increase their share of BSF food.

Third, the statistically significant average partial effects of expenditure on the selected commodity aggregates are consistently negative for maize, cassava and staple vegetables; consistently positive for wheat, beef, dairy and food away from home; and mixed for oils and fats. These consistent patterns highlight the heterogeneity of effects within the greater Bennett's Law aggregates.

Three findings regarding the effects of city size on urban consumption patterns are of note. First, larger cities positively influence households to purchase a larger share of the food they consume. The marginal effects of households living in primary cities and households living in secondary cities on purchased share were estimated and six of the seven estimates¹⁸ from the four countries were positive and statistically significant. As shown with the data from Tanzania and Zambia, the marginal effect of a household living in primary city on purchased share was twice the magnitude of a household living in a secondary city, indicating a positive relationship between city size and purchased food consumption. These findings are consistent with the conceptual model, where larger cities improve the access to markets due to the increased presence of markets and alter desirability to shop in a formal market due to the increased advertising in larger cities.

Second, city size does not significantly affect the consumption share of BSF food. Stage et al. (2010) found that although urban households consumed greater shares of BSF food than rural households, this difference is not driven by urbanization but rather by other household variables such as income. The findings here echo Stage et al. (2010). The primary effect of city size on the consumption of BSF foods is via increased income, and as this effect is controlled for with household expenditure the remaining effect of city size is insignificant.

Third, further disaggregation of BSF foods does reveal multiple city size effects. Based on the consumption patterns of Table 1-6, it was expected that large cities would positively impact the consumption of wheat, staple vegetables, beef, dairy and food away from home, and negatively affect the consumption of maize and cassava. Table 1-12 shows that when controlling

¹⁸ Note that no city in Malawi is large enough to qualify as “primary”, resulting in seven, rather than eight, estimated coefficients.

for various consumption determinants, the impacts of city size on these food consumption patterns follow expectations with the exceptions of beef and dairy, with limited and contradicting statistically significant marginal effects. These estimated effects indicate that although income and city size do not uniformly affect all Bennett's Law foods, income and city size do affect the consumption of certain commodities.

Five additional effects of non-expenditure determinant variables are of note, with their marginal effects shown in Tables 1-A-2 and 1-A-3 of the appendix. First, consistent with literature that indicates that nonfarm employment increases the opportunity cost of time and thereby influences consumption patterns (Senauer et al., 1986; Kennedy and Reardon, 1996), the marginal effects of nonfarm employment are positive and significant in urban Tanzania and urban Zambia for both purchased share and share of BSF foods. This finding highlights the significant effects on nonfarm labor on rural households' decisions to engage with markets. Second, dependency ratio, another variable related to opportunity cost of time that would contribute to the consumer demand factor of individual and household preferences, generally has a negative effect on purchased share and BSF food. Third, the amount of land farmed negatively affects the shares of purchased food as it directly provides the opportunity to consume food from own production.

Fourth, ownership of a telephone has a strongly significant positive effect on purchase share and share of BSF foods in more than half of the national estimates. This finding is in agreement with the hypothesis that ownership of a phone reveals that individual and household preferences are more open to modern lifestyle patterns. Fifth, the effect of the age of the household head is negative and significant on purchased share, affirming that elderly households are less open to modern lifestyle patterns.

The fractional probit model estimates the direct effects of city size on household consumption patterns, but as discussed earlier, city size also affects other determinant variables of consumption such as household income. The mediation model is used to estimate the indirect effects of city size on consumption patterns through household total expenditure, nonfarm labor, the amount of land farmed by the household and the maximum education attained within the household. These three variables are selected as mediator variables for the following reasons: difference in city size could impact household expenditure, it could expose household members to various nonfarm employment opportunities, it could vary household access to farmland, and larger cities could offer greater opportunities for education. There were no statistically significant indirect marginal effects of city size found using the Malawi data, nor were there significant indirect effects via education, but the remaining countries produced the following three findings of note (Table 1-13).

Table 1-13: Indirect effects of urban household settlement within primary and secondary cities, estimates by country

Mediator Variable	Food Consumption Pattern	Malawi	Tanzania		Uganda		Zambia	
		Secondary Cities	Primary Cities	Secondary Cities	Primary Cities	Secondary Cities	Primary Cities	Secondary Cities
Household total expenditure per AE	Purchased Share	-0.004 (0.657)	0.056 (0.014)	-0.003 (0.689)	0.043 (0.085)	0.029 (0.480)	0.018 (0.000)	0.001 (0.568)
	Beyond Staple	-0.006 (0.546)	0.106 (0.000)	-0.005 (0.644)	0.050 (0.031)	0.034 (0.462)	0.018 (0.000)	0.001 (0.562)
	Foods items							
Nonfarm employment	Purchased Share	0.000 (0.650)	0.008 (0.008)	0.006 (0.086)	0.004 (0.497)	0.000 (0.857)	0.005 (0.000)	0.001 (0.010)
	Beyond Staple	0.000 (0.696)	0.006 (0.004)	0.005 (0.081)	0.003 (0.363)	0.000 (0.822)	0.002 (0.001)	0.000 (0.050)
	Foods items							
Hectares of farmed land (log)	Purchased Share	0.001 (0.321)	0.010 (0.001)	0.009 (0.016)	0.012 (0.011)	0.019 (0.007)	0.014 (0.000)	0.008 (0.000)
	Beyond Staple	0.000 (0.740)	0.002 (0.354)	0.002 (0.385)	0.003 (0.301)	0.005 (0.240)	0.001 (0.161)	0.001 (0.150)
	Foods items							
Maximum education attained within the household	Purchased Share	0.000 (0.599)	0.001 (0.517)	0.001 (0.507)	0.009 (0.158)	0.012 (0.279)	0.000 (0.844)	0.000 (0.747)
	Beyond Staple	-0.001 (0.218)	-0.001 (0.418)	-0.001 (0.421)	0.000 (0.847)	0.000 (0.866)	0.000 (0.854)	0.000 (0.776)
	Foods items							

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

First, primary cities have a positive indirect effect on the consumption of purchased and BSF food via total household expenditure. Household expenditure is positively correlated with these food consumption patterns and the mediation model confirms that primary cities positively impact household expenditure. The positive indirect effect of city size on BSF food is important, given that city size did not have a significant direct effect on BSF food, by indicating that the type of urbanization does, albeit indirectly, influence the consumption patterns of these foods as indicated by Satterthwaite et al. (2010).

Second, nonfarm employment is also a significant mediator variable in the estimation of indirect effects. As observed in Tanzania and Zambia, larger cities have a positive effect on nonfarm employment, which, via increased opportunity cost of time, increases household demand for purchased share and the consumption of BSF food. These indirect effects amplify the direct effects of city size on purchased food, while indicating a second avenue in which city size has an indirect effect on the consumption of BSF food.

The final indirect effect of urban household city size on food consumption patterns relates to urban farming – both farming in urban areas and farming by urban households on rural land. Rural households engage in farming more often than urban households, but urban households continue to farm. The percentage of households that engage in farming has an inverse relationship with city size (primary city 9%, secondary city 29%, tertiary city 55%)¹⁹ such that city size is a statistically significant determinant of the amount of land farmed by households. Primary and secondary cities have positive indirect effects on purchased share

¹⁹ Average land farmed by settlement type: Primary 0.5ha, Secondary 1.2ha, Tertiary 2.0ha, Rural 4.1ha.

through their negative effect on the amount of land farmed. The use of the mediator model highlights these three indirect effects that are not accounted for when testing for direct effects.

Both non-parametric and parametric analysis show that distance to city affects the consumption patterns of rural households. Figures 1-8 and 1-9 display the LOWESS curve estimates of the relationships between distance to city and the primary food consumption patterns. Each figure exhibits inverse relationships with declining marginal effects between distance to cities and the food consumption aggregates²⁰. Along with these negative relationships, it is shown that proximity to larger cities has a larger effect on the consumption patterns than distances to smaller cities. Both of these observations are consistent with previous findings of new economic geography in that increased distance from urban centers reduces the interaction of rural households with urban markets, and larger markets, which would be more prevalent in larger cities, offer a greater variety of marketed goods that increase the capabilities for households to diversify their consumption patterns (Renkow 2007).

²⁰ The anomalous relationship between distance to tertiary cities and the consumption of high-income Bennett's Law food is driven by a small sample of households. Seventy-five percent of households are within seventy-five kilometers of a tertiary city, and it is the furthest ten percent of households that are driving the positive relationship between BSF food and distance to city.

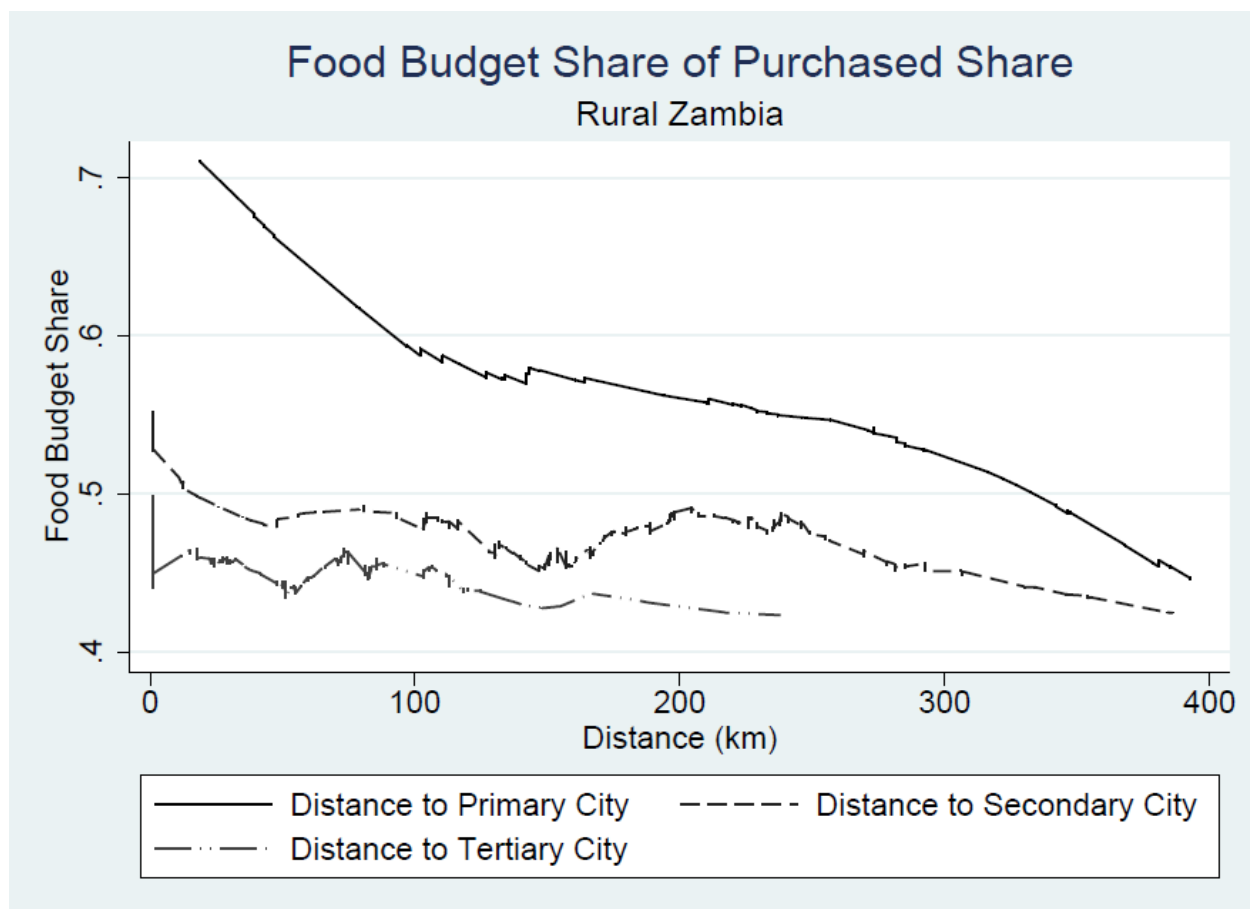


Figure 1-8: Locally weighted scatterplot smoothing curves representing the typical purchased share relative to household distance to city

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

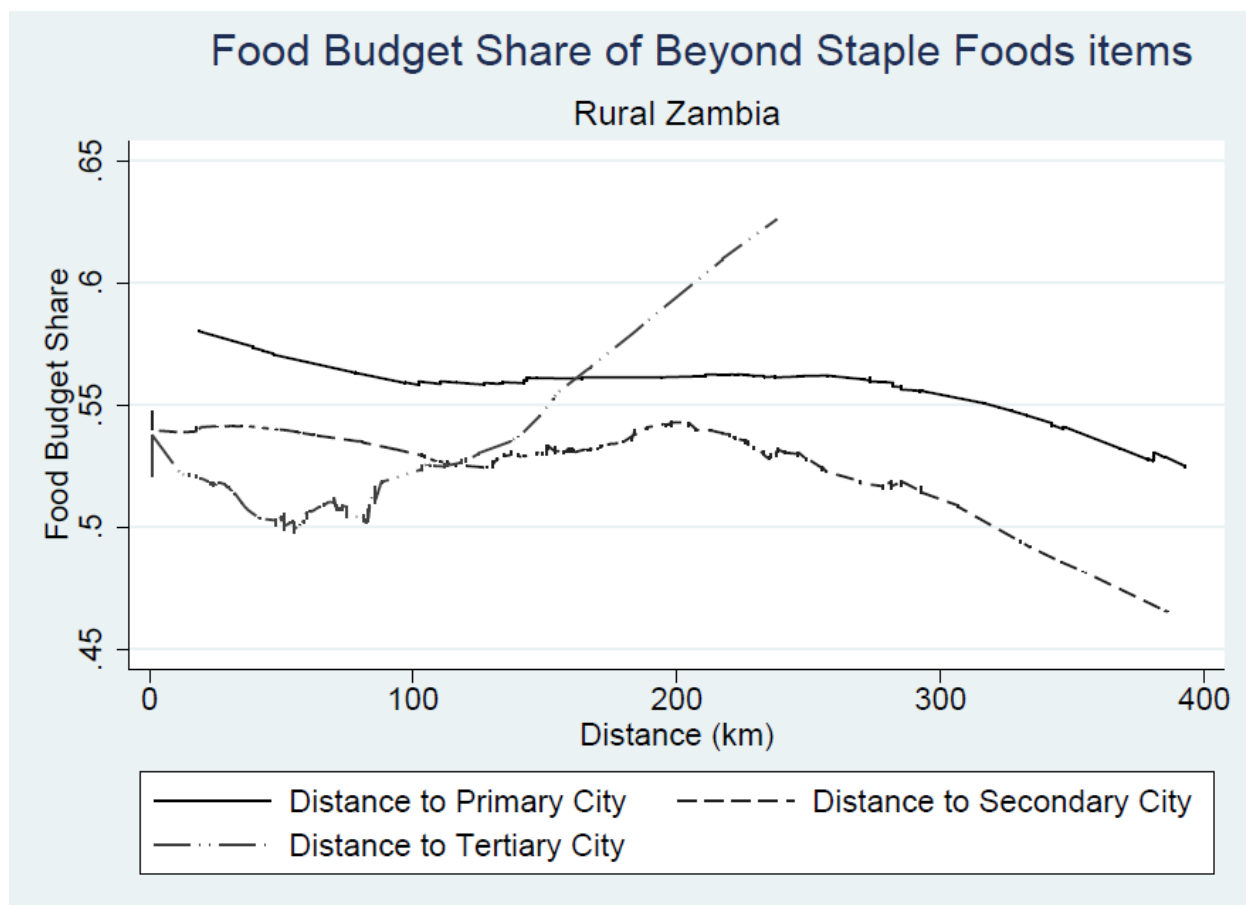


Figure 1-9: Locally weighted scatterplot smoothing curves representing the typical food budget shares of the consumption Beyond Staple Foods items relative to household distance to city

While the LOWESS analysis reveals correlations between the spatial effects and food consumption patterns, it does not control for the effects of other drivers of demand. Engel curve analysis provides the ability to estimate the effects of distance to city while controlling for multiple non-spatial determinant variables, leading to the observation of the following three patterns.

First, both of the key food consumption patterns are negatively impacted by distance to city (Table 1-14). Consistent with the von Thünen model, distance to markets increase transaction costs of acquiring the various foods from urban markets, limiting household ability to optimize their preferred bundle of foods to consume (Nelson, 2002). Owning a car reduces the

impact of the additional transaction costs, but with only one percent of rural households owning a car, the distance to urban areas is a significant limiting factor of consumption.

Table 1-14: Average partial effects of expenditure and distances to various city sizes for key food consumption patterns – rural Zambia

Rural Zambia	Purchased Share	Beyond Staple Foods items
Household total expenditure per AE	0.007 (0.150)	0.004 (0.269)
Distance to primary city (log)	-0.058 (0.000)	-0.024 (0.000)
Distance to secondary city (log)	-0.007 (0.118)	-0.012 (0.002)
Distance to tertiary city (log)	-0.008 (0.000)	-0.004 (0.020)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value

Second, the distance to a primary city has a far stronger impact on food consumption patterns than the distances to smaller cities (Table 1-14). Table 1-6 showed that households within primary cities consume greater shares of purchased food and BSF foods, so it follows that households near this city size would also have higher levels of the particular food consumption pattern. These findings are consistent with the hypothesized relationship between distance to primary cities and food consumption patterns that was based on the reach of commercial networks outside of primary cities.

Third, the pattern of average partial effects of distance on commodity consumption generally follows the expectations of Bennett's Law (Table 1-A-5). In particular, distance decreases the consumption of wheat, oils and fats, fresh vegetables, and beef, while increasing the consumption of cassava. Perhaps surprisingly, increased distance to cities decreases the consumption of maize. Cassava is commonly consumed in rural Zambia and could partially explain the negative effect of distance to cities on maize consumption. Results on dairy are also somewhat anomalous, with significant but opposite effects in primary and secondary cities.

Overall, however, the patterns of the relationships between distance to city and the consumption of commodities are as expected.

Indirect effects of distances to cities on rural food consumption patterns are estimated using total household expenditure, nonfarm employment, and the maximum attained education within the household as mediator variables. These variables are mediator variables as distance to cities could affect household expenditure, it could limit nonfarm employment opportunities that commonly occur in and around cities, and distance to cities could limit access to higher levels of education. The marginal effects of distances to cities on the three mediator variables generate indirect effects of the same sign for the distances to cities on the key measures of food consumption, as the marginal effects of these three mediator variables on the primary measures of food consumption are positive.

The overall finding from the mediation analysis is that increasing distance to cities decreases the primary food consumption patterns through indirect negative effects on household income, engagement in non-farm employment, and educational attainment (Table 1-15). This finding holds most clearly for the indirect effects of primary cities, where three of the six coefficients are negative and statistically significant.

Table 1-15: Indirect effects of distance from rural household settlement to primary, secondary or tertiary cities, estimates by country

Mediator Variable	Food Consumption Pattern	Distance to		
		Primary Cities	Secondary Cities	Tertiary Cities
Household total expenditure per AE	Purchased Share	-0.001 (0.043)	-0.001 (0.169)	0.000 (0.476)
	Beyond Staple Foods items	-0.001 (0.045)	0.000 (0.152)	0.000 (0.493)
Nonfarm employment	Purchased Share	-0.008 (0.000)	0.001 (0.447)	-0.001 (0.001)
	Beyond Staple Foods items	-0.001 (0.102)	0.000 (0.600)	0.000 (0.101)
Maximum education attained within the household	Purchased Share	0.000 (0.531)	0.000 (0.976)	0.000 (0.076)
	Beyond Staple Foods items	0.000 (0.518)	0.000 (0.977)	0.000 (0.117)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

1.5 Conclusions

I considered how the consumption patterns of purchased food and high-income Bennett's Law foods vary across income, by city size and by distances to cities of differing size. To my knowledge, this is the first paper to analyze the spatial effects of city size and distance to city on food consumption patterns in Africa. The analysis of the effects of household income and city size on the food consumption patterns of ESA are highlighted by the following results.

I find that the surprisingly high levels of purchased shares of food in total household consumed food value across ESA are a result of households transitioning food consumption towards purchased at subnumerary levels of income. This finding highlights that these transitions in consumption patterns are not solely a middle-class story, as conventionally assumed, but poor households are purchasing greater shares of their food and are consuming greater shares BSF foods as their incomes rise.

Nonparametric analysis shows purchased share is present in rural households at all levels of expenditure, to the extent that the fitted regressions of the share of purchased food do not fall below forty percent at any expenditure level in three of the four countries in this analysis.

Expenditure threshold analysis identifies that the consumption patterns of purchased share and the share of BSF foods are increasing below the poverty line. These significant initial levels and growth at low expenditure levels of purchased share indicate that poor households are currently able to access markets. Given the large portion of the population in ESA that currently has low income levels, coupled with the pattern of increased purchased share with income growth, this finding signals a strong future demand for increased food market infrastructure.

The Engel curve analysis highlighted many significant non-income determinants such as nonfarm employment that commonly occur at low income levels and in increasing magnitudes with rising incomes. These non-income determinants of consumption accentuate the income-induced patterns and contribute to the rise in the consumption patterns of purchased share and BSF foods at subnumerary levels of income.

Recognizing the level of market interaction and the extent of income-induced change in food consumption patterns across the entire income distribution, specifically among poor households, considerable pressure on the food systems in developing countries should be expected with future increases in household income.

The location of the future agrifood system investments is of due to the varying food demand across city size. Increasing city size has significant positive effects on purchased share in urban households, affirming the expectation that households will purchase more food when they have increased access to markets.

When controlling for other determinants of consumption, the consumption of BSF food is not shown to be positively affected by city size. Although, when considering the positive effects of city size on household expenditure and engagement in nonfarm employment, along with the positive effects that these determinant variables have on the consumption of BSF food, the

analysis shows that large cities indirectly have positive effects on the consumption of BSF food via these mediator variables that capture the effects of city size.

The distance to cities of any size has a diminishing negative effect on purchased share and the consumption of BSF foods by rural households. This relationship is consistent with the expectations that distance to urban centers will increase the cost of accessing urban markets, reducing the use of urban markets and resulting in lower shares of purchased or BSF food. The magnitude of the effect of distance to city varies by city size, with large cities having the strongest effects. Large cities are likely to have significant commercial networks that extend market reach further into rural areas, resulting in increased engagement with urban markets by rural households that live near large cities.

The spatial effects of city size and distance to city are shown in this analysis to be significant drivers of the food consumption patterns in Eastern and Southern Africa. As it is anticipated that urbanization will continue within this region and elsewhere in developing nations, the recognition of how urbanization occurs and where investments should be made should be considered when effort is put forth to meet the changing food demands of households within and near urban areas.

APPENDIX

APPENDIX

Table 1-A-1: Food budget shares of select commodity aggregates, country data by settlement

Malawi Data	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
National	28.6 (25.3)	2.7 (0.0)	2.7 (0.0)	3.2 (2.1)	10.2 (8.4)	1.3 (0.0)	1.0 (0.0)	1.3 (0.0)
Rural	30.6 (27.7)	2.1 (0.0)	2.9 (0.0)	2.6 (1.5)	10.2 (8.3)	1.0 (0.0)	0.8 (0.0)	1.1 (0.0)
Urban	17.8 (14.7)	6.2 (5.5)	1.4 (0.0)	6.6 (6.0)	10.6 (9.1)	3.2 (0.0)	2.5 (0.0)	2.1 (0.0)
Primary City								
Secondary City	17.5 (14.7)	6.3 (5.7)	1.2 (0.0)	6.9 (6.3)	10.9 (9.3)	3.3 (0.0)	2.5 (0.0)	2.1 (0.0)
Tertiary City	19.0 (14.7)	5.9 (5.1)	1.8 (0.0)	5.7 (4.8)	9.3 (8.5)	2.9 (0.0)	2.3 (0.0)	1.7 (0.0)

Tanzania Data	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
National	19.3 (15.2)	2.3 (0.0)	4.2 (0.0)	3.6 (3.0)	7.6 (6.2)	3.8 (0.0)	3.3 (0.0)	12.2 (2.9)
Rural	21.4 (18.2)	1.6 (0.0)	5.1 (0.0)	3.4 (2.7)	7.6 (6.0)	3.0 (0.0)	3.7 (0.0)	9.0 (0.0)
Urban	13.2 (8.8)	4.4 (3.4)	1.7 (0.0)	4.1 (3.6)	7.7 (6.5)	5.8 (4.7)	2.2 (0.0)	21.3 (15.2)
Primary City	6.8 (5.2)	5.6 (4.7)	0.8 (0.0)	3.6 (3.2)	7.4 (6.8)	6.4 (5.6)	1.7 (0.0)	33.0 (32.2)
Secondary City	12.9 (9.3)	4.7 (3.4)	1.5 (0.0)	4.7 (4.1)	7.6 (6.1)	6.4 (5.4)	2.0 (0.0)	18.1 (10.3)
Tertiary City	17.1 (12.6)	3.6 (2.0)	2.4 (0.0)	4.1 (3.6)	7.8 (6.6)	5.2 (2.4)	2.7 (0.0)	16.4 (7.4)

Table 1-A-1: (continued)

Uganda Data	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
National	9.3 (3.8)	1.3 (0.0)	10.5 (4.2)	2.1 (1.4)	3.1 (2.4)	4.3 (0.0)	3.5 (0.0)	5.2 (0.0)
Rural	10.0 (4.2)	0.8 (0.0)	12.3 (6.0)	2.0 (1.1)	2.8 (2.1)	3.7 (0.0)	3.3 (0.0)	3.3 (0.0)
Urban	6.6 (3.4)	3.2 (0.0)	4.1 (0.6)	2.8 (2.5)	4.2 (3.6)	6.2 (2.9)	4.3 (0.0)	11.8 (0.0)
Primary City	6.3 (3.2)	4.8 (3.5)	2.2 (0.0)	2.9 (2.8)	4.9 (4.3)	6.7 (5.3)	5.0 (3.1)	15.0 (0.0)
Secondary City	4.7 (2.7)	3.0 (0.0)	5.4 (0.0)	2.8 (2.9)	4.7 (4.2)	8.9 (8.6)	4.4 (3.1)	11.3 (0.0)
Tertiary City	7.0 (3.4)	2.6 (0.0)	4.7 (0.6)	2.8 (2.4)	3.8 (3.2)	5.7 (0.0)	4.0 (0.0)	10.7 (0.0)

Zambia Data	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
National	22.6 (17.9)	5.3 (1.7)	6.0 (0.0)	4.5 (3.6)	4.7 (3.6)	2.4 (0.0)	1.6 (0.0)	0.2 (0.0)
Rural	25.9 (21.8)	3.3 (0.0)	8.8 (0.0)	4.0 (2.9)	3.8 (2.5)	1.5 (0.0)	1.2 (0.0)	0.1 (0.0)
Urban	16.4 (13.5)	9.2 (7.3)	0.7 (0.0)	5.3 (4.5)	6.4 (5.3)	4.1 (0.0)	2.4 (0.7)	0.5 (0.0)
Primary City	14.6 (12.3)	9.9 (8.6)	0.1 (0.0)	4.6 (3.9)	6.6 (5.4)	4.9 (4.2)	3.0 (1.8)	0.8 (0.0)
Secondary City	17.3 (14.4)	9.4 (7.5)	0.3 (0.0)	5.5 (4.8)	6.4 (5.3)	3.8 (0.0)	2.4 (0.6)	0.4 (0.0)
Tertiary City	17.1 (13.7)	8.0 (5.4)	2.0 (0.0)	5.8 (4.8)	6.3 (5.1)	3.8 (0.0)	1.8 (0.0)	0.2 (0.0)

Table 1-A-2: Average partial effects of the household determinants for key food consumption patterns, urban household estimates by country

Urban	Purchased Share				High-Income Bennett's Law Food			
	Malawi	Tanzania	Uganda	Zambia	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.001 (0.756)	0.013 (0.001)	0.004 (0.296)	0.005 (0.296)	0.008 (0.000)	0.025 (0.000)	0.015 (0.000)	0.009 (0.000)
Household located in a primary city		0.090 (0.000)	0.063 (0.048)	0.037 (0.000)		0.017 (0.172)	-0.029 (0.130)	-0.010 (0.058)
Household located in a secondary city	0.090 (0.000)	0.051 (0.000)	0.036 (0.349)	0.022 (0.000)	0.007 (0.445)	-0.017 (0.208)	-0.035 (0.229)	-0.008 (0.087)
Nonfarm employment	-0.012 (0.639)	0.110 (0.000)	0.036 (0.419)	0.077 (0.000)	-0.007 (0.665)	0.089 (0.000)	0.025 (0.397)	0.030 (0.000)
Dependency ratio	-0.023 (0.374)	-0.026 (0.293)	0.132 (0.062)	-0.023 (0.009)	-0.043 (0.028)	-0.128 (0.000)	-0.217 (0.000)	-0.026 (0.006)
Household adult equivalents	0.000 (0.919)	0.003 (0.099)	-0.008 (0.215)	0.004 (0.000)	-0.002 (0.426)	0.009 (0.000)	-0.009 (0.014)	0.001 (0.375)
Farmed hectares of land (log)	-0.078 (0.004)	-0.037 (0.000)	-0.099 (0.000)	-0.028 (0.000)	0.012 (0.617)	-0.009 (0.348)	-0.026 (0.151)	-0.003 (0.205)
Own a gas or electric stove	0.049 (0.000)	-0.007 (0.740)		0.006 (0.201)	0.032 (0.052)	0.001 (0.977)		0.003 (0.587)
Own a refrigerator	0.038 (0.005)	0.040 (0.008)		0.007 (0.152)	0.027 (0.054)	0.015 (0.178)		-0.006 (0.380)
Maximum education attained within the household	0.003 (0.503)	0.000 (0.565)	0.004 (0.000)	0.002 (0.011)	0.007 (0.028)	-0.001 (0.423)	0.000 (0.823)	0.001 (0.082)
Age of household head	-0.001 (0.001)	-0.001 (0.002)	-0.001 (0.286)	-0.001 (0.000)	0.000 (0.918)	-0.001 (0.001)	-0.001 (0.244)	0.000 (0.010)
Female head of household	-0.022 (0.184)	0.009 (0.592)	0.049 (0.107)	-0.019 (0.000)	-0.003 (0.763)	0.002 (0.864)	-0.037 (0.072)	-0.005 (0.278)
Own a telephone	0.037 (0.018)	0.113 (0.000)	0.070 (0.088)	0.013 (0.011)	0.003 (0.749)	-0.002 (0.928)	0.010 (0.689)	0.000 (0.940)
Own a car	0.059 (0.002)	-0.027 (0.334)	0.057 (0.218)	0.000 (0.952)	0.051 (0.004)	-0.043 (0.024)	0.010 (0.737)	0.025 (0.001)
Own a motorcycle	0.027 (0.565)	0.018 (0.352)	0.045 (0.337)	-0.015 (0.428)	-0.003 (0.916)	-0.013 (0.521)	-0.032 (0.166)	-0.035 (0.145)
Own a bicycle	-0.082 (0.000)	-0.038 (0.014)	-0.106 (0.002)	-0.030 (0.000)	-0.003 (0.748)	0.009 (0.492)	-0.053 (0.018)	-0.012 (0.018)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 1-A-3: Average partial effects of the household determinants for key food consumption patterns, rural household estimates by country

Rural	Purchased Share				High-Income Bennett's Law Food			
	Malawi	Tanzania	Uganda	Zambia	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	-0.015 (0.000)	0.004 (0.525)	0.007 (0.214)	0.007 (0.150)	0.050 (0.000)	0.040 (0.000)	0.023 (0.000)	0.004 (0.269)
Distance to primary city (log)				-0.058 (0.000)				-0.024 (0.000)
Distance to secondary city (log)				-0.007 (0.118)				-0.012 (0.002)
Distance to tertiary city (log)				-0.008 (0.000)				-0.004 (0.020)
Nonfarm employment	0.138 (0.000)	0.175 (0.000)	0.130 (0.000)	0.314 (0.000)	0.013 (0.387)	0.053 (0.000)	0.097 (0.000)	0.035 (0.080)
Dependency ratio	-0.070 (0.000)	-0.091 (0.003)	-0.046 (0.195)	-0.052 (0.009)	-0.017 (0.191)	-0.083 (0.000)	-0.130 (0.000)	0.014 (0.372)
Household adult equivalents	-0.011 (0.000)	-0.004 (0.005)	-0.001 (0.827)	0.001 (0.730)	0.004 (0.113)	0.005 (0.000)	-0.010 (0.000)	-0.004 (0.029)
Farmed hectares of land (log)	-0.042 (0.008)	-0.060 (0.000)	-0.057 (0.000)	-0.031 (0.000)	0.009 (0.443)	-0.018 (0.003)	0.018 (0.008)	0.008 (0.017)
Own a gas or electric stove	0.051 (0.392)	0.107 (0.086)		0.116 (0.000)	0.054 (0.085)	-0.025 (0.357)		0.029 (0.258)
Own a refrigerator	0.263 (0.000)	0.118 (0.004)		0.014 (0.722)	-0.032 (0.255)	-0.017 (0.469)		0.001 (0.976)
Maximum education attained within the household	0.017 (0.000)	0.002 (0.061)	0.001 (0.470)	0.009 (0.000)	0.001 (0.721)	0.000 (0.593)	0.000 (0.648)	0.004 (0.000)
Age of household head	-0.002 (0.000)	0.000 (0.929)	-0.002 (0.003)	-0.002 (0.000)	0.000 (0.517)	0.000 (0.301)	0.000 (0.519)	0.000 (0.559)
Female head of household	-0.026 (0.020)	0.003 (0.888)	-0.014 (0.423)	-0.013 (0.249)	-0.020 (0.007)	-0.016 (0.167)	-0.007 (0.527)	-0.011 (0.159)
Own a telephone	0.045 (0.000)	0.087 (0.000)	0.061 (0.000)	0.065 (0.000)	0.010 (0.140)	0.039 (0.000)	0.031 (0.006)	0.011 (0.159)
Distance to market (log)	-0.017 (0.000)	0.004 (0.517)	-0.003 (0.656)	-0.018 (0.000)	-0.002 (0.411)	-0.002 (0.527)	0.004 (0.360)	0.000 (0.966)
Own a car	0.125 (0.043)	0.070 (0.262)	0.128 (0.060)	0.132 (0.006)	0.033 (0.193)	-0.022 (0.447)	0.010 (0.849)	0.103 (0.015)
Own a motorcycle	0.010 (0.752)	0.031 (0.426)	0.030 (0.295)	-0.022 (0.745)	0.063 (0.049)	0.003 (0.903)	0.023 (0.203)	0.044 (0.225)
Own a bicycle	0.007 (0.420)	-0.020 (0.183)	-0.055 (0.000)	0.008 (0.389)	0.014 (0.024)	-0.011 (0.235)	0.011 (0.280)	0.008 (0.266)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 1-A-4: Average partial effects of the household determinants for select commodity aggregates, urban household estimates by country

Urban Malawi	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.006 (0.008)	0.004 (0.000)	-0.001 (0.170)	0.001 (0.134)	-0.007 (0.000)	0.008 (0.002)	0.004 (0.006)	0.000 (0.200)
Household located in a secondary city	0.003 (0.730)	-0.003 (0.380)	-0.004 (0.019)	0.010 (0.001)	0.020 (0.000)	-0.001 (0.840)	-0.001 (0.588)	0.003 (0.167)
Nonfarm employment	-0.027 (0.040)	0.015 (0.020)	-0.002 (0.525)	-0.006 (0.324)	-0.009 (0.161)	0.000 (0.968)	0.001 (0.690)	0.000 (0.917)
Dependency ratio	0.031 (0.074)	-0.010 (0.240)	-0.004 (0.229)	-0.002 (0.732)	0.011 (0.159)	-0.003 (0.679)	0.006 (0.306)	-0.014 (0.015)
Household adult equivalents	-0.001 (0.787)	0.002 (0.208)	0.000 (0.938)	-0.001 (0.134)	-0.007 (0.000)	0.003 (0.001)	0.000 (0.717)	0.000 (0.921)
Farmed hectares of land (log)	-0.028 (0.177)	-0.008 (0.257)	-0.003 (0.514)	-0.010 (0.282)	0.015 (0.214)	-0.001 (0.866)	-0.003 (0.508)	-0.013 (0.113)
Own a gas or electric stove	-0.039 (0.000)	0.004 (0.409)	-0.007 (0.000)	0.002 (0.660)	0.014 (0.011)	0.009 (0.073)	0.010 (0.078)	-0.002 (0.504)
Own a refrigerator	-0.020 (0.049)	-0.001 (0.799)	0.003 (0.454)	0.013 (0.010)	0.004 (0.517)	0.002 (0.717)	0.005 (0.220)	-0.009 (0.001)
Maximum education attained within the household	-0.011 (0.000)	0.001 (0.462)	0.000 (0.547)	0.002 (0.058)	-0.001 (0.582)	0.003 (0.025)	0.001 (0.173)	-0.002 (0.014)
Age of household head	0.000 (0.143)	0.000 (0.275)	0.000 (0.965)	0.000 (0.335)	0.000 (0.420)	0.000 (0.388)	0.000 (0.535)	0.000 (0.251)
Female head of household	-0.019 (0.061)	-0.009 (0.069)	-0.003 (0.138)	0.010 (0.096)	-0.002 (0.749)	-0.005 (0.284)	-0.001 (0.731)	-0.001 (0.819)
Own a telephone	-0.010 (0.253)	0.009 (0.069)	-0.006 (0.013)	0.005 (0.292)	0.016 (0.000)	-0.001 (0.845)	0.010 (0.001)	0.004 (0.172)
Own a car	-0.047 (0.006)	0.008 (0.261)	-0.003 (0.179)	0.002 (0.614)	0.019 (0.021)	0.003 (0.582)	-0.001 (0.698)	0.008 (0.249)
Own a motorcycle	-0.037 (0.280)	0.037 (0.079)	0.000 (0.990)	0.005 (0.653)	0.021 (0.276)	-0.010 (0.134)	-0.006 (0.023)	-0.005 (0.512)
Own a bicycle	0.006 (0.423)	-0.015 (0.000)	0.000 (0.837)	-0.002 (0.669)	0.005 (0.205)	0.000 (0.920)	0.004 (0.203)	-0.002 (0.438)

Table 1-A-4: (continued)

Urban Tanzania	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.023 (0.000)	0.000 (0.850)	-0.003 (0.009)	-0.002 (0.000)	-0.007 (0.000)	0.002 (0.004)	0.001 (0.044)	0.020 (0.000)
Household located in a primary city	-0.026 (0.001)	0.017 (0.000)	-0.003 (0.383)	-0.002 (0.502)	0.019 (0.000)	-0.001 (0.867)	-0.013 (0.000)	0.037 (0.071)
Household located in a secondary city	0.002 (0.820)	0.008 (0.114)	-0.003 (0.557)	0.004 (0.152)	0.003 (0.537)	0.005 (0.531)	-0.010 (0.004)	-0.011 (0.584)
Nonfarm employment	-0.070 (0.000)	-0.003 (0.662)	-0.009 (0.055)	0.000 (0.941)	-0.013 (0.045)	-0.008 (0.378)	-0.019 (0.001)	0.125 (0.000)
Dependency ratio	0.023 (0.175)	0.017 (0.030)	0.013 (0.069)	0.008 (0.135)	0.011 (0.134)	0.012 (0.312)	0.006 (0.426)	-0.232 (0.000)
Household adult equivalents	-0.007 (0.000)	0.000 (0.836)	0.000 (0.156)	-0.001 (0.001)	-0.003 (0.000)	0.002 (0.000)	0.000 (0.273)	0.010 (0.000)
Farmed hectares of land (log)	0.014 (0.033)	-0.006 (0.100)	-0.001 (0.648)	0.000 (0.992)	-0.003 (0.421)	0.000 (0.989)	-0.001 (0.685)	-0.001 (0.960)
Own a gas or electric stove	0.006 (0.682)	0.000 (0.989)	0.001 (0.885)	-0.001 (0.823)	0.005 (0.501)	-0.002 (0.800)	0.015 (0.027)	-0.018 (0.474)
Own a refrigerator	-0.023 (0.028)	-0.005 (0.232)	-0.006 (0.041)	0.000 (0.907)	0.004 (0.448)	0.010 (0.160)	0.009 (0.025)	-0.025 (0.144)
Maximum education attained within the household	0.000 (0.456)	0.000 (0.260)	0.000 (0.048)	0.000 (0.791)	0.000 (0.869)	0.000 (0.539)	0.000 (0.348)	-0.002 (0.078)
Age of household head	0.001 (0.008)	0.000 (0.492)	0.000 (0.018)	0.000 (0.026)	0.000 (0.117)	0.000 (0.043)	0.000 (0.052)	0.000 (0.472)
Female head of household	-0.022 (0.020)	0.002 (0.670)	0.000 (0.987)	0.005 (0.085)	0.005 (0.212)	0.004 (0.506)	-0.002 (0.526)	-0.033 (0.048)
Own a telephone	-0.005 (0.688)	0.016 (0.002)	-0.003 (0.572)	0.012 (0.001)	0.000 (0.941)	0.012 (0.179)	0.000 (0.974)	-0.025 (0.464)
Own a car	0.012 (0.384)	0.007 (0.304)	-0.009 (0.014)	0.015 (0.063)	0.025 (0.007)	-0.008 (0.318)	0.009 (0.174)	-0.062 (0.018)
Own a motorcycle	-0.018 (0.148)	0.002 (0.788)	0.004 (0.447)	0.001 (0.844)	0.012 (0.133)	-0.003 (0.771)	-0.001 (0.813)	-0.062 (0.035)
Own a bicycle	-0.012 (0.197)	-0.006 (0.147)	0.005 (0.181)	-0.003 (0.169)	-0.003 (0.498)	0.008 (0.199)	0.008 (0.039)	-0.016 (0.348)

Table 1-A-4: (continued)

Urban Uganda	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.010 (0.000)	0.001 (0.012)	-0.004 (0.013)	0.000 (0.526)	-0.001 (0.129)	0.003 (0.001)	0.002 (0.011)	0.004 (0.004)
Household located in a primary city	0.014 (0.154)	0.011 (0.067)	-0.010 (0.129)	-0.001 (0.765)	0.010 (0.010)	-0.003 (0.721)	0.007 (0.320)	-0.002 (0.928)
Household located in a secondary city	-0.006 (0.618)	0.001 (0.945)	0.019 (0.228)	0.000 (0.949)	0.008 (0.089)	0.021 (0.150)	0.002 (0.831)	-0.034 (0.139)
Nonfarm employment	-0.002 (0.853)	0.007 (0.408)	0.001 (0.961)	0.000 (0.913)	-0.002 (0.678)	0.003 (0.783)	-0.007 (0.450)	0.039 (0.208)
Dependency ratio	0.024 (0.275)	0.015 (0.140)	0.033 (0.080)	0.005 (0.367)	-0.005 (0.503)	0.029 (0.107)	0.002 (0.877)	-0.223 (0.000)
Household adult equivalents	0.001 (0.704)	0.001 (0.572)	-0.001 (0.682)	-0.001 (0.110)	-0.002 (0.001)	0.000 (0.864)	-0.002 (0.089)	-0.003 (0.323)
Farmed hectares of land (log)	0.015 (0.028)	-0.006 (0.172)	0.016 (0.002)	-0.006 (0.054)	-0.007 (0.044)	0.008 (0.277)	0.010 (0.121)	-0.022 (0.409)
Maximum education attained within the household	0.000 (0.049)	0.000 (0.114)	0.000 (0.552)	0.000 (0.011)	0.000 (0.353)	0.001 (0.007)	0.000 (0.943)	-0.001 (0.020)
Age of household head	0.000 (0.355)	0.000 (0.283)	0.000 (0.435)	0.000 (0.844)	0.000 (0.674)	0.000 (0.923)	0.000 (0.437)	0.000 (0.830)
Female head of household	0.015 (0.139)	0.005 (0.455)	0.002 (0.756)	0.005 (0.050)	0.010 (0.006)	-0.003 (0.753)	-0.008 (0.225)	-0.059 (0.003)
Own a telephone	-0.014 (0.258)	0.012 (0.041)	0.001 (0.947)	0.000 (0.979)	-0.001 (0.810)	-0.004 (0.731)	-0.001 (0.960)	0.057 (0.011)
Own a car	-0.009 (0.431)	-0.001 (0.909)	-0.017 (0.035)	-0.005 (0.301)	0.009 (0.125)	0.007 (0.578)	0.005 (0.678)	-0.024 (0.434)
Own a motorcycle	-0.001 (0.909)	0.005 (0.459)	-0.016 (0.027)	0.004 (0.335)	0.002 (0.550)	-0.016 (0.065)	0.013 (0.127)	-0.050 (0.006)
Own a bicycle	-0.013 (0.158)	-0.014 (0.002)	0.018 (0.051)	0.004 (0.184)	0.004 (0.269)	-0.009 (0.247)	-0.003 (0.681)	-0.044 (0.030)

Table 1-A-4: (continued)

Urban Zambia	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.013 (0.000)	0.003 (0.000)	-0.003 (0.000)	-0.002 (0.000)	-0.003 (0.000)	0.003 (0.000)	0.002 (0.000)	0.000 (0.000)
Household located in a primary city	0.011 (0.004)	0.012 (0.001)	-0.007 (0.000)	-0.011 (0.000)	0.007 (0.000)	-0.002 (0.414)	0.006 (0.001)	0.002 (0.017)
Household located in a secondary city	0.015 (0.000)	0.010 (0.002)	-0.009 (0.000)	-0.003 (0.028)	0.003 (0.068)	-0.006 (0.006)	0.004 (0.016)	0.002 (0.056)
Nonfarm employment	-0.025 (0.000)	0.008 (0.142)	-0.006 (0.000)	-0.002 (0.456)	0.004 (0.180)	0.006 (0.048)	0.002 (0.595)	0.004 (0.014)
Dependency ratio	0.012 (0.102)	0.005 (0.435)	0.001 (0.586)	0.000 (0.872)	-0.010 (0.003)	0.000 (0.929)	0.005 (0.077)	-0.008 (0.000)
Household adult equivalents	-0.003 (0.000)	0.004 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.002 (0.000)	0.002 (0.033)	0.001 (0.002)	0.001 (0.004)
Farmed hectares of land (log)	-0.002 (0.136)	-0.002 (0.160)	0.002 (0.000)	-0.001 (0.363)	-0.002 (0.003)	-0.002 (0.014)	0.000 (0.434)	-0.001 (0.106)
Own a gas or electric stove	-0.006 (0.199)	-0.001 (0.749)	-0.002 (0.007)	0.007 (0.000)	0.001 (0.623)	-0.001 (0.741)	0.002 (0.373)	-0.001 (0.505)
Own a refrigerator	-0.005 (0.267)	0.000 (0.895)	0.001 (0.223)	0.004 (0.052)	-0.001 (0.700)	0.000 (0.888)	0.001 (0.471)	-0.004 (0.040)
Maximum education attained within the household	-0.001 (0.008)	0.000 (0.599)	-0.001 (0.000)	0.000 (0.213)	0.000 (0.908)	0.000 (0.110)	0.000 (0.513)	0.000 (0.174)
Age of household head	0.001 (0.000)	0.000 (0.082)	0.000 (0.173)	0.000 (0.479)	0.000 (0.622)	0.000 (0.786)	0.000 (0.036)	0.000 (0.000)
Female head of household	0.003 (0.457)	0.000 (0.919)	0.000 (0.664)	0.001 (0.402)	0.001 (0.754)	0.001 (0.576)	0.000 (0.794)	-0.001 (0.083)
Own a telephone	-0.007 (0.111)	0.013 (0.001)	-0.003 (0.000)	0.001 (0.678)	-0.004 (0.097)	0.005 (0.050)	0.004 (0.074)	0.001 (0.475)
Own a car	-0.011 (0.047)	-0.006 (0.234)	0.003 (0.251)	0.006 (0.003)	0.003 (0.210)	0.008 (0.004)	0.003 (0.172)	0.001 (0.347)
Own a motorcycle	0.027 (0.139)	0.022 (0.173)	-0.002 (0.208)	0.010 (0.195)	-0.002 (0.710)	-0.012 (0.046)	-0.005 (0.260)	-0.002 (0.061)
Own a bicycle	0.000 (0.915)	0.008 (0.022)	0.002 (0.008)	0.000 (0.805)	-0.001 (0.518)	-0.004 (0.050)	-0.002 (0.103)	-0.001 (0.397)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 1-A-5: Average partial effects of the household determinants for select commodity aggregates, rural household estimates by country

Rural Malawi	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.056 (0.000)	0.029 (0.000)	0.000 (0.685)	0.015 (0.000)	-0.021 (0.000)	0.040 (0.001)	0.048 (0.000)	0.003 (0.008)
Nonfarm employment	-0.042 (0.003)	0.009 (0.001)	0.001 (0.921)	0.010 (0.000)	-0.009 (0.104)	0.001 (0.685)	0.000 (0.997)	0.003 (0.137)
Dependency ratio	0.011 (0.365)	0.001 (0.773)	-0.008 (0.078)	-0.005 (0.019)	-0.009 (0.082)	-0.001 (0.783)	0.000 (0.893)	-0.002 (0.319)
Household adult equivalents	-0.007 (0.001)	0.000 (0.174)	0.002 (0.067)	0.000 (0.840)	-0.009 (0.000)	0.001 (0.014)	0.001 (0.001)	0.000 (0.416)
Farmed hectares of land (log)	0.002 (0.891)	0.002 (0.324)	-0.026 (0.000)	-0.002 (0.407)	0.004 (0.428)	0.003 (0.114)	0.003 (0.096)	0.000 (0.753)
Own a gas or electric stove	0.010 (0.666)	-0.004 (0.148)	0.004 (0.705)	0.014 (0.067)	0.038 (0.119)	0.007 (0.483)	0.005 (0.408)	0.001 (0.853)
Own a refrigerator	-0.006 (0.777)	0.003 (0.407)	-0.006 (0.383)	0.013 (0.015)	-0.006 (0.702)	0.015 (0.186)	-0.002 (0.429)	-0.001 (0.905)
Maximum education attained within the household	-0.003 (0.289)	0.001 (0.015)	0.001 (0.434)	0.003 (0.000)	0.002 (0.033)	0.001 (0.024)	0.002 (0.004)	0.000 (0.427)
Age of household head	0.000 (0.430)	0.000 (0.001)	0.000 (0.781)	0.000 (0.018)	0.000 (0.490)	0.000 (0.526)	0.000 (0.003)	0.000 (0.000)
Female head of household	0.013 (0.069)	-0.001 (0.447)	0.004 (0.138)	-0.001 (0.538)	0.004 (0.210)	0.001 (0.730)	-0.001 (0.727)	-0.001 (0.174)
Own a telephone	-0.007 (0.296)	0.006 (0.000)	-0.002 (0.463)	0.007 (0.000)	0.002 (0.464)	0.001 (0.421)	-0.001 (0.452)	0.002 (0.046)
Distance to market (log)	-0.004 (0.087)	0.000 (0.346)	0.003 (0.015)	-0.001 (0.080)	-0.002 (0.011)	0.001 (0.103)	0.000 (0.365)	0.000 (0.368)
Own a car	-0.002 (0.965)	0.010 (0.200)	-0.026 (0.000)	0.003 (0.457)	0.003 (0.906)	-0.006 (0.000)	0.014 (0.165)	0.001 (0.892)
Own a motorcycle	-0.012 (0.486)	0.001 (0.842)	-0.013 (0.025)	0.000 (0.974)	-0.012 (0.255)	-0.002 (0.702)	0.005 (0.397)	-0.002 (0.414)
Own a bicycle	-0.011 (0.052)	0.004 (0.004)	-0.003 (0.186)	0.004 (0.000)	0.006 (0.028)	0.001 (0.333)	0.002 (0.119)	-0.001 (0.308)

Table 1-A-5: (continued)

Rural Tanzania	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.037 (0.000)	0.004 (0.000)	-0.017 (0.000)	-0.002 (0.010)	-0.015 (0.000)	0.005 (0.001)	0.005 (0.003)	0.024 (0.000)
Nonfarm employment	-0.034 (0.017)	0.010 (0.000)	-0.008 (0.321)	-0.001 (0.774)	-0.002 (0.620)	-0.002 (0.619)	-0.039 (0.000)	0.085 (0.000)
Dependency ratio	0.012 (0.557)	-0.006 (0.119)	0.022 (0.095)	-0.006 (0.147)	0.001 (0.879)	-0.007 (0.329)	0.010 (0.407)	-0.077 (0.000)
Household adult equivalents	-0.003 (0.000)	0.000 (0.406)	-0.001 (0.035)	-0.001 (0.001)	-0.003 (0.000)	0.000 (0.588)	0.001 (0.002)	0.003 (0.000)
Farmed hectares of land (log)	0.035 (0.000)	-0.002 (0.073)	-0.005 (0.185)	-0.001 (0.214)	0.000 (0.962)	-0.002 (0.258)	0.006 (0.019)	0.003 (0.534)
Own a gas or electric stove	-0.023 (0.351)	-0.001 (0.809)	-0.004 (0.752)	0.003 (0.341)	0.020 (0.072)	-0.002 (0.804)	-0.005 (0.743)	-0.019 (0.252)
Own a refrigerator	-0.068 (0.003)	0.023 (0.004)	-0.010 (0.264)	-0.003 (0.319)	-0.002 (0.694)	0.011 (0.223)	0.025 (0.258)	-0.034 (0.001)
Maximum education attained within the household	0.000 (0.958)	0.000 (0.331)	0.000 (0.247)	0.000 (0.025)	0.000 (0.113)	0.000 (0.288)	-0.001 (0.002)	0.000 (0.821)
Age of household head	-0.001 (0.012)	0.000 (0.736)	0.000 (0.155)	0.000 (0.026)	0.000 (0.036)	0.000 (0.104)	0.000 (0.187)	0.000 (0.340)
Female head of household	0.005 (0.649)	0.003 (0.235)	-0.002 (0.705)	0.003 (0.128)	0.003 (0.395)	0.004 (0.348)	0.013 (0.040)	-0.018 (0.061)
Own a telephone	-0.002 (0.863)	0.004 (0.028)	-0.011 (0.049)	0.011 (0.000)	0.001 (0.684)	0.008 (0.016)	0.014 (0.002)	0.003 (0.718)
Distance to market (log)	0.004 (0.293)	-0.001 (0.337)	-0.007 (0.005)	0.001 (0.150)	0.003 (0.030)	0.000 (0.725)	-0.003 (0.118)	-0.004 (0.141)
Own a car	0.009 (0.730)	-0.006 (0.016)	-0.021 (0.010)	0.003 (0.551)	0.041 (0.023)	0.007 (0.562)	-0.016 (0.137)	-0.005 (0.802)
Own a motorcycle	-0.037 (0.070)	0.001 (0.705)	-0.014 (0.146)	-0.003 (0.415)	0.003 (0.628)	0.010 (0.220)	0.002 (0.809)	0.007 (0.686)
Own a bicycle	0.010 (0.310)	0.001 (0.428)	0.000 (0.989)	-0.003 (0.147)	-0.003 (0.287)	-0.004 (0.301)	0.001 (0.803)	-0.015 (0.060)

Table 1-A-5: (continued)

Rural Uganda	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.009 (0.001)	0.004 (0.000)	-0.024 (0.000)	0.001 (0.026)	0.000 (0.885)	0.008 (0.000)	0.005 (0.000)	0.007 (0.000)
Nonfarm employment	0.027 (0.016)	0.001 (0.618)	-0.047 (0.001)	0.000 (0.948)	0.004 (0.119)	0.016 (0.003)	-0.003 (0.580)	0.024 (0.000)
Dependency ratio	0.059 (0.003)	0.000 (0.975)	0.048 (0.026)	-0.003 (0.408)	-0.014 (0.004)	0.008 (0.331)	-0.013 (0.179)	-0.048 (0.000)
Household adult equivalents	-0.002 (0.379)	0.000 (0.338)	0.003 (0.106)	0.000 (0.260)	0.000 (0.277)	0.000 (0.700)	-0.001 (0.026)	-0.003 (0.034)
Farmed hectares of land (log)	-0.002 (0.796)	-0.001 (0.283)	0.010 (0.067)	0.001 (0.653)	-0.006 (0.000)	-0.003 (0.283)	0.010 (0.000)	0.000 (0.927)
Maximum education attained within the household	-0.001 (0.012)	0.000 (0.000)	0.000 (0.308)	0.000 (0.315)	0.000 (0.758)	0.000 (0.231)	0.000 (0.909)	0.000 (0.462)
Age of household head	0.000 (0.214)	0.000 (0.742)	-0.001 (0.062)	0.000 (0.009)	0.000 (0.228)	0.000 (0.456)	0.000 (0.275)	0.000 (0.331)
Female head of household	-0.012 (0.169)	0.003 (0.158)	-0.002 (0.845)	0.002 (0.325)	0.004 (0.073)	-0.009 (0.042)	0.000 (0.980)	-0.013 (0.065)
Own a telephone	0.008 (0.349)	0.003 (0.020)	-0.012 (0.217)	0.004 (0.031)	0.005 (0.010)	0.003 (0.467)	0.019 (0.000)	0.002 (0.649)
Distance to market (log)	-0.002 (0.466)	-0.001 (0.041)	0.001 (0.779)	0.000 (0.384)	-0.002 (0.018)	0.000 (0.847)	0.003 (0.020)	-0.005 (0.048)
Own a car	-0.016 (0.518)	0.000 (0.919)	-0.044 (0.005)	0.005 (0.445)	0.006 (0.511)	0.000 (0.972)	0.035 (0.081)	-0.024 (0.000)
Own a motorcycle	-0.011 (0.404)	0.000 (0.796)	-0.026 (0.038)	0.005 (0.150)	0.008 (0.009)	0.005 (0.369)	0.010 (0.143)	0.001 (0.933)
Own a bicycle	-0.012 (0.127)	-0.002 (0.047)	0.027 (0.003)	0.003 (0.038)	0.003 (0.104)	-0.002 (0.554)	-0.003 (0.422)	-0.007 (0.186)

Table 1-A-5: (continued)

Rural Zambia	Staple Foods items			Beyond Staple Foods items				
	Maize	Wheat	Cassava	Oils & Fats	Staple Vegetables	Beef	Dairy	Food Away from Home
Household total expenditure per AE	-0.028 (0.000)	0.012 (0.000)	0.000 (0.938)	-0.004 (0.000)	0.000 (0.798)	0.004 (0.006)	0.003 (0.000)	0.000 (0.569)
Distance to primary city (log)	-0.045 (0.000)	-0.008 (0.000)	0.137 (0.000)	-0.003 (0.000)	-0.006 (0.000)	-0.004 (0.001)	-0.006 (0.000)	0.000 (0.458)
Distance to secondary city (log)	-0.007 (0.052)	-0.003 (0.001)	0.034 (0.000)	-0.001 (0.105)	-0.002 (0.005)	-0.001 (0.240)	0.002 (0.003)	0.000 (0.668)
Distance to tertiary city (log)	0.001 (0.746)	0.000 (0.934)	0.007 (0.000)	0.000 (0.698)	-0.002 (0.000)	-0.001 (0.134)	0.000 (0.219)	0.000 (0.086)
Nonfarm employment	-0.042 (0.030)	0.018 (0.001)	-0.031 (0.084)	0.018 (0.000)	0.013 (0.011)	0.002 (0.635)	0.001 (0.750)	0.000 (0.286)
Dependency ratio	0.017 (0.307)	-0.012 (0.009)	-0.003 (0.805)	-0.002 (0.567)	-0.002 (0.634)	0.003 (0.455)	0.002 (0.474)	-0.001 (0.063)
Household adult equivalents	-0.003 (0.064)	0.002 (0.001)	0.005 (0.001)	-0.001 (0.000)	0.000 (0.993)	0.000 (0.857)	0.001 (0.005)	0.000 (0.179)
Farmed hectares of land (log)	0.007 (0.034)	0.001 (0.290)	-0.017 (0.000)	0.001 (0.202)	0.000 (0.629)	0.002 (0.001)	0.002 (0.018)	0.000 (0.351)
Own a gas or electric stove	-0.067 (0.001)	-0.007 (0.136)	-0.048 (0.035)	-0.007 (0.347)	0.006 (0.191)	-0.002 (0.539)	-0.004 (0.055)	0.000 (0.041)
Own a refrigerator	0.014 (0.637)	0.007 (0.397)	-0.014 (0.612)	0.012 (0.293)	0.002 (0.613)	-0.001 (0.834)	0.003 (0.540)	0.000 (0.718)
Maximum education attained within the household	-0.003 (0.045)	0.001 (0.105)	-0.003 (0.001)	0.002 (0.000)	0.001 (0.010)	0.000 (0.077)	0.000 (0.049)	0.000 (0.762)
Age of household head	0.000 (0.671)	0.000 (0.028)	0.000 (0.017)	0.000 (0.651)	0.000 (0.330)	0.000 (0.236)	0.000 (0.688)	0.000 (0.907)
Female head of household	0.023 (0.015)	0.000 (0.856)	-0.003 (0.632)	-0.002 (0.325)	-0.002 (0.369)	0.001 (0.751)	0.000 (0.938)	0.000 (0.741)
Own a telephone	0.013 (0.144)	0.009 (0.001)	-0.034 (0.000)	0.011 (0.000)	0.003 (0.089)	0.003 (0.243)	0.003 (0.055)	0.000 (0.072)
Distance to market (log)	-0.002 (0.350)	-0.006 (0.000)	0.007 (0.000)	0.000 (0.985)	-0.004 (0.000)	0.000 (0.417)	0.001 (0.114)	0.000 (0.142)
Own a car	-0.079 (0.000)	-0.001 (0.886)	-0.080 (0.000)	0.017 (0.181)	0.006 (0.608)	0.008 (0.194)	0.005 (0.511)	0.000 (0.966)
Own a motorcycle	0.007 (0.854)	-0.009 (0.215)	-0.079 (0.000)	0.012 (0.115)	0.019 (0.287)	-0.006 (0.044)	0.001 (0.767)	0.000 (0.673)
Own a bicycle	-0.031 (0.000)	0.006 (0.009)	0.009 (0.169)	0.004 (0.007)	0.001 (0.603)	-0.001 (0.626)	-0.004 (0.020)	0.000 (0.162)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

REFERENCES

REFERENCES

- Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *Review of Economics and statistics*, 79(4), 527-539.
- Battersby, J., & Peyton, S. (2014, June). The geography of supermarkets in Cape Town: Supermarket expansion and food access. In *Urban Forum* (Vol. 25, No. 2, pp. 153-164). Springer Netherlands.
- Bennett, M. K. (1941). Wheat in national diets. *Wheat Studies*, (02).
- Berdegúe, J. A., Carriazo, F., Jara, B., Modrego, F., & Soloaga, I. (2015). Cities, Territories, and Inclusive Growth: Unraveling Urban–Rural Linkages in Chile, Colombia, and Mexico. *World Development*.
- Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17), 7301-7306.
- Christiaensen, L., Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: the role of rural diversification and secondary towns¹. *Agricultural Economics*, 44(4-5), 435-447.
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368), 829-836.
- Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596-610.
- D’Haese, M., & Van Huylenbroeck, G. (2005). The rise of supermarkets and changing expenditure patterns of poor rural households case study in the Transkei area, South Africa. *Food Policy*, 30(1), 97-113.
- Fafchamps, M., & Shilpi, F. (2003). The spatial division of labour in Nepal. *The Journal of Development Studies*, 39(6), 23-66.
- Ferré, C., Ferreira, F. H., & Lanjouw, P. (2012). Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, lhs007.
- Garrett, J. L., & Ruel, M. T. (1999). Are determinants of rural and urban food security and nutritional status different? Some insights from Mozambique. *World Development*, 27(11), 1955-1975.

- Gómez, M. I., Barrett, C. B., Raney, T., Pinstруп-Andersen, P., Meerman, J., Croppenstedt, A., Carishma, B., & Thompson, B. (2013). Post-green revolution food systems and the triple burden of malnutrition. *Food Policy*, 42, 129-138.
- Gómez, M. I., & Ricketts, K. D. (2013). Food value chain transformations in developing countries: Selected hypotheses on nutritional implications. *Food Policy*, 42, 139-150.
- Haggblade, S., Hazell, P., & Reardon, T. (2010). The rural non-farm economy: Prospects for growth and poverty reduction. *World Development*, 38(10), 1429-1441.
- Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate behavioral research*, 45(4), 627-660.
- Hino, H. (2014). Shopping at different food retail formats: Understanding cross-shopping behavior through retail format selective use patterns. *European Journal of Marketing*, 48(3/4), 674-698.
- Huang, J., & Bouis, H. (2001). Structural changes in the demand for food in Asia: empirical evidence from Taiwan. *Agricultural Economics*, 26(1), 57-69.
- Leser, C. E. V. (1963). Forms of Engel functions. *Econometrica: Journal of the Econometric Society*, 694-703.
- Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The manchester school*, 22(2), 139-191.
- Nelson, G. C. (2002). Introduction to the special issue on spatial analysis for agricultural economists. *Agricultural Economics*, 27(3), 197-200.
- 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*, volume 11, 619-632.
- Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1), 121-133.
- Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition reviews*, 70(1), 3-21.
- Reardon, T., & Timmer, C. P. (2007). Transformation of markets for agricultural output in developing countries since 1950: How has thinking changed?. *Handbook of agricultural economics*, 3, 2807-2855.
- Renkow, M. (2007). Cities, Towns, and the Rural Nonfarm Economy in S. Haggblade, Hazell, P. B., & Reardon, T. (Eds.). (pp. 183-198). International Food Policy Res Institute.

- Rischke, R., Kimenju, S. C., Klasen, S., & Qaim, M. (2015). Supermarkets and food consumption patterns: The case of small towns in Kenya. *Food Policy*, 52, 9-21.
- Satterthwaite, D., McGranahan, G., & Tacoli, C. (2010). Urbanization and its implications for food and farming. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554), 2809-2820.
- Serber, G. A. F., & Wild, C. J. (1989). *Nonlinear Regression*. John Wiley & Sons, 435.
- Sharma, A. (2016). Urban Proximity and Spatial Pattern of Land Use and Development in Rural India. *The Journal of Development Studies*, 1-19.
- Stage, J., Stage, J., & McGranahan, G. (2010). Is urbanization contributing to higher food prices? *Environment and Urbanization*, 22(1), 199-215.
- Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. *The Economic Journal*, 64(255), 511-527.
- Strydom, J. W. (2011). Retailing in disadvantaged communities: the outshopping phenomenon revisited. *Journal of Contemporary Management*, 8, 150-172.
- Subramanian, S., & Deaton, A. (1996). The demand for food and calories. *Journal of political economy*, 104(1), 133-162.
- Thiele, S., & Weiss, C. (2003). Consumer demand for food diversity: evidence for Germany. *Food Policy*, 28(2), 99-115.
- Turrell, G., & Kavanagh, A. M. (2006). Socio-economic pathways to diet: modelling the association between socio-economic position and food purchasing behaviour. *Public health nutrition*, 9(03), 375-383.
- Von Braun, J., De Haen, H., & Blanken, J. (1991). *Commercialization of agriculture under population pressure: Effects on production, consumption, and nutrition in Rwanda* (Vol. 85). Intl Food Policy Res Inst.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Working, H. (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, 38(221), 43-56.
- World Bank. (2015, October 4). The international poverty line has just been raised to \$1.90 a day, but global poverty is basically unchanged. How is that even possible? [Blog post]. Retrieved from <http://blogs.worldbank.org/developmenttalk/international-poverty-line-has-just-been-raised-190-day-global-poverty-basically-unchanged-how-even>

ESSAY 2: CONSUMPTION OF PROCESSED FOOD IN THE DIETS OF POOR RURAL AND URBAN CONSUMERS IN DEVELOPING EASTERN AND SOUTHERN AFRICA

2.1 Introduction

I explore how the consumption patterns of processed food in developing Eastern and Southern Africa (ESA)²¹ vary across the household income distribution and spatial considerations. Having a clear understanding of processed food consumption patterns has value, because the food people consume – specifically the share of processed food – generates profound implications on the structure of the agrifood system, the level and types of public and private investments, the level and distribution of employment, and nutrition and health.

Household consumption patterns of processed food are expected to vary much as the consumption of staple food varies across the income distribution (Bennett, 1941; Pingali, 2007); an expectation that households would consume more processed food with rising incomes and increased urbanicity (de Haen et al., 2003). Furthermore, household consumption patterns are expected to change among urban middle class households that attain the supernumerary income (Stone 1954) and sufficient market access to indulge their desire to luxury goods, such as processed food.

By contrast, I find that both urban and rural households are rapidly transitioning towards diets with higher shares of processed food, beginning at income levels below the international poverty line²². Given the surprisingly low income level at which I observe diet change, I seek to

²¹ Countries include Botswana, Burundi, Ethiopia, Kenya, Lesotho, Malawi, Mozambique, Namibia, Rwanda, South Sudan, Swaziland, Tanzania, Uganda, Zambia, and Zimbabwe.

²² The World Bank (2015) defines the poverty line \$1.90 daily per adult equivalent expenditure, calculated with purchasing power parity adjusted 2011 constant international dollars.

understand the pattern in diet change – its speed and direction – across the entire income distribution.

I recognize that non-income factors such as rural nonfarm employment (Senauer et al., 1986), education (Turrell & Kavanagh, 2006), preferences from urbanization and proximity to urban areas (Thiele and Weiss, 2003), and even “lifestyle changes” (Huang and Bouis, 2001) can affect the food consumption patterns, and I include these in the analysis. Beyond the urban/rural distinction, city size has been shown to affect other lifestyle patterns of households within low-income countries such as income and poverty – where income is greater in larger cities, and medium and small cities lead to greater quantities of migrating households escaping from poverty (Ferré et al., 2012; Christiaensen et al., 2013; Berdegue et al., 2015). Therefore, I choose to test for the effects of city size on the consumption pattern of processed food.

The paper proceeds as follows. Section two presents the conceptual model. Section three addresses the context, data, definitions, model, and estimation procedures. Section four presents results, and section five concludes.

2.2 Conceptual Framework

Households select their food consumption patterns to maximize their utility from the food that they consume. The heterogeneity of households’ results in each household possessing a unique utility function based on the consumer demand characteristics regarding preferences for taste, energy density, and convenience; on social and individual norms and beliefs regarding informal institutions, beauty, and lifestyle; and on health concerns including foods’ nutritional content and perceived safety. Households maximize utility functions conditional the affordability of these goods subject to their income (m).

$$\text{Max } u(\mathbf{x}), \text{ s.t. } \mathbf{p}\mathbf{x} \leq m \quad (2.1)$$

optimal bundle: $x(\mathbf{p}, m)$

Prices include market prices as well as shadow prices for the production and preparation as households can purchase food from markets or source food from own production.

I begin the conceptual framework with Bennett's Law providing a relationship between household income and food consumption patterns. Stone's theory of supernumerary income adds to this relationship by suggesting that income induced diet change results from purchases by households with increasing levels of discretionary income. I relax the assumption of required supernumerary income as nearly two thirds of the population within ESA²³ lives below the international poverty line²⁴, which otherwise would have consolidated much of the recent income-induced diet change among the non-poor. If I find a change in the direction or speed of diet change that occurs at income levels below those that would have normally associated with increasing discretionary income, then I would want to know the specific empirical configuration of the pattern; specifically, at what income levels does the pattern begin and what is the shape of the pattern? I would want also to understand its drivers of consumption patterns that affect the consumption patterns of the poor. More formally, controlling for income, what non-income factors drive consumption patterns of poor households?

Rischke et al. (2010) conceptualized a framework that incorporates non-income drivers by combining consumer demand, as discussed above, with food environment – highlight supply-side factors such as access, availability, price and market-induced desirability of food. Access refers to the cost to access foods, which is subject to the density of retailers and households'

²³ Authors' calculation using the World Bank's PovcalNet database.

²⁴ The poverty line is defined as \$1.90 daily per adult equivalent expenditure, calculated with purchasing power parity adjusted 2011 constant international dollars (World Bank, 2015)

costs of reaching them. Availability refers to the variety of foods offered by formal and informal markets. Prices vary across products and across retailers (Gómez and Ricketts 2013). Desirability is the market-induced effects of shopping atmosphere and marketing.

Our conceptual framework expands that of Rischke et al. to include the effects of city size (defined by population) on the urban environment that are exogenous drivers of food environment and consumer demand (Figure 2-1). The urban environment of the city in which an urban household is located would directly affect the household consumption patterns. Urban environment affects rural household consumption patterns in the same way as they would affect urban household consumption patterns, although the increasing costs of accessing cities and the food markets therein as the distance from a household to cities increases would attenuate these affects.

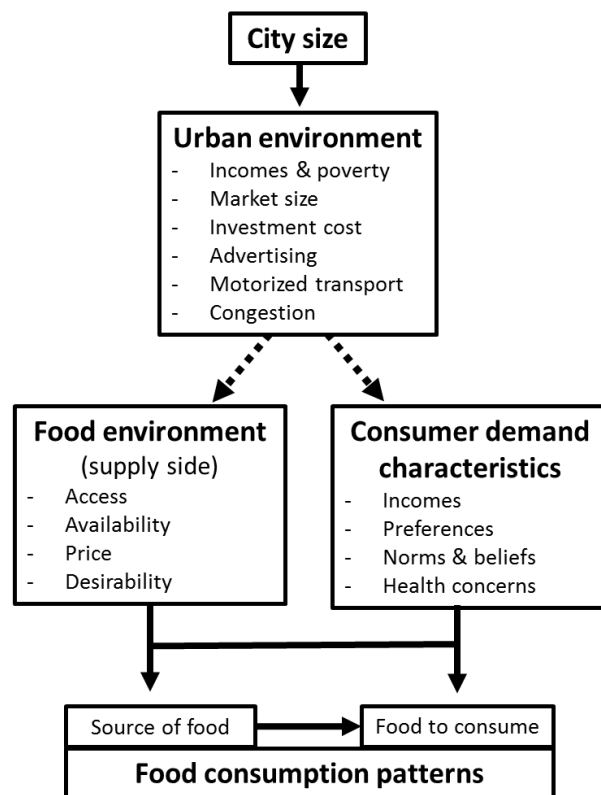


Figure 2-1: Conceptual framework

I suggest six effects of rising city size on the factors of urban environment that are pertinent to food environment and consumer demand.

First, incomes rise and poverty rates fall as shown in recent empirical work across Africa (Ferré et al., 2012, World Bank, 2009; Christiaensen et al., 2013). Both should increase the consumption of processed food.

Second, the market size increases more than proportionally to city size due to the positive relationship between city size and mean consumer incomes. Increased size of market attracts investment in retail food markets; thereby improving physical access to them, which markets expands the variety of foods that are available.

Third, infrastructure investment costs rises, given higher costs of land associated with larger cities. Higher investment cost asymmetrically affects investment: greater impact on formal food retailers that buy land compared to informal retailers that often do not pay for the space they occupy. These costs are transferred to food prices, incentivizing consumers the use of informal retail that may not have the same variety of food as formal retailers. The higher land costs within and near larger cities increase the costs of urban and peri-urban farming, reducing the likelihood of food production for own consumption.

Fourth, increased consumer exposure to modern advertising. Advertising affects consumer preferences, encouraging the consumption of advertised food from modern food retailers.

Third, increases access to public and private motorized transport. Motorized transport improves physical access for outlets such as supermarkets that are larger and fewer in numbers, but should have little or no effect on access to other outlet types, which have developed historically to serve largely foot-bound consumers. Access to transport thus favors the use of

supermarket shopping, which provide consumers with lower prices of processed food (Neven et al., 2006).

Fourth, congestion increases, diminishing the access advantage conveyed by motorized transport and so increasing the cost of shopping in a formal retail market. The increased cost of transportation also increases the costs of cross shopping²⁵, reducing the benefits that multiple markets offer in terms of available variety of food, increasing household benefits of shopping from formal retail market, which offer a variety of food from one location.

Table 2-1 summarizes my expectations of the effects of city size on processed food consumption via urban environment's effect on the food environment and consumer demand factors. Of the six urban environment factors, four drive strictly positive relationships between city size and processed food consumption: mean incomes, market size, exposure to advertising, and access to motorized transport. The cost of infrastructure investment and congestion have mixed effects.

²⁵ The practice of cross shopping is where households shop at various locations. Cross shopping is common in China (Goldman 2000), Israel (Hino 2014) and in South Africa, where households travel to locations away from their area of residence to acquire food (D'Haese and van Huylenbroeck 2005, Strydom 2011).

Table 2-1: Impact of city size on food choice: urban environment, food environment, and demand factors

Change in exogenous urban environment factor as city size rises	Impact on supply (food environment) and demand factors		Impact on processed food demand
	Food environment	Demand factors	
Increased incomes, lower poverty		Higher incomes; rising health concerns; more “modern” norms & beliefs	Positive
Disproportionate increase in market size	Increased access to formal retail; increased availability of processed foods		Positive
Higher investment cost	Reduced physical access to formal markets; increased cost of owning farmland		Positive or Negative
Greater exposure to advertising	Increased desirability of processed food	More “modern” norms & beliefs around food	Positive
Greater access to motorized transport	Asymmetric impact on physical access of markets – increases it for supermarkets but little or no effect for other outlet types		Positive
Higher congestion	Reduces advantage conveyed by motorized transport to purchase all food from formal markets	Higher opportunity cost of time	Positive or Negative

The net effect of city size on processed food consumption depends on the size and effect of each factor. This framework does not provide a singular hypothesis for the effect of city size; therefore, I considered the cumulative nature of the suggested effects and now hypothesize a positive net effect of city size on processed food consumption.

2.3 Methods

2.3.1 Context

Developing Eastern and Southern Africa (ESA) is rapidly urbanizing. United Nations Population Division²⁶ estimates that twenty-one percent of the population in ESA lived in urban settlements in 2000; by 2015, the urbanization level was 26 percent, and it projects forty-four percent of the population will live in urban settlements by 2050. To represent this urbanizing region of ESA I will use the nations of Malawi, Tanzania, Uganda and Zambia. These nations include a variety of population based city “types”, including three primary cities with populations greater than one million, thirty-eight secondary cities with populations between one million and one hundred thousand, and over three hundred tertiary cities with populations below one hundred thousand.

In addition to rapid urbanization, ESA is characterized by rising average per capita income and a modernizing food system. Between 2000 and 2010, the population weighted annual growth in per capita GNI²⁷ of these four nations was positive each year and averaged nearly four percent since 2004. In spite of the rapid growth of incomes, household data shows that daily per adult equivalent incomes remain low, averaging incomes of just above three dollars²⁸ per day (Table 2-2). In the midst of the low average incomes, the “supermarket revolution” is occurring in ESA, evidenced by the expansion of supermarket and other formal food retail outlets (Reardon and Timmer, 2007). Formal retailers typically enter markets by establishing their retail outlets near relatively affluent areas within the greater market (Battersby

²⁶ World Urbanization Prospects: The 2014 Revision

²⁷ World Bank data (<http://data.worldbank.org/>): GNI/capita, PPP (constant 2011 international \$)

²⁸ Daily expenditure levels valued in purchasing power parity (PPP) adjusted constant 2011 international dollars.

and Peyton, 2014), which has resulted in supermarkets directing significant investment into larger cities and lesser investments into secondary and tertiary cities, thereby altering the food retail options across city size.

Table 2-2: National expenditure and population by expenditure strata and city size

	Total Expenditure (per adult equivalent)					Represented Population: Total ('000s) and Cumulative Density by Expenditure Strata								
	Malawi	Tanzania	Uganda	Zambia	Pooled	Total	< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00
National	2.48 (1.67)	3.03 (2.26)	3.51 (2.54)	3.13 (1.74)	3.11 (2.19)	101,923	2.1	13.8	30.2	45.2	67.4	86.2	96.7	100
Rural	1.97 (1.52)	2.38 (1.97)	2.78 (2.23)	1.78 (1.30)	2.37 (1.91)	77,041	2.6	16.9	36.3	53.2	76.6	93.4	99.1	100
Urban	5.31 (3.09)	4.91 (3.50)	6.10 (4.37)	5.66 (3.74)	5.39 (3.78)	24,882	0.4	4.1	11.2	20.4	38.8	64.0	89.2	100
Primary City		7.85 (5.83)	8.11 (6.39)	7.41 (4.62)	7.82 (5.72)	6,126	0.0	1.5	3.4	6.6	17.0	41.7	77.4	100
Secondary City	5.25 (3.14)	4.35 (3.34)	8.67 (4.56)	5.20 (3.68)	5.06 (3.43)	7,178	0.2	3.7	11.3	22.1	42.7	67.4	90.8	100
Tertiary City	5.52 (2.84)	3.52 (2.76)	5.10 (3.77)	4.48 (2.99)	4.31 (3.13)	11,578	0.8	5.6	15.2	26.6	48.0	73.6	94.4	100

Source: authors' calculations using national household level surveys

Notes: Daily average expenditure levels per adult equivalent are presented in purchasing power parity adjusted constant 2011 international dollars. Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

2.3.2 Data

Four nationally representative household surveys provide the household data for this analysis: 2010/2011 Malawi Integrated Household Survey (sample size of 12,271), 2010/2011 Tanzania National Panel Survey (3,924), 2009/2010 Uganda National Panel Survey (2,975), and 2010 Zambia Living Conditions Monitoring Survey (19,397). Due to the nations' geographic proximity, similarities in national demographics, the timing of the surveys, and the similarity of the data collection instruments I selected these datasets. In-person interviews performed by enumerators of these surveys collected household data on consumption, employment, and other household characteristics (Table 2-3).

Table 2-3: Household descriptive characteristics

	Pooled Data					
	National	Rural	Urban	Primary Cities	Secondary Cities	Tertiary Cities
Total household expenditure per AE	3.11 (2.19)	2.37 (1.91)	5.39 (3.78)	7.82 (5.72)	5.06 (3.43)	4.31 (3.13)
Nonfarm employment	27.1 (20.0)	21.4 (0.0)	44.7 (50.0)	52.6 (50.0)	41.9 (40.0)	42.1 (40.0)
Dependency ratio	47.1 (50.0)	49.8 (50.0)	38.9 (40.0)	33.4 (33.3)	38.7 (40.0)	41.9 (42.9)
Household adult equivalents	6.7 (5.6)	6.8 (5.6)	6.4 (5.4)	6.5 (5.5)	6.2 (5.2)	6.4 (5.4)
Hectares of farmed land	3.5 (0.9)	4.1 (1.2)	1.4 (0.0)	0.5 (0.0)	1.2 (0.0)	2.0 (0.2)
Own a gas or electric stove	6.9 (0.0)	1.9 (0.0)	21.5 (0.0)	33.2 (0.0)	23.8 (0.0)	12.0 (0.0)
Own a refrigerator	8.7 (0.0)	2.3 (0.0)	27.7 (0.0)	48.8 (0.0)	29.3 (0.0)	12.9 (0.0)
Maximum education attained within the household	13.5 (14.0)	12.3 (14.0)	17.1 (17.0)	20.7 (17.0)	12.9 (13.0)	17.8 (17.0)
Age of household head	45.6 (43.0)	46.2 (44.0)	43.7 (41.0)	43.9 (41.0)	42.8 (41.0)	44.2 (41.0)
Female headed household	22.1 (0.0)	21.2 (0.0)	24.9 (0.0)	26.1 (0.0)	21.2 (0.0)	26.6 (0.0)
Own a phone	55.4 (100.0)	46.7 (0.0)	82.4 (100.0)	93.3 (100.0)	84.3 (100.0)	75.5 (100.0)
Kilometers to Market	9.5 (4.5)	10.7 (5.7)	5.7 (2.0)	10.3 (2.0)	4.7 (2.0)	3.9 (1.4)
Own a car	2.9 (0.0)	1.2 (0.0)	8.2 (0.0)	12.1 (0.0)	10.7 (0.0)	4.6 (0.0)
Own a motorcycle	4.7 (0.0)	4.6 (0.0)	5.2 (0.0)	3.9 (0.0)	4.3 (0.0)	6.4 (0.0)
Own a bicycle	47.0 (0.0)	51.9 (100.0)	31.9 (0.0)	13.7 (0.0)	35.4 (0.0)	39.3 (0.0)

Source: authors' calculations using national household level surveys

Notes: Daily average expenditure levels per adult equivalent are presented in purchasing power parity adjusted constant 2011 international dollars. Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

The expenditure data are valued in purchasing power parity (PPP) constant 2011 international dollars, using historic exchange rates obtained from XE.com and conversion factors for constant PPP valuation from worldbank.org. I use per adult equivalent (AE)²⁹ levels of expenditure to calculate total household expenditure, total food consumption, and totals for various food consumption aggregates. Each survey collected data on the value of food consumption with a seven day recall, with the exception of Zambia, which used a two or four week recall depending on the food item. The valuation of consumed own production varies by dataset. Stated values of food consumed from own production are collected by the Uganda and Zambia surveys³⁰, while the Malawi and Tanzania datasets include quantities of consumed own production that I value using imputed household and community prices of marketed food. To estimate these values, I calculated prices of purchased food items that match items from consumed own production. The medians of these purchased food prices are then calculated based on locality and multiplied by the quantities of consumed own production to estimate the value of consumption.

City population data from <http://www.citypopulation.de/>, a website that assembles recent census data, paired with household level data identify the population size of the cities where urban households reside. Distance measurements for rural Zambia are approximated using distance from center of rural constituency to the nearest primary, secondary and tertiary cities with the distance calculator at <https://www.daftlogic.com/>.

²⁹ Per adult equivalent values are calculated as one AE for each household member fifteen years of age or older, 0.75 AE for each child aged five to fourteen and 0.50 AE for each child younger than five years old.

³⁰ The observations where Uganda reported farm-gate prices, I imputed market purchase prices to represent the value of own production.

2.3.3 Definitions

Processed share is the term I use to represent the share of processed food value in total food expenditure. Processed food is defined as food that prior to household acquisition has been transformed from its original state, beyond removal from the plant and (for non-perishables) drying. Processed food therefore identifies food embedded with the labor of humans or machines.

To analyze the consumption of processed food with more granularity I create five processing sub-aggregates through the following three steps. First, I categorize food items into maize and non-maize food aggregates; as maize is the primary staple food in the region and could dominate the consumption patterns. Second, I disaggregate both the maize and non-maize food aggregates into processed and unprocessed aggregates as defined above. Third, I further disaggregate the processed non-maize aggregate into processed low and processed high food aggregates. I assign to the high category to food items if they satisfy at least two of the following three conditions: multiple ingredients; physical change induced by heating, freezing, extrusion, or chemical processes (i.e., more than simple physical transformation); and packaging more complex than simple paper or plastic. The resulting sub-aggregates are unprocessed maize, processed maize, unprocessed non-maize, low processed non-maize, and high processed non-maize (Table 2-4). Although this aggregation delineates between unprocessed, processed low and processed high based on the structure of the food system and the value-added that they represent, this aggregation closely mirror the processed food aggregates considered in nutrition-based literature³¹ (Monteiro et al., 2010).

³¹ Table 2-A-1 shows a sample of the food items and how the items map into both the Monteiro et al. (2010) aggregates and the processing aggregates.

Table 2-4: Processed food aggregates

Share	Definition
Processed consumption	Share of any food acquired in processed form
Unprocessed Maize	Maize acquired in unprocessed form, e.g. maize grain
Processed Maize	Maize acquired in processed form, e.g. maize meal
Unprocessed	Non-maize food items acquired as unprocessed, e.g. common bean
Low Processed	Non-maize food items acquired as low processed, e.g. wheat flour
High Processed	Non-maize food items acquired as high processed, e.g. bread

City size is the term used to classify cities by population. For analysis, I categorize cities by population thresholds: primary cities have populations greater than one million, secondary cities have populations below one million and above than 100,000, and tertiary cities have populations below one hundred thousand.

2.3.4 General model

I follow a four-part analysis to estimate the effects of city size and distance to cities on food consumption patterns. I establish the empirical configuration of the pattern of diet change across the income distribution in the first two parts, and explore drivers of the consumption patterns, that include the effects of city size, in the last two parts.

Part 1 is nonparametric analysis in the form of LOWESS curves, which impose no structure on the data, allowing the data to reveal patterns of consumption across the expenditure distribution. This analysis provides a first assessment of identifying at what expenditure level changes in consumption patterns begin, and whether changes are sudden and rapid or gradual.

Part 2 formally establishes expenditure breakpoints, and rates of change before and after them, by testing for thresholds where significant change occurs in (a) the rate of change or (b) the direction of the relationship between food consumption patterns and total expenditure. By imposing linear relations on the food consumption patterns before and after an estimated expenditure breakpoint, part 2 formally estimates particular expenditure levels at which patterns

change, and estimates patterns of change at expenditures below the breakpoint (among lower income households) and above it (among higher income households).

Part 3 examines drivers of consumption change with the use of Engel curve based regressions to estimate the marginal effects of income and non-income determinants on food consumption shares. Parts 1 & 2 direct the selection of an appropriate functional form of the Engel curve model that I use in the regression analysis.

Part 4 uses a mediation model to estimate the indirect effects of city size and distance to city on food consumption patterns occurring via other determinant variables. The total effects of city size and distance to cities incorporate the estimated effects from the regression analysis in part 3 with the indirect effects from part 4.

I complete each part of the analysis separately for rural and urban households and by country, with descriptive and non-parametric analysis highlighting city size differences that I further analyze in the parametric analysis.

2.3.5 Empirical model

The nonparametric analysis and the expenditure threshold analysis will estimate the marginal effect of income (Y^*), without controlling for any other determinants of consumption, on processed share and five sub-aggregates that depict processed food consumption.

$$D_{Food\ Aggregate} = f(Y^*) \quad (2.2)$$

In parts 3 & 4 of the analysis, I use an Engel curve model to estimate the marginal effects of income and non-income household variables, including city size and distance to city, on each of the dependent variables shown in Table 2-4. The conceptual model posits that city size influences the urban environment, which affects both the food environment and consumer

demand characteristics. To specify the model empirically, I include variables that capture as many of the supply and demand side factors as possible. In the urban analysis, I include dummy variables for primary and secondary cities that reflect city size differences, and in the rural analysis, I include variables for the distances to various city sizes. I use the conceptual model to interpret the meaning of the city size variables and generate defensible conclusions in instances where data to represent factors of city size are unavailable or imperfectly capture the concepts in the conceptual model.

Data from the national datasets sufficiently capture consumer demand characteristics, with variables that proxy for incomes, preferences, norms and beliefs, and health concerns. The national datasets provide less comprehensive coverage of the food environment. Four variables from the datasets are relevant to accessing food: own a car, own a motorcycle, own a bicycle, and distance to nearest market (for rural households), but measures of availability, price and food desirability are not available. The national datasets do not capture household level prices well due to the high prevalence of non-standard units, limiting the ability to apply prices to food aggregates. I capture systematic differences across cities with the inclusion of city dummy variables in the pooled regressions.

I am able to directly control for two of the six urban environment factors within the conceptual model: income and access to motorized transport. City size dummies capture the four remaining factors of city size: market size, congestion, investment cost, and advertising. I estimate the marginal effects of income, city size, distance to cities and the other determinant variables with the following Engel's Curve Model:

$$D_{Dependent\ Variable} = f(Ac^*, Av^*, P^*, D^*, Y^*, Z^*, W^*, H^*) \quad (2.3)$$

The dependent variables are processed share and the processing sub-aggregates (Table 2-4). The determinants identified in equation three are the factors of food environment and consumer demand characteristics, each with household variables that proxy for their effects as shown in Table 2-5 and described below.

Table 2-5: Determinants of demand in the Engel's Curve Model

Ac*	Access	Own a car, Own a motorcycle, Own a bicycle, Distance to market (rural), City size (urban), Distance to cities (rural)	Y*	Disposable Income	Daily total expenditure per adult equivalent, City size (urban), Distance to cities (rural)
Av*	Availability	City size (urban), Distance to cities (rural)	Z*	Individual and Household Preferences	Nonfarm employment, Dependency ratio, Adult equivalents, Farmed hectares of land, Own a stove, Own a refrigerator, Maximum level of education, Age of household head, Female head of household, City size (urban), Distance to cities (rural)
P*	Price	City size (urban), Distance to cities (rural)	W*	Social and Individual Norms and Beliefs	Own a stove, Own a refrigerator, Maximum level of education, Age of household head, Female head of household, Own a telephone, City size (urban), Distance to cities (rural)
D*	Desirability	City size (urban), Distance to cities (rural)	H*	Health Concerns	Maximum level of education, Female head of household, Dependency ratio, City size (urban), Distance to cities (rural)

City size variables that affect the urban environment in the urban analyses are *dummy variables for residence in primary- and secondary cities (tertiary cities are excluded)*, and in the rural analyses are *distance to nearest primary city, distance to nearest secondary city, distance to nearest tertiary city*. In the urban analyses, these variables simply identify the general city size where the household resides, in turn identifying the urban environment. In the rural analyses, these variables account for accessing the urban environment and are included in logarithmic form to account for diminishing marginal effects of distance. As distances increase, there is a reduced effect of urban environment on food environment, consumer demand characteristics and ultimately food consumption decisions. The attenuation of these effects are a result of increased costs of accessing urban areas and the diminished spillover effects of the urban environment at

greater distances. I expect that households near larger cities will have consumption patterns similar to those of urban residents due to the density of commercial activity within larger cities that is likely to lead to greater commercial networks immediately surrounding the cities.

I use the dummy variables *ownership of a car*, *ownership of a motorcycle*, and *ownership of a bicycle* to proxy for access to foods. Ownership of any of these transportation vehicles reduce the cost of access, while increasing the ability to shop in bulk with decreasing magnitude in the following rank: car, motorcycle, and bicycle. *Distance to market* also acts as a proxy for access to foods as greater distances increase the shadow price of transportation to markets, reducing household ability to purchase processed food. Distance is measured in kilometers and it enters the estimation in a logarithmic form to account for diminishing effects of distance. Variables of household distance from household to market are available for Tanzania and Zambia, and distance from community to market are available for Malawi and Uganda. Distance to market variables are only included in rural household analysis, as certain distances are only available at the community level.

Total daily household expenditure per AE proxies for income, a key factor of consumer demand. The estimation methods section explains how expenditure enters the estimation equations.

I include nine variables, in addition to city size and distance to cities variables, to proxy for individual and household preferences. I calculate *nonfarm employment* as the percentage of working age³² adults employed (including self-employment) in nonfarm activities. I calculate the household's *dependency ratio* by dividing the number of dependents by the total members within the household. Both nonfarm employment and dependency ratio increase household opportunity

³² Ages 15-64

cost of time, thereby incentivizing households to reduce their time allocated towards preparing food, which incentivizes the consumption of processed foods. I calculate *household adult equivalents* as the total number of adult equivalents within a household because more household equivalents reduce the opportunity cost of time based on economies of scale within households, therefore reducing the benefit of processed food consumption.

Household shadow prices regarding household ability to produce, prepare and/or preserve food vary based on household characteristics such as the amount of *land farmed*, *ownership of a gas/electric stove* and *ownership of a refrigerator*. *Land farmed* is the variable representing the hectares of land that a household farmed during the previous harvest season, which enters the analysis in logarithmic form to account for diminishing marginal effects of the amount of land farmed on food consumption patterns. Farming one's own land increases the likelihood of consuming food from own production, which would directly reduce the share of processed food consumption that by definition households acquired in the processed form. Ownership of a *gas/electric stove* will reduce the preparation cost of food, diminishing the benefits of processed food. Ownership of a *refrigerator* reduces the cost of food preservation, reducing the benefit of low perishability that is commonly associated with processed food.

The *maximum level of education* within a household is an ordinal value based on the level of education acquired by any member of the household. Education is positively associated with healthier consumption patterns, which could affect the consumption of processed foods (Turrell and Kavanagh, 2006). Elderly households are more likely to maintain traditional diets, with low levels of processed food; therefore, I include *age of the head of the household* in the empirical model. Finally, as females have a greater propensity to delay gratification and consume healthier

food options (Smith, 2003), I include the *gender of the head of the household* in the empirical model.

Ownership of a telephone is a proxy for social and individual norms in addition to ownership of a gas/electric stove, ownership of a refrigerator, maximum level of education, age of household head, and female head of household, which I already described. Ownership of a phone is a zero one dummy variable that indicates household tendency towards modernization and willingness to modernize food consumption patterns.

The variables that proxy for health concerns are maximum level of education, female head of household, and dependency ratio, which I have previously described.

2.3.6 Estimation methods

I use LOWESS curve regressions (Cleveland, 1979; Cleveland and Devlin, 1988) to observe the nonparametric relationship between the food budget share and the household per adult equivalent total expenditure. My second application of LOWESS curves is to map the relationship between distance to various city types and rural food consumption patterns. The LOWESS procedure is a robust form of local polynomial smoothing that does not impose functional form on the relationship between dependent and independent variables. Compared with other forms of local polynomial smoothing, this specific form of local polynomial smoothing is more resistant to statistical outliers because it applies smaller local weights to estimates with large residuals, which is valuable given the nature of household survey data. I completed the LOWESS analysis on expanded data where I generated duplicate observations according to population weights to represent the sample population.

I test for expenditure thresholds in the second part of the analysis, to observe changes in the direction or rate of change in the food consumption patterns. I use a non-linear estimation to find the best fit of two linear curves that join at a singular point that optimizes the fit of the two curves. This estimation method is as a piecewise regression (Serber and Wild, 1989) that takes the following form:

$$E[y|x] = \begin{cases} \beta_{10} + \beta_{11}x, & x \leq \alpha \\ \beta_{20} + \beta_{21}x, & x > \alpha \end{cases} \quad (2.4)$$

$$s. t. \quad \beta_{10} + \beta_{11}\alpha = \beta_{20} + \beta_{21}\alpha$$

In the piecewise regressions, the household food budget shares of food item aggregates (y) are regressed against household total expenditure per adult equivalent (x), generating inflection point values and slopes of each curve around the inflection points with robust standard errors. This estimation method shows statistically significant changes in the consumption patterns when the piecewise regression identifies a kink in the consumption curve at the inflection point.

The third part of the analysis incorporates the nonparametric analysis and the threshold analysis by contributing to the identification of the functional form of the Engel curve. Four common forms of Engel's curves are: the linear relationship suggested in Engel's original work (1857), the linear relationship between log income and the dependent variable as identified by Working (1943), Working's model with the addition of the inverse of income as proposed by Leser in 1963, and the functional form proposed by Banks, Blundell and Lewbel (1996) that includes both log income and squared log income. Each of these forms follow the base form of a share of consumption that is dependent on income and other household variables. The inclusion of non-linear relationships between income and the consumption of goods has been an improvement to the linear functional form as the non-linear relationships enable the analysis of

consumption patterns that exhibit differing marginal, potentially inverted, effects of income on the relative shares of consumption of goods over strata of household income.

I follow a two-step process to select the appropriate functional form of the Engel curve. First, I use the results of the nonparametric and semi-parametric analyses to narrow the options of the functional form based on the general shapes of the patterns between total expenditure and processed share. Second, I complete regressions using the reduced list of functional forms and compare the fit of the regressions with adjusted R^2 or chi-squared values to identify the best fit. I treat total expenditure as exogenous in food demand analysis as recommended by Subramanian and Deaton (1996).

Consumption shares bound the dependent variables in the Engel's Curve Model between zero and one. The standard estimation method for such a dependent variable is to convert it with a log-odds transformation and then to estimate with ordinary least squares (Wooldridge, 2010). Food consumption data from in one-week survey results in many households with food budget shares of zero or one hundred percent, disallowing the use of the log-odds transformation. Given the characteristics of the dependent variable and the assumption of asymptotically normal standard errors, the appropriate method of estimation is the fractional probit model (Papke and Wooldridge, 1996 and 2008). This model estimates the marginal effects of the predictors of a dependent variable that takes values of a closed set of zero to one.

In the fourth part of the analysis, I use a mediation model to test for indirect effects of the variables *city size* and *distance to city* on food consumption patterns. City size and distance to city are predictor variables that could affect other determinant variables (mediator variables), which also affect food consumption patterns. I would not capture the indirect effects of city size and distance to city variables that occur through the mediator variables without the use of a

mediation model. I calculate the indirect effects of predictor variables by multiplying the estimated effect of a mediator variable on a dependent variable with the estimated effect of the predictor variable on the mediator variable. I use OLS to estimate the city size and distance to city effects on the mediator variables, and bootstrap the standard errors of the indirect effects (Wooldridge, 2010). The estimation of indirect effects follows the methodology described in Hayes and Preacher (2010).

2.4 Results

I use descriptive statistics, LOWESS curves analysis and piecewise regressions to show the relationship of processed food consumption with income and then across city size in urban areas and distance to cities in rural areas. Then I will discuss the results of the full Engel's Curve Model regression and the results of the mediation model analysis.

2.4.1 Descriptive statistics and nonparametric analysis of the relationship between processed food consumption and total household expenditure

Tables 2-6 & 2-7 show the consumption pattern of processed food across the expenditure strata of urban and rural households, while the LOWESS curves (Figure 2-2) visually represent these same relationships. These descriptive statistics and estimated curves do not show causation, but are valuable in observing preliminary patterns between household total expenditure and processed share.

Table 2-6: Food budget shares of processed food, urban households by country and expenditure

Urban	Total	Average Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	62.5 (66.3)	54.5 (60.0)	56.4 (58.5)	51.4 (55.6)	56.2 (57.0)	62.0 (65.4)	65.7 (68.2)	68.9 (71.4)	66.4 (71.4)	50.0 (52.7)	57.8 (59.6)	65.7 (68.7)
Tanzania	66.2 (69.8)	44.4 (56.7)	63.1 (66.8)	56.2 (59.3)	54.9 (59.2)	59.3 (65.5)	67.9 (70.3)	74.1 (76.2)	78.8 (80.3)	56.8 (60.4)	58.7 (63.4)	72.2 (75.0)
Uganda	54.7 (56.2)	73.3 (78.6)	44.7 (52.5)	37.9 (42.1)	46.4 (47.1)	42.5 (38.9)	52.8 (55.6)	60.9 (61.0)	69.1 (67.7)	44.8 (47.1)	44.9 (43.8)	59.9 (60.5)
Zambia	65.4 (68.4)	46.0 (46.7)	50.3 (53.5)	56.4 (58.2)	59.3 (62.2)	63.7 (66.7)	67.3 (68.8)	69.8 (71.1)	72.5 (73.8)	51.7 (54.3)	60.4 (63.5)	69.0 (70.8)
Pooled	62.7 (66.8)	52.9 (56.7)	56.8 (60.8)	52.9 (57.0)	54.2 (57.6)	56.7 (61.3)	63.3 (66.7)	68.7 (70.7)	73.6 (75.0)	52.8 (56.0)	55.7 (60.1)	67.6 (69.7)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

Table 2-7: Food budget shares of processed food, rural households by country and expenditure

Rural	Total	Average Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	37.0 (35.1)	38.5 (32.9)	34.7 (29.0)	34.0 (30.3)	35.3 (32.9)	38.3 (37.4)	43.0 (42.9)	52.0 (52.3)	56.3 (69.1)	34.7 (29.8)	35.2 (32.3)	42.6 (42.3)
Tanzania	39.0 (36.5)	50.7 (51.5)	36.1 (31.1)	34.6 (30.4)	35.0 (31.8)	38.1 (36.1)	43.0 (41.5)	54.4 (53.7)	65.7 (65.5)	36.0 (32.0)	37.4 (35.1)	46.3 (46.9)
Uganda	30.7 (26.6)	25.9 (9.9)	26.7 (26.8)	26.5 (20.2)	27.4 (22.2)	28.5 (25.9)	33.9 (29.0)	43.6 (45.9)	64.7 (71.4)	26.4 (22.0)	28.6 (24.4)	40.6 (38.9)
Zambia	33.2 (28.9)	29.1 (24.2)	27.3 (23.1)	30.8 (26.5)	34.3 (31.1)	37.6 (34.5)	43.6 (43.9)	56.1 (59.0)	59.1 (67.3)	28.3 (24.1)	34.3 (30.8)	45.2 (45.7)
Pooled	35.6 (32.3)	41.2 (36.6)	32.1 (26.8)	31.9 (27.4)	32.8 (28.5)	34.7 (32.8)	39.6 (37.1)	49.9 (51.2)	63.5 (67.7)	32.0 (26.8)	34.1 (31.2)	43.9 (43.7)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

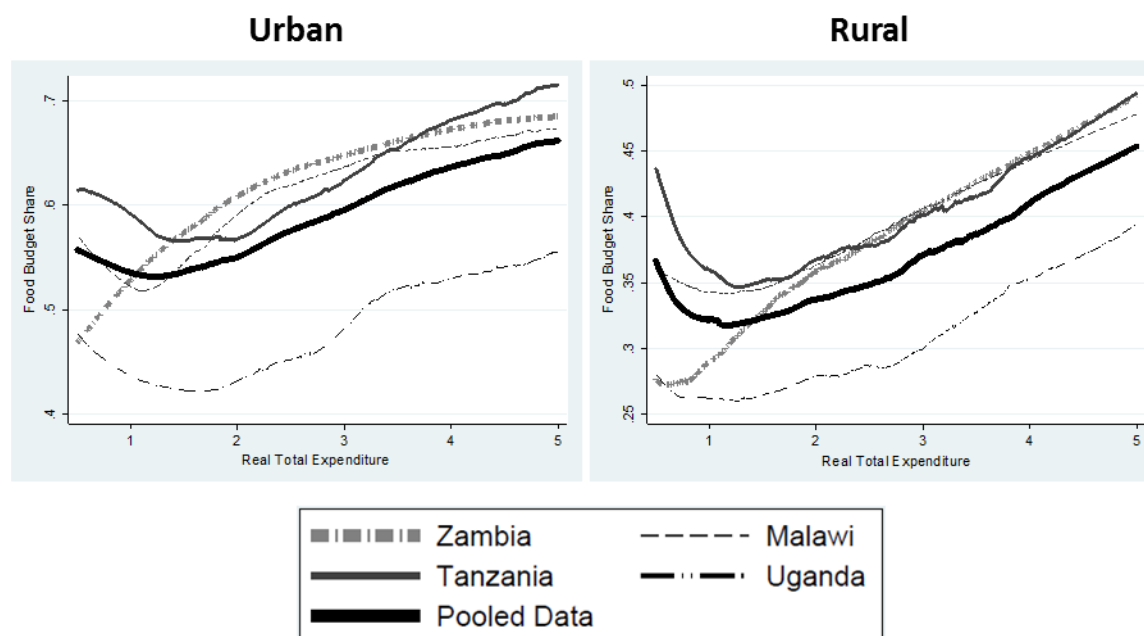


Figure 2-2: Nonparametric (LOWESS curve) analysis of consumption patterns of processed food
 Note: LOWESS curve estimates use a bandwidth equal to 0.4.

Figure 2-2 shows a general pattern in both urban and rural households of initial decline in the share of processed food followed by a monotonic rise of processed share as total expenditures rise. The only exception to this pattern is in urban Zambia, where there is no initial decline. The overall positive correlation between processed food consumption and expenditure is consistent with the patterns observed in low-income countries in Asia, where households consume greater shares of processed food at higher levels of household expenditure (Pingali 2007). The pattern shown here resembles the functional forms of Engel curves for luxury goods as suggested by Leser (1963) or Banks, Blundell and Lewbell (1997). Beyond the general pattern, there are three results of note.

First, the consumption of processed food is greater for urban households (62.7%) than for rural households (35.6%). The larger average shares of purchased food³³ by urban households

³³ Author's calculation from pooled data

(83.4%) than by rural households (45.6%) are likely to drive this relationship, as processed food by definition cannot be acquired by own production.

Second, the positive pattern of processed share begins below the poverty line in all countries for both urban and rural households. The LOWESS estimates of the pooled data show that lowest estimated share of processed food in urban households (53.1%) occurs at total expenditure of \$1.22 and the lowest share of processed food in rural households (31.8%) occurs at total expenditure of \$1.16. These estimates, which are consistent with the average processed shares by expenditure strata, show two findings: One, the demand for processed food occurs at all levels of income; it is not merely a middle-class story. Two, processed food are a significant portion of households' diets at any income level, as evidenced by the lowest share of processed food consumption being more than half of the food value consumed by urban households, and nearly a third of the food value consumed by rural households.

Third, the poorest households do not consume the lowest share of processed food value across the expenditure strata. I suggest two explanations of this. One, the poorest households have the least amount of land that could be used to produce food for own consumption. This lack of capital results in poor households not being able to produce sufficient food for own consumption necessitating their purchase of food, including processed food. Two, the poorest households are likely to receive the greatest share of their income in the form of remittances, which leads to these households purchasing a significant share of their food value. The data do not consistently show the share of income from remittances, but the data show that the share of food value received in the forms of food in kind or gifts³⁴, which I assume is correlated with remittance income, is greatest among the poorest households.

³⁴ Ten percent of food value among households with total daily expenditure below \$1.00, six percent of food value among households with total daily expenditure above \$1.00.

Figure 2-3 and Table 2-A-2 of the appendix show additional depth regarding patterns of processed food consumption in the form of processing level aggregates. Highly processed food, one of the processing level aggregates, has garnered much attention in nutrition literature focused on low-income countries (Popkin et al., 2012; Gómez et al., 2013; Monteiro et al., 2013). Highly processed food contains many utility enhancing attributes such as convenience and taste in the form of increased salt and fat. My findings show that highly processed foods exhibit the largest increase in food budget share among both rural (9% - 21%) and urban (17% - 31%) households across the expenditure stratum³⁵, with the transition towards greater shares of highly processed food beginning at total expenditures of under one dollar per day. Unprocessed maize shows the greatest decline in food budget share, where three of the four countries consume their maximum shares of unprocessed maize at expenditures levels below \$1.00. These patterns reinforce and expound upon the patterns identified in nutrition literature.

³⁵ These results are for pooled households with daily total expenditure below \$5.00.

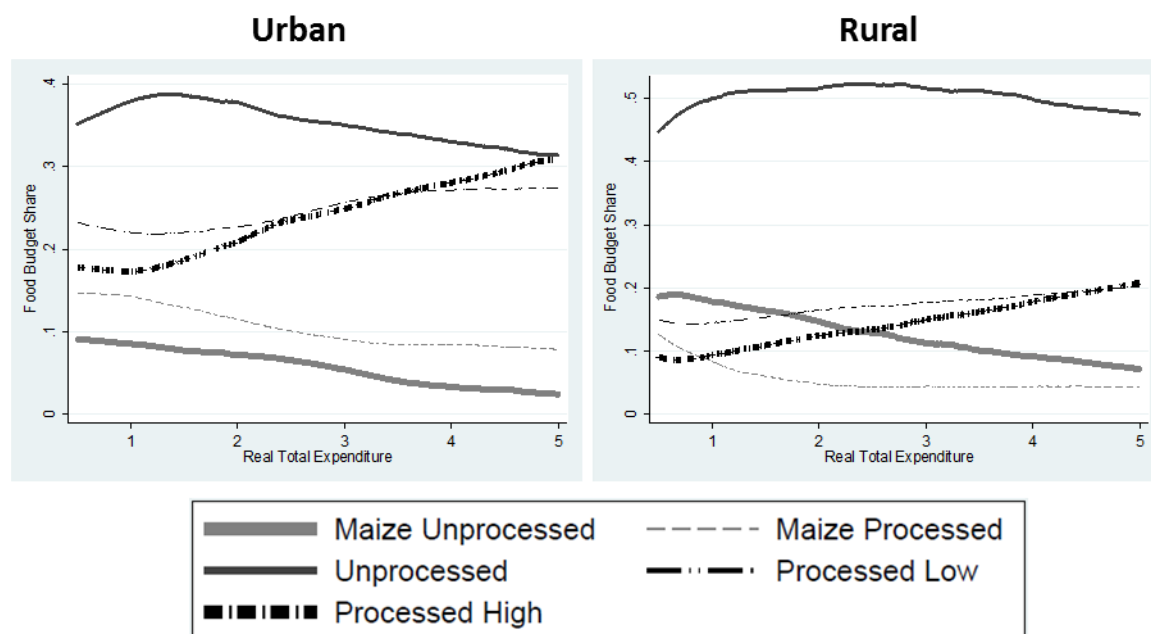


Figure 2-3: Nonparametric (LOWESS curve) analysis of processed food aggregates – pooled data

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

2.4.2 Expenditure threshold analysis

The expenditure threshold analysis shows that patterns of rural and urban processed share vary across income levels, and that the rise of processed share commonly begins among poor households (Table 2-8). Expenditure breakpoints for rural Malawi and Tanzania are \$1.10 and \$0.87 respectively, breakpoints that reveal negative slopes in the consumption patterns below the breakpoint and positive beyond the breakpoint. Positive relationships between expenditure and processed share bracket the expenditure breakpoints of rural Uganda and Zambia. These results suggest that supernumerary income is not a necessary condition for households to increasing processed food consumption.

Table 2-8: Expenditure thresholds and slopes prior to and after expenditure thresholds in processed food consumption patterns

Countries	Rural			Urban		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.10 (0.000)	-0.055 (0.070)	0.038 (0.000)	\$3.55 (0.000)	0.058 (0.000)	-0.019 (0.544)
Tanzania	\$0.87 (0.000)	-0.325 (0.004)	0.038 (0.000)	\$1.53 (0.000)	-0.089 (0.309)	0.057 (0.000)
Uganda	\$3.29 (0.000)	0.014 (0.159)	0.095 (0.000)	\$2.12 (0.041)	-0.025 (0.714)	0.053 (0.062)
Zambia	\$2.14 (0.003)	0.061 (0.000)	0.041 (0.000)	\$2.08 (0.000)	0.091 (0.000)	0.024 (0.000)
Pooled	\$0.86 (0.000)	-0.200 (0.000)	0.034 (0.000)	\$1.40 (0.000)	-0.053 (0.438)	0.042 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

The urban expenditure threshold analysis is less statistically significant than the rural analysis, but the statistically significant results indicate a positive relationship between household expenditure and processed food consumption, regardless of the expenditure level. Urban Zambia provides the only instance where the slopes of this relationship are significant above and below the breakpoint, and it is of note that the slope prior to the expenditure breakpoint is greater than beyond the expenditure breakpoint.

Tables 2-9 & 2-10 display the results of the expenditure breakpoint analysis for the processing sub-aggregates, with five findings noted here. First, each³⁶ of 13 statistically significant pairs of slopes representing rural or urban country level processing aggregate consumption patterns have the same direction and are of greater absolute value before the expenditure breakpoints than beyond the breakpoints: negative slopes for maize and unprocessed non-maize, positive for non-maize processed low and processed high foods. This pattern of the relative magnitude of slopes highlights the growth of processed food consumption among the poor.

³⁶ This excludes the slopes that bracket the expenditure breakpoint for highly processed non-maize in Uganda, which has a statistically significant expenditure breakpoint at \$0.45. Half of one percent of the observations in Uganda have expenditures less than \$0.45.

Table 2-9: Expenditure thresholds and slopes prior to and after expenditure thresholds for processed food aggregates – urban data

Countries	Maize Unprocessed			Maize Processed		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$2.50 (0.000)	-0.048 (0.007)	-0.001 (0.745)	\$0.97 (0.000)	-0.492 (0.000)	-0.018 (0.001)
Tanzania	\$2.62 (0.000)	0.012 (0.615)	-0.038 (0.001)	\$2.68 (0.004)	-0.026 (0.051)	-0.007 (0.412)
Uganda		-0.011 (0.423)	-0.006 (0.285)	\$1.95 (0.001)	-0.054 (0.202)	-0.004 (0.682)
Zambia	\$2.00 (0.000)	-0.051 (0.000)	-0.009 (0.000)	\$0.84 (0.001)	0.118 (0.375)	-0.020 (0.000)
Pooled	\$2.59 (0.164)	-0.010 (0.437)	-0.020 (0.008)	\$2.69 (0.000)	-0.034 (0.000)	-0.006 (0.219)

Countries	Unprocessed			Processed Low			Processed High		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$1.19 (0.000)	0.047 (0.401)	-0.023 (0.000)	\$2.66 (0.000)	0.052 (0.000)	0.004 (0.628)	\$2.57 (0.000)	0.063 (0.000)	0.024 (0.006)
Tanzania	\$1.47 (0.001)	0.061 (0.362)	-0.035 (0.000)	\$1.42 (0.000)	-0.094 (0.191)	0.021 (0.001)	\$1.47 (0.064)	0.001 (0.988)	0.051 (0.000)
Uganda	\$2.28 (0.010)	0.033 (0.613)	-0.053 (0.049)	\$0.90 (0.067)	-0.129 (0.619)	0.022 (0.042)	\$0.44 (0.000)	-3.256 (0.000)	0.032 (0.005)
Zambia	\$2.09 (0.000)	-0.041 (0.000)	-0.015 (0.000)	\$1.88 (0.000)	0.070 (0.000)	0.020 (0.000)	\$2.81 (0.000)	0.039 (0.000)	0.019 (0.001)
Pooled	\$1.46 (0.000)	0.053 (0.194)	-0.026 (0.000)	\$1.24 (0.001)	-0.038 (0.488)	0.021 (0.000)		0.037 (0.000)	0.036 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 2-10: Expenditure thresholds and slopes prior to and after expenditure thresholds for processed food aggregates – rural data

Countries	Maize Unprocessed			Maize Processed		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi		-0.045 (0.000)	-0.040 (0.000)	\$1.48 (0.000)	-0.140 (0.000)	-0.009 (0.000)
Tanzania	\$1.93 (0.000)	0.022 (0.178)	-0.046 (0.000)	\$1.79 (0.000)	-0.062 (0.000)	0.000 (0.936)
Uganda	\$3.20 (0.000)	0.002 (0.784)	-0.024 (0.036)	\$2.44 (0.000)	-0.017 (0.032)	0.010 (0.241)
Zambia	\$2.29 (0.043)	-0.036 (0.000)	-0.024 (0.002)	\$0.55 (0.000)	-0.168 (0.090)	0.005 (0.009)
Pooled	\$1.56 (0.173)	-0.024 (0.071)	-0.033 (0.000)	\$1.81 (0.000)	-0.055 (0.000)	0.001 (0.663)

Countries	Unprocessed			Processed Low			Processed High		
	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint	Expenditure Breakpoint	Slope Before Breakpoint	Slope After Breakpoint
Malawi	\$2.00 (0.000)	0.035 (0.000)	-0.006 (0.269)	\$2.18 (0.000)	0.050 (0.000)	0.011 (0.016)	\$1.90 (0.000)	0.051 (0.000)	0.029 (0.000)
Tanzania	\$0.77 (0.000)	0.232 (0.083)	-0.009 (0.170)			0.008 (0.015)	\$2.00 (0.000)	0.016 (0.097)	0.045 (0.000)
Uganda	\$2.56 (0.000)	-0.007 (0.616)	-0.051 (0.000)	\$1.73 (0.036)	0.006 (0.827)	0.025 (0.000)	\$3.20 (0.000)	0.008 (0.053)	0.034 (0.010)
Zambia	\$3.72 (0.000)	-0.024 (0.000)	0.025 (0.564)	\$1.98 (0.005)	0.034 (0.000)	0.020 (0.004)	\$2.11 (0.001)	0.027 (0.000)	0.017 (0.000)
Pooled	\$2.10 (0.000)	0.036 (0.000)	-0.018 (0.011)	\$0.75 (0.009)	-0.044 (0.489)	0.015 (0.000)	\$3.08 (0.000)	0.024 (0.000)	0.040 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Second, among rural households the slopes of processed maize consumption patterns are negative below the expenditure breakpoint (\$1.46 pooled), and positive above the breakpoint. This result is consistent with poor rural households increasing their income with increased production of maize, a substitute for processed maize, yet choose to acquire processed grain at low levels of total expenditure.

Third, non-maize processed food drives the income-induced growth of processed share in urban areas as shown by positive slopes of processed share at expenditures above the breakpoint and negative slopes of processed maize at expenditures.

Overall, the non-parametric and semi-parametric expenditure analysis suggest that there is significant change in processed food consumption patterns among poor households in ESA, and that the changing consumption patterns are often more rapid for the poor than for the non-poor.

2.4.3 Descriptive statistics of processed food consumption by city size

Table 2-11 shows the food budget shares of processed food across settlement types that are consistent with the hypothesis of increased consumption of processed food by households within larger cities, which the results show with each country having a positive monotonic relationship between processed share and city size.

Table 2-11: Food budget shares of processed food, aggregated country and settlement

	Processed Share				
	Pooled	Malawi	Tanzania	Uganda	Zambia
National	42.2 (40.6)	40.9 (40.0)	46.1 (45.7)	35.9 (32.8)	44.4 (44.7)
Rural	35.6 (32.3)	37.0 (35.1)	39.0 (36.5)	30.7 (26.6)	33.2 (28.9)
Urban	62.7 (66.8)	62.5 (66.3)	66.2 (69.8)	54.7 (56.2)	65.4 (68.4)
Primary City	71.0 (73.1)		76.8 (77.8)	61.3 (59.8)	69.2 (70.6)
Secondary City	66.1 (68.4)	64.2 (67.7)	68.8 (70.8)	53.2 (52.7)	66.5 (68.8)
Tertiary City	56.2 (60.1)	56.6 (60.1)	58.7 (62.3)	52.4 (55.5)	59.4 (63.5)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

The food budget shares of the processed food aggregates across city size in Table 2-12 from which I highlight two consumption patterns. First, three of the five aggregates display a monotonic relationship between processed share and increasing city size; negative for unprocessed maize and unprocessed non-maize, positive for processed low non-maize. Table 2-A-3, the country-specific form of Table 2-12 found in the appendix shows these same monotonic patterns³⁷, strengthening the pattern observed in the pooled data. These patterns are consistent with the patterns of Table 2-11, as they show greater consumption of processed food within larger cities.

Table 2-12: Food budget shares of processed food aggregates, pooled data by settlement

Pooled Data	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
National	11.6 (1.4)	6.5 (0.0)	46.2 (43.1)	18.6 (16.1)	17.1 (11.4)
Rural	14.0 (6.5)	5.9 (0.0)	50.4 (48.9)	16.5 (13.3)	13.2 (8.6)
Urban	4.0 (0.0)	8.4 (5.1)	33.2 (30.6)	25.2 (24.9)	29.1 (24.5)
Primary City	0.4 (0.0)	8.0 (5.4)	28.7 (26.6)	25.0 (24.5)	37.9 (34.1)
Secondary City	3.1 (0.0)	11.6 (8.5)	30.8 (28.8)	26.9 (26.8)	27.7 (24.5)
Tertiary City	6.6 (0.0)	6.7 (2.8)	37.2 (33.6)	24.2 (24.1)	25.4 (20.3)

Source: authors' calculations using national household level surveys

³⁷ Unprocessed maize and unprocessed non-maize food budget shares in Uganda minimally deviate from the monotonicity of the pattern.

Second, low processed foods are likely the “entry point” for processed foods among rural- and low-income urban households, as they show no clear pattern with respect to city size gradations (Table 2-12). Rising incomes and denser urban settlement likely drive a transition towards more highly processed foods, suggesting that highly processed foods are the driving force behind the consumption patterns of processed food.

2.4.4 Nonparametric analysis of the relationship between distance to cities and processed food consumption

Figure 2-4 displays the LOWESS curve estimates of the relationship between distance to city and processed share. Each curve, representing the relationships between processed share and distances to cities of difference size, displays an inverse relationship between distance to city and processed share, which is consistent with previous studies that have studied the effects of distance to urban centers on lifestyle patterns (Fafchamps and Shilpi, 2003; Sharma, 2016). In addition to these negative relationships, the proximity to larger cities has a larger effect on processed share than distances to smaller cities. The difference in the effects of distance by city size is consistent with the conceptual model that includes the attenuation of the effects of urban environment that are sensitive to city size.

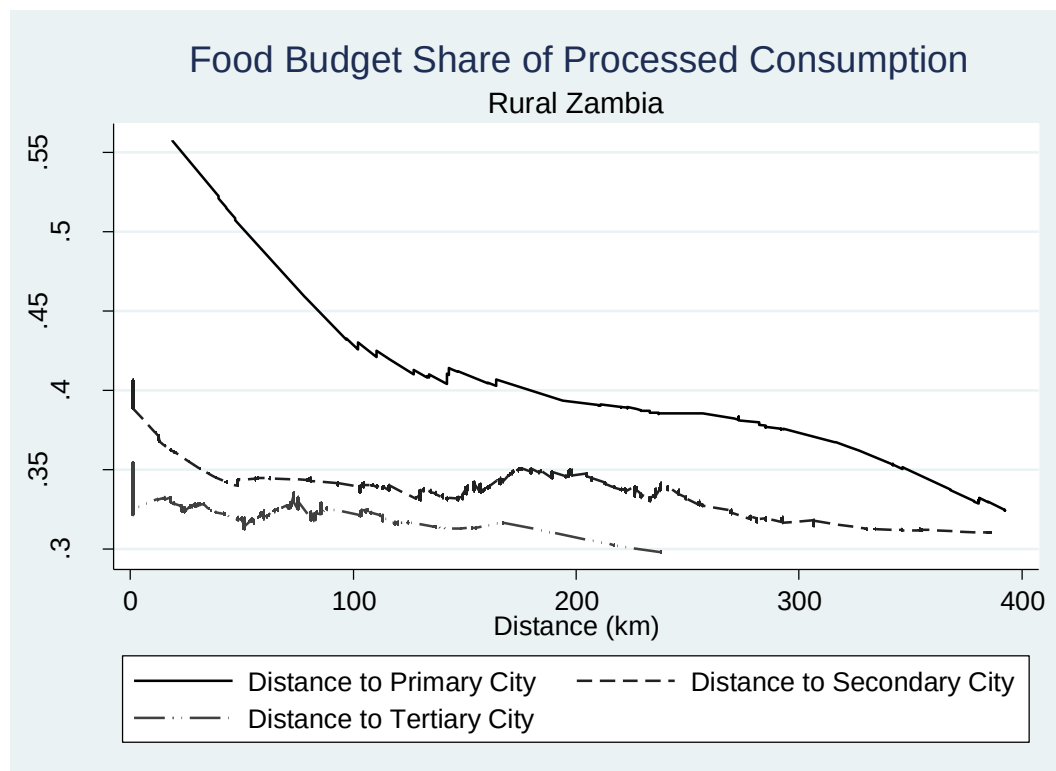


Figure 2-4: Locally weighted scatterplot smoothing curves representing the typical processed share relative to household distance to city

Note: LOWESS curve estimates use a bandwidth equal to 0.4.

2.4.5 Engel's Curve Model regression analysis

The fractional probit regressions analysis builds upon the previous analyses by isolating the impact of determinant variables using the full Engel's Curve Model. I selected the functional form of the Engel curve suggested by Banks, Blundell and Lewbel (1997). The descriptive analyses highlighted the need for the functional form to allow for changing slopes of the food consumption patterns, which narrowed the functional form options to those suggested by Banks et al. (1997) and Leser (1963). Due to superior goodness of fit of the fractional probit regressions I selected the Banks et al. suggested functional form. This functional form includes non-income independent variables and two independent variables for expenditure: log total household expenditure and the squared log of total household expenditure. As expenditure is included in the

model as two separate variables, I calculate the average partial effects of expenditure with the following equation:

$$\frac{\partial}{\partial(\text{expenditure})} = \left[a \frac{1}{\text{expenditure}} + 2b \frac{\ln(\text{expenditure})}{\text{expenditure}} \right] * E\phi[X\beta] \quad (2.5)$$

where a equals the marginal effects of log expenditure, b equals the marginal effects of squared log expenditure, and the scaling factor in the right of the equation is equal to the average of the normal densities of the predicted values of the dependent variable.

Here I first consider the effects of expenditure on processed food consumption and then the effects non-expenditure variables, including the effects of city size and distance to cities, on processed food consumption.

Households with higher total expenditures consume larger shares of processed food, as shown by the positive average partial effects of total expenditure on processed share in Tables 2-13 & 2-14. This finding is consistent with previous analyses on the consumption of processed food (Veeck & Veeck, 2000; Pingali, 2007; Popkin et al., 2012; Tschirley et al., 2015), and provides a base relationship for the following three findings.

Table 2-13: Average partial effects of expenditure and city size on processed share – urban data

Urban	Processed Share			
	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.007 (0.004)	0.017 (0.000)	0.009 (0.001)	0.010 (0.000)
Household located in a primary city		0.068 (0.000)	0.000 (1.000)	0.037 (0.000)
Household located in a secondary city	0.053 (0.000)	0.047 (0.003)	-0.053 (0.049)	0.043 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 2-14: Average partial effects of expenditure and distances to various city sizes on processed share – rural data
(marginal effects of distance only for rural Zambia)

Rural	Processed Share			
	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.004 (0.287)	0.009 (0.062)	0.022 (0.000)	0.010 (0.009)
Distance to primary city (log)				-0.045 (0.000)
Distance to secondary city (log)				-0.002 (0.554)
Distance to tertiary city (log)				-0.005 (0.001)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

First, the marginal effects of expenditure on processed share are positive at all levels of income (Figure 2-5)³⁸, and among rural households, the marginal effects are positive in all countries at expenditure levels above \$1.55. These effects show no supernumerary level of expenditure in urban areas and, if one exists, in rural areas this level of expenditure would be below the poverty line. The differences from the slopes of processed share represented by LOWESS curves (Figure 2-2) necessitates the later analysis of additional determinants affecting processed food consumption.

³⁸ The graphed negative marginal effect of expenditure on processed share in Uganda is theoretical, as the dataset does not include households at this level of total expenditure.

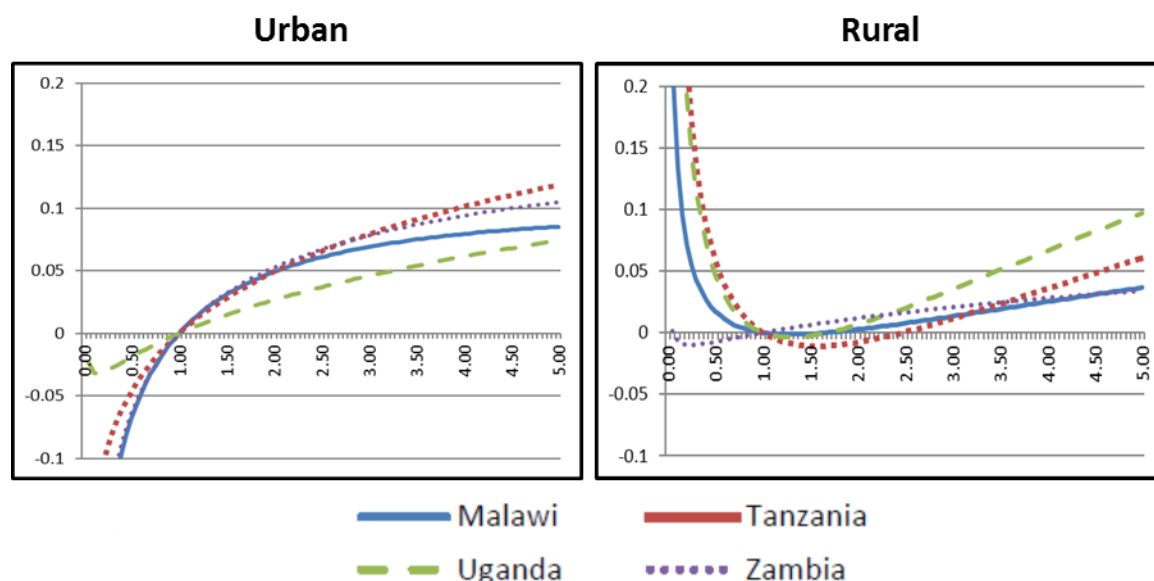


Figure 2-5: Estimated effects of expenditure on purchased share and the consumption of processed food

Second, processed maize is not driving the transition towards processed food within urban areas, as expenditure positively affects the consumption of processed non-maize aggregates and negatively affects maize and unprocessed non-maize aggregates (Table 2-A-6).

Third, there is evidence of the adherence to Bennett's Law within rural areas, and that household desire of non-staple food at higher incomes contributes to the increased consumption of processed food. Total expenditure has a positive effect on all non-maize aggregates other than in rural Uganda, where households commonly consider matoke to be the primary staple food and matoke is primarily included in the non-maize unprocessed sub-aggregate (Table 2-A-7).

The fractional probit regression analysis reveals three key results specific to non-expenditure determinants of urban processed food consumption. First, larger city size positively affects the consumption of processed food (Table 2-13). Three of the four countries show significant marginal effects of the city size dummy variables that indicate a positive relation, with only secondary cities in Uganda consuming a smaller share of processed food than do

households in tertiary cities. There is not a pattern of relative marginal effects between primary and secondary cities, merely the significant difference between cities with populations above one hundred thousand and those below. This result supports the hypothesis that larger cities influence households to consume greater shares of potentially timesaving processed food, relative to households within smaller cities.

Second, under further disaggregation of the consumption of processed food, I show that household settlement in larger cities continues to affect processed food consumption (Table 2-A-6). In each of the four countries, the marginal effects of city size on unprocessed and processed maize shares show that households in larger cities favor the purchase of the latter. In three of the four countries, the marginal effects of primary and secondary cities on the consumption of unprocessed food is negative when significant, indicating that increased city size is associated with less consumption of unprocessed food. Lastly, the significant marginal effects of city size on low processed food are mixed and on high processed food are positive. Given the positive marginal effect of large cities on processed consumption, this result supports the earlier descriptive results that indicate that highly processed food drives overall growth of processed food consumption in urban areas.

Third, five additional effects of non-expenditure determinant variables for the urban consumption of processed food are of note (Tables 2-A-4 & 2-A-6). First, households with greater participation in nonfarm employment consume greater shares of processed food, specifically greater shares of highly processed food and lesser shares of unprocessed maize. This finding is consistent with literature that indicates that nonfarm employment increases the opportunity cost of time, which would adjust household preferences to favor timesaving food items (Senauer et al., 1986; Kennedy and Reardon, 1996). Second, dependency ratio, another

consumer demand factor of individual and household preferences, generally has a negative effect on processed food consumption, specifically highly processed food that show negative effects of dependency ratio in each country. Third, the amount of land farmed negatively affects processed share in each country as the production of food enabled by the farming of land directly provides the opportunity to consume food from own production, which by definition is unprocessed. Fourth, ownership of a telephone has a positive effect on processed share, which is in agreement with the hypothesis that ownership of a phone suggests individual and household preferences that are consistent with modern lifestyle patterns. Fifth, the effect of the age of the household head, although minimal, is negative and significant on processed share, affirming that elderly households are less open to modern lifestyle patterns.

I will highlight four results pertaining to the marginal effects of non-expenditure determinants on processed food consumption by rural households.

First, distance to cities has a negative effect on processed share in rural Zambia (Table 2-14). This is consistent with the von Thünen model that says that distance to cities increase transaction costs of acquiring foods from urban markets (Nelson, 2002), thereby increasing the effective price and reducing household access to processed foods.

Second, the effect of distance to a city has a monotonically increasing negative effect on processed share across city sizes. Table 2-6 shows greater consumption of processed food in primary cities, therefore it is consistent that households near primary cities would also consume greater shares of processed food.

Third, further examining processed foods (Table 2-A-7), I find that, distance to primary and secondary cities are significant determinants for all processing sub-aggregates, other than low processed food. Distance to tertiary cities is only statistically significant for high processed

foods, and this determinant has a negative marginal effect that is in agreement with the marginal effects of distance to primary or secondary cities on high processed food. The negative marginal effects of distance to all cities, including tertiary cities, reveal household preference for high processed food and the ability of cities of any size to provide it. Unexpected results are the negative marginal effects of distances to larger cities on unprocessed maize. The positive marginal effect of distance to cities on unprocessed non-maize partially explains the unexpected result as unprocessed non-maize is an aggregate that includes cassava, which households in rural Zambia consider as a substitute for unprocessed maize.

Fourth, determinant variables other than proximity to city have significant effects on processed food consumption providing four observations of note. First, increased opportunity cost of time from household participation in nonfarm employment is significant to the processed food consumption of rural households as indicated by positive marginal effects on processed food in each country (Tables 2-A-5). Nonparametric analysis of the relationship between nonfarm employment and total household expenditure shows a positive relationship between these variables beginning at incomes as low as one dollar per day; this relationship contributes to the rise in processed share beginning among the poor. Haggblade et al. (2010) noted that households in rural Africa increasingly depend on nonfarm employment; therefore, this finding leads to an expectation of future increased processed food consumption at low expenditures. Second, the average partial effects of rural farmed hectares of land are negative for processed share and all statistically significant sub-aggregates for processed food. The amount of farmed hectares of land positively relates to household total expenditure; therefore, the amount of farmed hectares of land partially offsets a portion of the positive effects of expenditure on processed share. Third, high levels of education positively influence the consumption of

processed share in Malawi and Zambia, and when statistically significant it reduces the unprocessed sub-aggregates and increases the processed sub-aggregates (Table 2-A-7). Fourth, female head of household has a negative effect on highly processed food, which, assuming that highly processed foods are less healthy aligns with female-headed households selecting healthier eating patterns (Smith, 2003).

2.4.6. Mediation model analysis

City size affects multiple determinant variables within the Engel's Curve Model, therefore I use mediation model analysis to estimate the indirect effects of city size on processed share that occur through household total expenditure, nonfarm labor, the amount of land farmed by the household and the maximum education attained within the household. I select these four variables as mediator variables for the following reasons: difference in city size could affect household expenditure (Ferré et al., 2012), it could expose household members to various nonfarm employment opportunities, it could vary household access to farmland, and larger cities could offer greater opportunities for education. I do not find statistically significant indirect effects of city size with the Malawi data, nor via education, but the remaining countries and mediator variables led to the following results.

Table 2-15 shows the indirect effects that primary and secondary cities have on processed share with three results of note. First, primary cities in Tanzania and Zambia have positive indirect effects on processed share, suggesting a stronger positive effect of primary cities on processed share than shown in Tables 2-13 & 2-A-4. Second, primary cities in Tanzania along with primary and secondary cities in Zambia significantly increase household participation in nonfarm employment, which positively effects processed share, resulting in positive and

significant indirect effects of city size. Third, city size indirectly increases the consumption of processed food via the amount of household farmed land. Larger cities provide households with less opportunity for farming larger areas of land³⁹, and land can be used for the production of processed food, therefore city size increases processed share by reducing the consumption of an alternative food aggregate. The positive indirect effects of city size via hectares of farmed land on processed share show that city size reduces farmed land, which would have reduced processed share, therefore city size increases processed food consumption beyond what direct effects would show.

Table 2-15: Indirect effects of urban household settlement within primary and secondary cities on processed share, estimates by country

Mediator Variable	Malawi		Tanzania		Uganda		Zambia	
	Primary Cities	Secondary Cities	Primary Cities	Secondary Cities	Primary Cities	Secondary Cities	Primary Cities	Secondary Cities
Household total expenditure per AE		-0.005 (0.574)	0.073 (0.000)	-0.004 (0.658)	0.025 (0.153)	0.017 (0.522)	0.029 (0.000)	0.002 (0.565)
Nonfarm employment		0.000 (0.732)	0.007 (0.011)	0.005 (0.090)	0.002 (0.646)	0.000 (0.897)	0.002 (0.000)	0.001 (0.015)
Hectares of farmed land (log)		0.001 (0.382)	0.009 (0.019)	0.008 (0.047)	0.007 (0.073)	0.011 (0.023)	0.015 (0.000)	0.008 (0.000)
Maximum education		-0.001 (0.286)	0.000 (0.831)	0.000 (0.835)	0.004 (0.182)	0.005 (0.290)	0.000 (0.948)	0.000 (0.933)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

I estimate the indirect effects of distances to cities on rural processed food consumption patterns using the same mediator variables that I use while estimating urban indirect effects (Table 2-16). The overall finding from the rural mediation analysis is that increasing distance to cities decreases processed share through indirect negative effects on household income, nonfarm employment, and hectares of farmed land. The one anomalous result is the positive indirect for secondary cities on processed share via hectares of farmed land. Increasing rural household

³⁹ The percentage of urban households that engage in farming also has an inverse relationship with city size (primary city 9%, secondary city 29%, tertiary city 55%).

distance from secondary cities results in statistically significant reduction in farmed land, causing this positive effect. The use of the mediator model highlights the indirect effects that testing for direct effects in typical regression analysis do not capture.

Table 2-16: Indirect effects of distance from rural household settlement to primary, secondary or tertiary cities on processed share – rural Zambia

Mediator Variable	Distance to		
	Primary Cities	Secondary Cities	Tertiary Cities
Household total expenditure per AE	-0.001 (0.075)	0.000 (0.203)	0.000 (0.337)
Nonfarm employment	-0.005 (0.000)	0.000 (0.455)	-0.001 (0.002)
Hectares of farmed land (log)	-0.002 (0.006)	0.002 (0.000)	0.000 (0.538)
Maximum education attained within the household	0.000 (0.534)	0.000 (0.975)	0.000 (0.079)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

2.5 Conclusions

I considered how the consumption patterns of processed food varies across income, by city size and by distances to cities of differing size. To my knowledge, this is the first paper to analyze the spatial effects of city size and distance to city on processed food consumption in Africa.

Processed foods have penetrated the diets of rural (36%) and urban (63%) households at all levels across the income distribution of ESA, which is both a result of increased market access because of public and private investments and an indication of demand for future processing investments within the agrifood system. I find that income-induced diet change that begins among the poor contribute to these high levels of processed food consumption. The pattern of a rising share of processed food value with increasing total expenditure in the diets of poor households consistently appears in the analysis of descriptive statistics, nonparametric

analysis, expenditure threshold analysis and Engel's curve regression analysis. These findings highlight that the rise in processed food consumption is not solely a middle-class story, as conventionally assumed, but poor households are purchasing significant shares of processed food.

Our analysis of sub-aggregates of processed food suggests that the consumption of highly processed food is the driving force behind the consumption pattern of processed food in general, where low processed food serves as a transition from unprocessed to highly processed food. This contributes to discussions regarding nutrition and the sophistication of the processing sector.

Given the large portion of the population in ESA that currently has low income, coupled with the pattern of increased processed share with income growth, these findings signal a strong future demand for increased food market infrastructure.

The future growth of investments into the agrifood system should consider proximity to urban areas as I found that the consumption patterns of processed food significantly vary by households' settlement in or near cities of various size. I consistently observe these patterns across the four-step analysis that includes descriptive, nonparametric, fractional probit regression, and mediation model analyses.

City size has significant positive effects on processed food consumption in urban households, affirming the expectation that city size will alter the urban environment such that households will have greater access to processed foods and the consumer demand characteristics will exhibit an increased desire of processed food.

The distance to cities of any size has a diminishing negative effect on processed food consumption by rural households. The distance to city effects with the greatest magnitudes were for primary cities, which are likely to have significant commercial networks that extend market

reach further into rural areas that result in increased engagement with urban markets by rural households.

I show that the spatial effects of city size and distance to city are significant drivers of processed food consumption in ESA. As urbanization is expected to continue within this region and elsewhere in developing nations, the recognition of how urbanization occurs should be considered when effort is put forth to meet the changing food demands of households near and within urban areas.

APPENDIX

APPENDIX

Table 2-A-1: Relationship between food classification scheme in this paper and that in Monteiro et al. (2010)

Processing Aggregates	Unprocessed	Low Processed		High Processed	
Monteiro Aggregates	Unprocessed or Minimally Processed	Processed Culinary or Ingredients		Ultra-Processed	
	Whole grains	Rice	Sugar	Vegetable oils	Breads
	Beans & pulses	Butchered meat	Spices	Vegetable fats	Biscuits
	Fresh fruits	Dried fish	Wheat flour	Pasta	Soft drinks
	Fresh vegetables		Maize meal		Cheeses
	Roots & tubers		Other flours		Processed meats
	Fresh fish			<i>Pasteurized milk*</i>	Prepared food away from home
	Eggs				

Note: Pasteurized milk is part of Monteiro et al.'s Unprocessed or Minimally Processed aggregate

Table 2-A-2: Food budget shares of processed food aggregates, by country and expenditure

Rural	Total	Average Maize Unprocessed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50- 1.00	\$1.00- 1.50	\$1.50- 2.00	\$2.00- 3.00	\$3.00- 5.00	\$5.00- 10.00	> \$10.00	Low	Middle	High
Malawi	20.7 (17.5)	20.5 (0.0)	25.5 (25.3)	23.7 (23.7)	21.2 (20.2)	17.6 (16.4)	13.0 (11.1)	9.3 (6.5)	12.7 (3.9)	24.7 (23.5)	21.2 (20.3)	14.4 (12.2)
Tanzania	15.6 (10.3)	14.4 (0.0)	16.8 (5.3)	18.4 (11.0)	19.7 (17.4)	15.7 (14.2)	11.7 (8.6)	5.9 (1.8)	2.0 (0.0)	18.2 (11.5)	16.6 (14.3)	10.1 (6.9)
Uganda	5.6 (0.0)	1.9 (0.0)	7.1 (0.0)	5.4 (0.0)	6.5 (0.0)	5.7 (0.0)	5.3 (0.0)	3.0 (0.0)	2.2 (0.0)	6.2 (0.0)	6.0 (0.0)	3.9 (0.0)
Zambia	21.2 (16.5)	25.7 (18.5)	24.2 (20.3)	22.4 (19.0)	20.8 (17.7)	18.4 (13.9)	15.7 (9.3)	8.5 (5.1)	8.6 (2.4)	24.2 (19.7)	20.3 (16.8)	14.7 (8.4)
Pooled	14.0 (6.5)	17.7 (0.0)	18.7 (9.1)	16.6 (9.1)	16.2 (10.8)	12.6 (7.1)	9.6 (5.3)	5.2 (1.1)	3.9 (0.0)	16.4 (5.3)	14.5 (9.5)	9.3 (5.1)

Table 2-A-2: (continued)

Urban	Total	Average Maize Unprocessed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	5.3 (0.0)	11.3 (0.0)	8.8 (0.0)	11.5 (0.0)	7.1 (0.0)	4.4 (0.0)	3.4 (0.0)	1.9 (0.0)	7.6 (0.0)	14.2 (0.0)	6.0 (0.0)	3.8 (0.0)
Tanzania	5.4 (0.0)	14.5 (24.0)	8.2 (0.0)	6.5 (0.0)	9.6 (0.0)	10.1 (0.0)	4.4 (0.0)	1.6 (0.0)	0.7 (0.0)	8.2 (0.0)	9.8 (0.0)	2.7 (0.0)
Uganda	1.6 (0.0)	1.3 (0.0)	2.6 (0.0)	7.1 (0.0)	3.1 (0.0)	2.2 (0.0)	1.6 (0.0)	0.9 (0.0)	0.2 (0.0)	4.5 (0.0)	1.8 (0.0)	1.0 (0.0)
Zambia	3.4 (0.0)	16.2 (0.0)	10.5 (0.0)	7.4 (0.0)	6.0 (0.0)	3.8 (0.0)	2.5 (0.0)	1.4 (0.0)	0.8 (0.0)	10.3 (0.0)	5.2 (0.0)	1.8 (0.0)
Pooled	4.0 (0.0)	11.3 (0.0)	7.9 (0.0)	7.4 (0.0)	7.3 (0.0)	6.8 (0.0)	3.2 (0.0)	1.4 (0.0)	1.1 (0.0)	7.9 (0.0)	6.7 (0.0)	2.2 (0.0)

Rural	Total	Average Maize Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	9.9 (0.0)	24.8 (0.0)	16.2 (0.0)	10.3 (0.0)	7.3 (0.0)	5.4 (0.0)	5.1 (0.0)	4.5 (0.0)	3.8 (1.3)	15.6 (0.0)	7.6 (0.0)	5.0 (0.0)
Tanzania	5.8 (0.0)	13.6 (0.0)	10.1 (0.0)	7.1 (0.0)	4.7 (0.0)	4.1 (0.0)	4.0 (0.0)	4.7 (0.0)	3.2 (2.5)	8.1 (0.0)	4.3 (0.0)	4.1 (0.0)
Uganda	4.5 (0.0)	8.4 (0.0)	6.6 (0.0)	4.7 (0.0)	5.5 (0.0)	3.7 (0.0)	4.5 (0.0)	2.5 (0.0)	2.3 (0.0)	5.6 (0.0)	3.8 (0.0)	3.9 (0.0)
Zambia	4.7 (0.0)	7.2 (0.0)	3.9 (0.0)	4.6 (0.0)	4.5 (0.0)	5.1 (0.0)	5.4 (0.0)	5.3 (2.1)	5.0 (2.8)	4.5 (0.0)	4.6 (0.0)	5.3 (0.0)
Pooled	5.9 (0.0)	14.2 (0.0)	9.5 (0.0)	6.7 (0.0)	5.3 (0.0)	4.2 (0.0)	4.3 (0.0)	3.8 (0.0)	3.0 (0.0)	8.0 (0.0)	4.7 (0.0)	4.3 (0.0)

Urban	Total	Average Maize Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	12.5 (10.1)	41.1 (51.1)	25.7 (27.1)	15.1 (15.8)	15.8 (14.9)	13.9 (14.3)	10.8 (9.9)	9.0 (7.3)	6.3 (4.3)	18.8 (14.7)	16.8 (16.5)	10.2 (8.7)
Tanzania	7.8 (5.1)	3.2 (0.0)	12.8 (9.0)	14.3 (9.2)	9.2 (6.3)	8.8 (5.9)	7.7 (6.2)	5.6 (4.9)	3.0 (2.4)	12.8 (9.2)	8.7 (5.7)	6.1 (4.6)
Uganda	5.0 (2.0)	7.8 (0.0)	13.5 (9.3)	9.0 (0.0)	7.6 (0.0)	6.2 (2.7)	5.3 (3.0)	3.8 (2.0)	1.6 (0.8)	9.5 (0.0)	5.3 (2.0)	4.1 (2.2)
Zambia	13.0 (10.7)	17.9 (6.3)	19.0 (15.0)	18.1 (17.8)	16.8 (16.3)	15.7 (15.3)	13.0 (12.3)	10.0 (9.0)	6.3 (5.3)	18.2 (15.5)	16.9 (16.5)	10.8 (9.4)
Pooled	8.4 (5.1)	12.0 (0.0)	15.4 (10.4)	14.4 (10.9)	11.1 (7.3)	9.9 (6.6)	8.2 (6.3)	5.9 (4.6)	3.6 (2.5)	12.9 (8.1)	10.1 (5.8)	6.8 (4.8)

Table 2-A-2: (continued)

Rural	Total	Average Non-Maize Unprocessed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	42.3 (41.0)	41.1 (39.2)	39.8 (38.5)	42.2 (40.9)	43.5 (43.0)	44.2 (43.7)	44.0 (41.4)	38.7 (35.9)	31.0 (26.5)	40.5 (38.9)	43.6 (43.7)	42.9 (41.2)
Tanzania	45.3 (43.0)	35.0 (29.7)	47.1 (43.4)	47.0 (43.1)	45.3 (41.9)	46.2 (44.8)	45.3 (43.0)	39.7 (38.8)	32.3 (32.2)	45.9 (42.6)	46.0 (44.4)	43.6 (41.7)
Uganda	63.8 (65.6)	72.1 (82.6)	66.3 (65.0)	68.1 (70.1)	66.1 (71.3)	65.7 (67.6)	60.8 (63.8)	53.4 (51.0)	33.1 (25.7)	67.4 (70.1)	65.4 (67.7)	55.5 (55.5)
Zambia	45.6 (43.1)	45.2 (42.9)	48.5 (45.8)	46.9 (45.4)	44.9 (43.4)	44.0 (42.0)	40.8 (37.7)	35.4 (31.7)	32.3 (28.1)	47.5 (45.5)	45.4 (43.2)	40.0 (37.7)
Pooled	50.4 (48.9)	41.1 (38.0)	49.2 (46.7)	51.5 (49.5)	51.0 (49.4)	52.7 (51.2)	50.8 (50.8)	44.9 (45.3)	32.5 (28.7)	51.7 (50.1)	51.5 (50.1)	46.8 (45.7)

Urban	Total	Average Non-Maize Unprocessed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	32.2 (30.9)	34.1 (31.8)	34.9 (37.2)	37.1 (36.7)	36.6 (34.3)	33.7 (32.4)	30.8 (29.8)	29.2 (27.9)	26.0 (25.6)	35.8 (35.5)	36.2 (34.6)	30.4 (28.9)
Tanzania	28.4 (26.1)	41.1 (31.1)	28.7 (26.6)	37.3 (37.4)	35.5 (31.8)	30.5 (29.1)	27.7 (25.1)	24.2 (22.3)	20.5 (19.7)	35.0 (34.1)	31.5 (30.0)	25.1 (23.5)
Uganda	43.7 (42.2)	25.4 (21.4)	52.7 (46.4)	55.0 (52.3)	50.5 (52.9)	55.3 (55.9)	45.6 (41.9)	38.2 (38.3)	30.7 (32.2)	50.8 (49.8)	53.3 (54.4)	39.1 (38.3)
Zambia	31.2 (29.4)	37.8 (35.7)	39.2 (35.9)	36.2 (34.5)	34.7 (33.2)	32.5 (30.3)	30.3 (29.2)	28.8 (27.7)	26.7 (25.5)	38.0 (35.3)	34.3 (32.8)	29.1 (28.0)
Pooled	33.2 (30.6)	35.8 (31.1)	35.2 (34.1)	39.7 (37.4)	38.5 (35.0)	36.5 (32.1)	33.5 (30.6)	29.9 (28.1)	25.3 (24.2)	39.3 (37.2)	37.5 (32.7)	30.2 (28.3)

Rural	Total	Average Non-Maize Low Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	18.9 (16.3)	12.3 (9.2)	14.8 (12.4)	17.6 (15.0)	19.5 (17.2)	22.1 (20.5)	23.3 (22.0)	27.0 (26.1)	25.1 (24.7)	15.1 (12.5)	19.5 (17.2)	23.4 (21.9)
Tanzania	14.9 (11.4)	21.1 (12.5)	12.6 (5.8)	12.7 (7.9)	14.4 (10.6)	15.9 (13.2)	15.3 (13.3)	19.3 (17.8)	25.5 (24.7)	13.4 (8.2)	15.5 (12.0)	16.6 (14.4)
Uganda	16.6 (13.8)	9.9 (2.3)	14.0 (9.3)	14.5 (8.8)	14.3 (10.8)	16.6 (13.4)	19.2 (17.7)	22.0 (22.0)	20.0 (15.6)	13.9 (9.2)	16.7 (13.8)	20.9 (19.7)
Zambia	18.9 (15.2)	15.3 (10.4)	16.2 (12.3)	17.5 (13.6)	19.9 (16.4)	21.1 (18.0)	24.2 (22.7)	30.0 (29.1)	30.5 (27.3)	16.3 (12.3)	19.7 (16.1)	24.7 (22.7)
Pooled	16.5 (13.3)	17.0 (10.0)	14.1 (9.9)	14.8 (10.5)	15.7 (12.6)	17.2 (14.3)	18.1 (16.0)	21.7 (19.4)	23.2 (23.1)	14.2 (9.8)	17.0 (13.8)	19.7 (17.8)

Table 2-A-2: (continued)

Urban	Total	Average Non-Maize Low Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	27.7 (27.3)	11.6 (17.0)	22.3 (19.2)	22.8 (23.1)	25.0 (25.0)	28.4 (28.9)	31.0 (30.2)	30.3 (31.0)	25.1 (25.3)	20.8 (19.0)	25.5 (25.4)	29.5 (29.6)
Tanzania	22.4 (22.4)	25.1 (22.9)	25.1 (24.4)	20.6 (18.8)	20.5 (21.1)	21.2 (21.7)	24.4 (25.0)	23.2 (23.3)	19.7 (18.1)	21.8 (21.4)	21.1 (21.7)	23.2 (23.0)
Uganda	25.6 (25.1)	29.9 (30.9)	20.2 (13.7)	21.3 (14.7)	23.2 (18.8)	19.6 (16.5)	27.8 (25.8)	28.7 (27.6)	23.6 (25.1)	21.7 (14.7)	23.4 (20.2)	27.1 (26.2)
Zambia	30.4 (29.9)	19.0 (16.2)	18.0 (16.2)	23.9 (22.3)	25.9 (24.5)	28.6 (27.7)	31.3 (30.5)	34.3 (33.8)	36.8 (36.0)	20.2 (18.3)	26.4 (25.5)	33.3 (32.5)
Pooled	25.2 (24.9)	23.3 (22.4)	22.5 (21.0)	21.7 (20.5)	22.6 (22.6)	22.9 (22.5)	27.1 (26.6)	27.3 (26.7)	25.1 (24.4)	21.5 (19.6)	23.0 (22.7)	26.8 (26.3)

Rural	Total	Average Non-Maize High Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	8.2 (5.1)	1.4 (0.0)	3.7 (1.1)	6.1 (3.8)	8.5 (6.0)	10.8 (8.8)	14.6 (12.1)	20.5 (19.2)	27.4 (26.0)	4.1 (1.4)	8.0 (5.6)	14.2 (11.4)
Tanzania	18.3 (14.0)	16.0 (11.7)	13.5 (9.8)	14.8 (11.0)	15.9 (12.9)	18.1 (13.9)	23.7 (18.5)	30.4 (27.8)	37.0 (36.3)	14.5 (10.7)	17.6 (13.7)	25.6 (21.7)
Uganda	9.5 (4.7)	7.6 (0.0)	6.1 (3.7)	7.3 (2.8)	7.5 (3.9)	8.2 (4.6)	10.2 (5.9)	19.0 (15.6)	42.4 (27.0)	7.0 (3.3)	8.1 (4.6)	15.9 (9.0)
Zambia	9.6 (6.9)	6.6 (3.1)	7.2 (4.8)	8.7 (6.3)	9.9 (7.7)	11.4 (8.7)	13.9 (11.4)	20.8 (19.2)	23.6 (19.8)	7.4 (4.9)	10.0 (7.7)	15.2 (12.0)
Pooled	13.2 (8.6)	9.9 (5.6)	8.5 (4.7)	10.4 (6.9)	11.7 (8.3)	13.3 (9.6)	17.1 (12.0)	24.4 (21.2)	37.4 (32.1)	9.8 (5.8)	12.4 (8.8)	19.9 (15.2)

Urban	Total	Average Non-Maize High Processed Share by Expenditure Strata								Average Processed Share by Expenditure Terciles		
		< \$0.50	\$0.50-1.00	\$1.00-1.50	\$1.50-2.00	\$2.00-3.00	\$3.00-5.00	\$5.00-10.00	> \$10.00	Low	Middle	High
Malawi	22.2 (21.1)	1.8 (0.0)	8.4 (4.6)	13.5 (12.2)	15.5 (14.4)	19.6 (19.3)	23.9 (22.5)	29.6 (27.9)	35.0 (35.8)	10.4 (9.3)	15.6 (14.4)	26.1 (24.6)
Tanzania	36.0 (33.5)	16.2 (14.7)	25.2 (21.5)	21.3 (16.5)	25.2 (20.2)	29.3 (25.8)	35.8 (33.1)	45.3 (45.9)	56.1 (56.3)	22.2 (19.1)	29.0 (25.6)	43.0 (42.3)
Uganda	24.1 (17.7)	35.6 (10.7)	11.0 (6.4)	7.6 (3.0)	15.6 (6.4)	16.7 (10.1)	19.7 (16.4)	28.4 (22.6)	43.9 (33.5)	13.6 (5.8)	16.3 (10.1)	28.7 (22.8)
Zambia	21.9 (20.3)	9.1 (5.2)	13.3 (11.4)	14.3 (12.0)	16.7 (14.1)	19.4 (18.3)	22.9 (21.6)	25.6 (24.1)	29.4 (27.6)	13.3 (11.4)	17.1 (15.0)	24.9 (23.4)
Pooled	29.1 (24.5)	17.6 (10.7)	18.9 (13.8)	16.8 (13.0)	20.5 (16.5)	23.9 (19.5)	28.0 (24.3)	35.5 (30.9)	44.9 (40.8)	18.3 (13.4)	22.6 (17.7)	34.0 (29.3)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

Table 2-A-3: Food budget shares of processed food aggregates, country data by settlement

Malawi Data	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
National	18.3 (14.0)	10.3 (0.0)	40.7 (38.8)	20.3 (18.3)	10.4 (6.6)
Rural	20.7 (17.5)	9.9 (0.0)	42.3 (41.0)	18.9 (16.3)	8.2 (5.1)
Urban	5.3 (0.0)	12.5 (10.1)	32.2 (30.9)	27.7 (27.3)	22.2 (21.1)
Primary City	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Secondary City	4.2 (0.0)	13.2 (11.1)	31.6 (30.4)	28.1 (27.7)	22.8 (21.7)
Tertiary City	9.0 (1.6)	10.0 (3.9)	34.4 (32.9)	26.4 (25.4)	20.3 (17.8)

Tanzania Data	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
National	13.0 (4.9)	6.3 (0.0)	41.0 (37.4)	16.9 (14.3)	22.9 (17.8)
Rural	15.6 (10.3)	5.8 (0.0)	45.3 (43.0)	14.9 (11.4)	18.3 (14.0)
Urban	5.4 (0.0)	7.8 (5.1)	28.4 (26.1)	22.4 (22.4)	36.0 (33.5)
Primary City	0.2 (0.0)	6.6 (5.2)	23.1 (22.2)	22.1 (21.4)	48.1 (48.8)
Secondary City	3.0 (0.0)	10.0 (7.0)	28.2 (26.1)	24.5 (24.5)	34.3 (31.4)
Tertiary City	9.8 (1.5)	7.3 (3.6)	31.5 (30.0)	21.4 (22.1)	30.0 (26.3)

Uganda Data	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
National	4.7 (0.0)	4.6 (0.0)	59.4 (61.0)	18.6 (16.0)	12.7 (6.4)
Rural	5.6 (0.0)	4.5 (0.0)	63.8 (65.6)	16.6 (13.8)	9.5 (4.7)
Urban	1.6 (0.0)	5.0 (2.0)	43.7 (42.2)	25.6 (25.1)	24.1 (17.7)
Primary City	0.5 (0.0)	5.7 (2.6)	38.2 (39.9)	25.8 (26.1)	29.7 (23.0)
Secondary City	0.1 (0.0)	4.6 (2.6)	46.6 (47.3)	24.2 (23.0)	24.5 (17.5)
Tertiary City	2.2 (0.0)	4.8 (1.8)	45.5 (42.4)	25.7 (24.8)	21.9 (15.5)

Zambia Data	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
National	15.0 (7.1)	7.6 (0.0)	40.6 (36.7)	22.9 (20.9)	13.9 (10.5)
Rural	21.2 (16.5)	4.7 (0.0)	45.6 (43.1)	18.9 (15.2)	9.6 (6.9)
Urban	3.4 (0.0)	13.0 (10.7)	31.2 (29.4)	30.4 (29.9)	21.9 (20.3)
Primary City	0.6 (0.0)	14.0 (12.0)	30.2 (29.0)	30.9 (30.4)	24.3 (23.2)
Secondary City	3.0 (0.0)	14.3 (11.7)	30.5 (29.0)	30.4 (29.6)	21.9 (20.1)
Tertiary City	7.2 (0.0)	9.9 (7.5)	33.5 (31.1)	30.0 (29.4)	19.4 (17.3)

Source: authors' calculations using national household level surveys

Notes: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

Table 2-A-4: Average partial effects of the household determinants on processed share, urban household estimates by country

Urban	Processed Share			
	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.007 (0.004)	0.017 (0.000)	0.009 (0.001)	0.010 (0.000)
Household located in a primary city		0.068 (0.000)	0.000 (1.000)	0.037 (0.000)
Household located in a secondary city	0.053 (0.000)	0.047 (0.003)	-0.053 (0.049)	0.043 (0.000)
Nonfarm employment	0.008 (0.726)	0.094 (0.000)	0.015 (0.619)	0.037 (0.000)
Dependency ratio	-0.043 (0.091)	-0.075 (0.005)	-0.146 (0.001)	-0.001 (0.918)
Household adult equivalents	0.001 (0.684)	0.007 (0.000)	-0.006 (0.098)	0.006 (0.000)
Farmed hectares of land (log)	-0.098 (0.005)	-0.032 (0.010)	-0.056 (0.009)	-0.028 (0.000)
Own a gas or electric stove	0.035 (0.007)	-0.009 (0.702)		0.001 (0.902)
Own a refrigerator	0.021 (0.138)	-0.002 (0.851)		0.005 (0.404)
Maximum education attained within the household	0.007 (0.087)	0.000 (0.837)	0.001 (0.008)	0.000 (0.787)
Age of household head	-0.001 (0.139)	-0.001 (0.010)	-0.002 (0.014)	-0.001 (0.001)
Female head of household	-0.025 (0.066)	-0.008 (0.638)	-0.018 (0.381)	-0.008 (0.096)
Own a telephone	-0.009 (0.514)	0.073 (0.005)	0.067 (0.026)	0.029 (0.000)
Own a car	0.014 (0.518)	-0.033 (0.159)	-0.031 (0.329)	0.000 (0.954)
Own a motorcycle	-0.006 (0.856)	-0.008 (0.687)	-0.038 (0.207)	0.020 (0.292)
Own a bicycle	-0.065 (0.000)	-0.025 (0.102)	-0.100 (0.000)	-0.023 (0.000)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 2-A-5: Average partial effects of the household determinants on processed share, rural household estimates by country

Rural	Processed Share			
	Malawi	Tanzania	Uganda	Zambia
Household total expenditure per AE	0.004 (0.287)	0.009 (0.062)	0.022 (0.000)	0.010 (0.009)
Distance to primary city (log)				-0.045 (0.000)
Distance to secondary city (log)				-0.002 (0.554)
Distance to tertiary city (log)				-0.005 (0.001)
Nonfarm employment	0.080 (0.000)	0.117 (0.000)	0.107 (0.000)	0.183 (0.000)
Dependency ratio	-0.049 (0.004)	-0.075 (0.006)	-0.047 (0.071)	-0.032 (0.045)
Household adult equivalents	-0.009 (0.001)	-0.002 (0.073)	-0.003 (0.249)	-0.003 (0.073)
Farmed hectares of land (log)	-0.029 (0.040)	-0.044 (0.000)	-0.034 (0.000)	-0.023 (0.000)
Own a gas or electric stove	-0.029 (0.517)	0.000 (0.992)		0.039 (0.065)
Own a refrigerator	0.128 (0.003)	0.007 (0.794)		0.037 (0.117)
Maximum education attained within the household	0.014 (0.001)	0.001 (0.164)	0.000 (0.914)	0.007 (0.000)
Age of household head	-0.001 (0.008)	0.000 (0.794)	0.000 (0.539)	0.000 (0.039)
Female head of household	-0.008 (0.421)	0.018 (0.267)	-0.010 (0.491)	-0.004 (0.633)
Own a telephone	0.028 (0.002)	0.063 (0.000)	0.053 (0.000)	0.049 (0.000)
Distance to market (log)	-0.013 (0.000)	-0.004 (0.418)	-0.008 (0.095)	-0.008 (0.001)
Own a car	0.121 (0.008)	0.077 (0.201)	-0.057 (0.250)	0.118 (0.000)
Own a motorcycle	0.022 (0.514)	0.027 (0.421)	0.023 (0.237)	0.041 (0.447)
Own a bicycle	-0.021 (0.008)	-0.023 (0.086)	-0.034 (0.003)	0.009 (0.193)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 2-A-6: Average partial effects of the household determinants for processed food aggregates, urban household estimates by country

Urban Malawi	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.001 (0.680)	-0.004 (0.002)	-0.007 (0.000)	0.005 (0.001)	0.012 (0.000)
Household located in a secondary city	-0.034 (0.000)	0.037 (0.000)	-0.018 (0.031)	0.008 (0.287)	0.008 (0.249)
Nonfarm employment	0.017 (0.362)	-0.037 (0.026)	-0.020 (0.210)	0.022 (0.095)	0.022 (0.081)
Dependency ratio	0.013 (0.461)	0.019 (0.225)	0.034 (0.068)	-0.012 (0.464)	-0.046 (0.005)
Household adult equivalents	0.002 (0.212)	-0.004 (0.060)	-0.004 (0.034)	0.005 (0.013)	0.000 (0.983)
Farmed hectares of land (log)	0.044 (0.000)	-0.180 (0.001)	0.045 (0.089)	-0.008 (0.730)	-0.016 (0.330)
Own a gas or electric stove	-0.011 (0.228)	-0.025 (0.006)	-0.023 (0.028)	0.014 (0.175)	0.032 (0.004)
Own a refrigerator	-0.020 (0.023)	0.002 (0.840)	0.001 (0.918)	0.008 (0.459)	0.001 (0.906)
Maximum education attained within the household	-0.006 (0.045)	-0.005 (0.083)	-0.001 (0.760)	0.005 (0.056)	0.004 (0.091)
Age of household head	0.000 (0.088)	-0.001 (0.008)	0.000 (0.450)	0.000 (0.673)	0.000 (0.888)
Female head of household	0.005 (0.691)	-0.020 (0.066)	0.022 (0.086)	-0.004 (0.695)	-0.002 (0.832)
Own a telephone	0.006 (0.535)	-0.016 (0.173)	0.002 (0.839)	-0.007 (0.485)	0.020 (0.033)
Own a car	-0.021 (0.141)	-0.019 (0.160)	0.020 (0.230)	-0.006 (0.706)	0.026 (0.058)
Own a motorcycle	-0.032 (0.120)	0.005 (0.834)	0.043 (0.135)	-0.039 (0.059)	0.025 (0.279)
Own a bicycle	0.047 (0.000)	-0.039 (0.000)	0.019 (0.047)	-0.005 (0.542)	-0.019 (0.012)

Table 2-A-6: (continued)

Urban Tanzania	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.017 (0.000)	-0.010 (0.000)	-0.012 (0.000)	-0.002 (0.289)	0.023 (0.000)
Household located in a primary city	-0.057 (0.000)	0.025 (0.003)	-0.022 (0.080)	-0.009 (0.440)	0.051 (0.006)
Household located in a secondary city	-0.028 (0.001)	0.037 (0.001)	-0.011 (0.423)	0.007 (0.549)	0.006 (0.723)
Nonfarm employment	-0.053 (0.000)	-0.010 (0.420)	-0.028 (0.108)	0.004 (0.776)	0.109 (0.000)
Dependency ratio	-0.003 (0.869)	0.022 (0.149)	0.070 (0.002)	0.077 (0.000)	-0.176 (0.000)
Household adult equivalents	-0.004 (0.004)	-0.004 (0.000)	-0.004 (0.002)	0.003 (0.001)	0.008 (0.000)
Farmed hectares of land (log)	0.023 (0.000)	-0.028 (0.000)	0.002 (0.808)	-0.019 (0.035)	0.000 (0.988)
Own a gas or electric stove	0.015 (0.513)	-0.005 (0.688)	-0.007 (0.634)	-0.004 (0.770)	-0.001 (0.964)
Own a refrigerator	-0.004 (0.773)	-0.015 (0.090)	0.007 (0.546)	0.031 (0.013)	-0.018 (0.259)
Maximum education attained within the household	0.000 (0.378)	0.000 (0.930)	0.001 (0.376)	0.001 (0.124)	-0.001 (0.216)
Age of household head	0.000 (0.228)	0.000 (0.077)	0.001 (0.052)	-0.001 (0.053)	-0.001 (0.033)
Female head of household	-0.015 (0.112)	-0.006 (0.479)	0.024 (0.074)	0.014 (0.198)	-0.020 (0.187)
Own a telephone	-0.012 (0.300)	0.018 (0.075)	-0.036 (0.097)	0.060 (0.000)	0.008 (0.767)
Own a car	-0.027 (0.012)	0.022 (0.060)	0.038 (0.056)	0.008 (0.635)	-0.057 (0.028)
Own a motorcycle	-0.018 (0.079)	-0.006 (0.578)	0.028 (0.098)	0.039 (0.026)	-0.049 (0.056)
Own a bicycle	0.007 (0.405)	-0.020 (0.014)	0.017 (0.192)	0.008 (0.471)	-0.012 (0.446)

Table 2-A-6: (continued)

Urban Uganda	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.004 (0.075)	-0.006 (0.000)	-0.008 (0.003)	0.001 (0.726)	0.009 (0.000)
Household located in a primary city	-0.008 (0.056)	0.020 (0.024)	0.007 (0.714)	-0.017 (0.272)	0.002 (0.914)
Household located in a secondary city	-0.015 (0.000)	0.012 (0.292)	0.068 (0.012)	-0.023 (0.204)	-0.037 (0.128)
Nonfarm employment	-0.003 (0.669)	0.002 (0.849)	-0.009 (0.762)	-0.012 (0.619)	0.033 (0.287)
Dependency ratio	0.017 (0.112)	0.005 (0.797)	0.126 (0.004)	0.096 (0.009)	-0.221 (0.000)
Household adult equivalents	0.000 (0.758)	0.001 (0.270)	0.007 (0.080)	0.000 (0.897)	-0.007 (0.036)
Farmed hectares of land (log)	0.010 (0.014)	-0.002 (0.762)	0.038 (0.042)	-0.011 (0.514)	-0.053 (0.040)
Maximum education attained within the household	0.000 (0.397)	0.001 (0.005)	-0.001 (0.013)	0.002 (0.000)	-0.001 (0.264)
Age of household head	0.000 (0.062)	0.000 (0.630)	0.001 (0.053)	-0.001 (0.103)	-0.001 (0.378)
Female head of household	0.004 (0.481)	0.009 (0.266)	0.012 (0.545)	0.033 (0.045)	-0.060 (0.002)
Own a telephone	-0.001 (0.936)	-0.015 (0.167)	-0.067 (0.023)	0.033 (0.147)	0.060 (0.011)
Own a car	-0.002 (0.883)	-0.009 (0.260)	0.031 (0.341)	0.011 (0.681)	-0.013 (0.705)
Own a motorcycle	0.003 (0.720)	-0.006 (0.539)	0.037 (0.224)	0.002 (0.942)	-0.033 (0.101)
Own a bicycle	0.008 (0.242)	-0.021 (0.001)	0.093 (0.000)	-0.021 (0.241)	-0.058 (0.006)

Table 2-A-6: (continued)

Urban Zambia	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.007 (0.000)	-0.008 (0.000)	-0.006 (0.000)	0.009 (0.000)	0.008 (0.000)
Household located in a primary city	-0.033 (0.000)	0.050 (0.000)	-0.002 (0.713)	-0.025 (0.000)	0.015 (0.001)
Household located in a secondary city	-0.019 (0.000)	0.045 (0.000)	-0.017 (0.000)	-0.011 (0.018)	0.013 (0.001)
Nonfarm employment	-0.017 (0.000)	-0.007 (0.311)	-0.019 (0.010)	0.016 (0.053)	0.031 (0.000)
Dependency ratio	0.001 (0.824)	0.009 (0.202)	-0.002 (0.820)	0.015 (0.127)	-0.029 (0.001)
Household adult equivalents	-0.002 (0.002)	-0.001 (0.193)	-0.004 (0.000)	0.004 (0.000)	0.004 (0.000)
Farmed hectares of land (log)	0.011 (0.000)	-0.024 (0.000)	0.011 (0.000)	-0.004 (0.092)	-0.005 (0.013)
Own a gas or electric stove	-0.003 (0.235)	-0.004 (0.327)	0.000 (0.965)	0.004 (0.436)	-0.001 (0.905)
Own a refrigerator	-0.005 (0.061)	-0.002 (0.637)	-0.003 (0.631)	0.007 (0.210)	-0.002 (0.730)
Maximum education attained within the household	-0.001 (0.051)	-0.001 (0.114)	0.000 (0.772)	0.001 (0.042)	0.000 (0.629)
Age of household head	0.000 (0.002)	0.001 (0.001)	0.000 (0.025)	0.000 (0.015)	-0.001 (0.000)
Female head of household	0.001 (0.583)	0.002 (0.642)	0.008 (0.092)	0.003 (0.556)	-0.013 (0.002)
Own a telephone	-0.002 (0.501)	-0.005 (0.321)	-0.026 (0.000)	0.021 (0.001)	0.019 (0.000)
Own a car	0.001 (0.914)	-0.011 (0.034)	0.000 (0.992)	-0.005 (0.492)	0.009 (0.151)
Own a motorcycle	0.008 (0.377)	0.019 (0.328)	-0.024 (0.153)	-0.031 (0.104)	0.032 (0.210)
Own a bicycle	0.011 (0.000)	-0.013 (0.002)	0.011 (0.016)	-0.004 (0.409)	-0.006 (0.159)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

Table 2-A-7: Average partial effects of the household determinants for processed food aggregates, rural household estimates by country

Rural Malawi	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.025 (0.000)	-0.023 (0.000)	0.019 (0.000)	0.026 (0.000)	0.054 (0.000)
Nonfarm employment	-0.068 (0.000)	0.022 (0.121)	-0.032 (0.051)	0.022 (0.069)	0.031 (0.000)
Dependency ratio	0.045 (0.002)	-0.032 (0.009)	0.006 (0.660)	-0.007 (0.487)	-0.013 (0.025)
Household adult equivalents	0.003 (0.262)	-0.010 (0.000)	0.006 (0.002)	-0.002 (0.218)	0.003 (0.002)
Farmed hectares of land (log)	0.027 (0.015)	-0.045 (0.004)	0.002 (0.892)	-0.008 (0.360)	0.003 (0.569)
Own a gas or electric stove	0.032 (0.443)	-0.020 (0.487)	-0.002 (0.970)	-0.043 (0.173)	0.013 (0.274)
Own a refrigerator	-0.089 (0.007)	0.101 (0.035)	-0.076 (0.035)	0.046 (0.090)	0.010 (0.331)
Maximum education attained within the household	-0.002 (0.630)	-0.001 (0.671)	-0.013 (0.000)	0.008 (0.001)	0.005 (0.000)
Age of household head	0.000 (0.064)	0.000 (0.169)	0.000 (0.217)	0.000 (0.591)	0.000 (0.000)
Female head of household	0.006 (0.531)	0.006 (0.386)	0.002 (0.807)	-0.009 (0.119)	-0.009 (0.005)
Own a telephone	-0.007 (0.381)	-0.001 (0.891)	-0.020 (0.006)	0.010 (0.051)	0.014 (0.000)
Distance to market (log)	0.007 (0.029)	-0.009 (0.000)	0.006 (0.007)	-0.001 (0.672)	-0.002 (0.100)
Own a car	-0.022 (0.556)	0.014 (0.706)	-0.110 (0.020)	0.027 (0.386)	0.050 (0.002)
Own a motorcycle	0.007 (0.709)	-0.048 (0.039)	-0.023 (0.501)	0.040 (0.177)	-0.009 (0.430)
Own a bicycle	0.016 (0.022)	-0.026 (0.000)	0.006 (0.383)	0.004 (0.461)	0.002 (0.463)

Table 2-A-7: (continued)

Rural Tanzania	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.023 (0.000)	-0.015 (0.000)	0.009 (0.068)	-0.001 (0.823)	0.025 (0.000)
Nonfarm employment	-0.044 (0.003)	0.004 (0.690)	-0.080 (0.000)	0.005 (0.635)	0.104 (0.000)
Dependency ratio	-0.008 (0.719)	0.017 (0.222)	0.083 (0.002)	-0.004 (0.795)	-0.084 (0.000)
Household adult equivalents	-0.002 (0.118)	-0.002 (0.002)	0.004 (0.000)	-0.002 (0.002)	0.002 (0.028)
Farmed hectares of land (log)	0.044 (0.000)	-0.014 (0.006)	-0.003 (0.712)	-0.029 (0.000)	-0.006 (0.269)
Own a gas or electric stove	-0.007 (0.829)	-0.017 (0.307)	-0.005 (0.920)	0.034 (0.132)	-0.021 (0.307)
Own a refrigerator	-0.057 (0.020)	-0.019 (0.065)	0.014 (0.571)	0.045 (0.007)	-0.029 (0.034)
Maximum education attained within the household	0.000 (0.696)	0.000 (0.637)	-0.001 (0.069)	0.001 (0.004)	0.000 (0.765)
Age of household head	-0.001 (0.050)	0.000 (0.449)	0.000 (0.349)	0.000 (0.719)	0.000 (0.436)
Female head of household	-0.009 (0.421)	0.011 (0.243)	-0.009 (0.548)	0.015 (0.119)	-0.011 (0.263)
Own a telephone	-0.019 (0.056)	0.019 (0.013)	-0.046 (0.001)	0.028 (0.000)	0.018 (0.038)
Distance to market (log)	0.009 (0.022)	-0.005 (0.090)	-0.005 (0.342)	0.002 (0.534)	-0.001 (0.677)
Own a car	-0.047 (0.360)	0.064 (0.034)	-0.055 (0.090)	0.034 (0.149)	-0.002 (0.934)
Own a motorcycle	-0.022 (0.331)	-0.013 (0.194)	-0.004 (0.879)	0.033 (0.081)	0.003 (0.870)
Own a bicycle	0.015 (0.135)	-0.003 (0.644)	0.010 (0.405)	-0.004 (0.603)	-0.013 (0.107)

Rural Uganda	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.005 (0.062)	-0.004 (0.035)	-0.019 (0.000)	0.010 (0.000)	0.013 (0.000)
Nonfarm employment	-0.009 (0.333)	0.033 (0.000)	-0.100 (0.000)	0.025 (0.043)	0.044 (0.000)
Dependency ratio	0.038 (0.009)	0.021 (0.157)	0.013 (0.635)	0.024 (0.204)	-0.084 (0.000)
Household adult equivalents	0.000 (0.992)	-0.002 (0.107)	0.003 (0.119)	0.003 (0.232)	-0.003 (0.029)
Farmed hectares of land (log)	0.004 (0.393)	-0.005 (0.445)	0.031 (0.000)	-0.022 (0.001)	-0.006 (0.171)
Maximum education attained within the household	0.000 (0.379)	-0.001 (0.016)	0.000 (0.675)	0.001 (0.024)	0.000 (0.546)
Age of household head	0.000 (0.453)	0.000 (0.006)	0.000 (0.899)	0.000 (0.828)	0.000 (0.231)
Female head of household	-0.014 (0.019)	0.003 (0.711)	0.024 (0.092)	0.007 (0.494)	-0.020 (0.013)
Own a telephone	-0.010 (0.124)	0.019 (0.002)	-0.042 (0.001)	0.023 (0.014)	0.013 (0.054)
Distance to market (log)	-0.002 (0.351)	0.000 (0.975)	0.010 (0.036)	-0.002 (0.594)	-0.006 (0.055)
Own a car	0.002 (0.927)	-0.016 (0.240)	0.057 (0.274)	-0.015 (0.551)	-0.020 (0.421)
Own a motorcycle	-0.012 (0.169)	0.002 (0.854)	-0.010 (0.622)	0.014 (0.283)	0.010 (0.368)
Own a bicycle	0.009 (0.157)	-0.020 (0.000)	0.026 (0.030)	-0.001 (0.923)	-0.011 (0.111)

Table 2-A-7: (continued)

Rural Zambia	Maize		Non-Maize		
	Unprocessed	Processed	Unprocessed	Processed Low	Processed High
Household total expenditure per AE	-0.022 (0.000)	-0.007 (0.000)	0.010 (0.012)	0.010 (0.000)	0.007 (0.000)
Distance to primary city (log)	-0.022 (0.000)	-0.019 (0.000)	0.070 (0.000)	-0.005 (0.156)	-0.016 (0.000)
Distance to secondary city (log)	-0.012 (0.002)	0.006 (0.005)	0.013 (0.001)	0.004 (0.190)	-0.009 (0.000)
Distance to tertiary city (log)	0.002 (0.334)	-0.001 (0.211)	0.004 (0.037)	-0.002 (0.049)	-0.002 (0.008)
Nonfarm employment	-0.114 (0.000)	0.037 (0.000)	-0.111 (0.000)	0.064 (0.000)	0.057 (0.000)
Dependency ratio	0.036 (0.033)	-0.017 (0.022)	-0.005 (0.788)	-0.003 (0.823)	-0.012 (0.066)
Household adult equivalents	-0.001 (0.567)	-0.002 (0.006)	0.004 (0.021)	0.000 (0.910)	0.000 (0.653)
Farmed hectares of land (log)	0.024 (0.000)	-0.013 (0.000)	0.002 (0.590)	-0.006 (0.011)	-0.001 (0.517)
Own a gas or electric stove	-0.110 (0.000)	0.003 (0.778)	0.029 (0.215)	0.037 (0.023)	-0.016 (0.079)
Own a refrigerator	0.020 (0.509)	0.003 (0.846)	-0.061 (0.011)	-0.014 (0.399)	0.041 (0.009)
Maximum education attained within the household	-0.003 (0.012)	0.001 (0.254)	-0.005 (0.000)	0.004 (0.000)	0.003 (0.000)
Age of household head	0.000 (0.150)	0.000 (0.009)	0.001 (0.001)	0.000 (0.054)	0.000 (0.000)
Female head of household	0.018 (0.052)	0.006 (0.266)	-0.013 (0.146)	0.001 (0.900)	-0.010 (0.005)
Own a telephone	0.001 (0.953)	0.011 (0.037)	-0.052 (0.000)	0.026 (0.000)	0.012 (0.001)
Distance to market (log)	0.001 (0.679)	-0.004 (0.004)	0.007 (0.012)	0.001 (0.725)	-0.005 (0.000)
Own a car	-0.096 (0.000)	0.009 (0.547)	-0.048 (0.096)	0.064 (0.004)	0.030 (0.058)
Own a motorcycle	-0.018 (0.682)	0.024 (0.258)	-0.027 (0.487)	0.019 (0.635)	-0.003 (0.767)
Own a bicycle	-0.024 (0.002)	-0.006 (0.100)	0.016 (0.041)	0.007 (0.210)	0.010 (0.002)

Source: Authors' calculations

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value.

REFERENCES

REFERENCES

- Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *Review of Economics and statistics*, 79(4), 527-539.
- Battersby, J., & Peyton, S. (2014, June). The geography of supermarkets in Cape Town: Supermarket expansion and food access. In *Urban Forum* (Vol. 25, No. 2, pp. 153-164). Springer Netherlands.
- Bennett, M. K. (1941). Wheat in national diets. *Wheat Studies*, (02).
- Berdegúe, J. A., Carriazo, F., Jara, B., Modrego, F., & Soloaga, I. (2015). Cities, Territories, and Inclusive Growth: Unraveling Urban–Rural Linkages in Chile, Colombia, and Mexico.
- Christiaensen, L., Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: the role of rural diversification and secondary towns¹. *Agricultural Economics*, 44(4-5), 435-447.
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368), 829-836.
- Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596-610.
- De Haen, H., Stamoulis, K., Shetty, P., & Pingali, P. (2003). The world food economy in the twenty-first century: challenges for international co-operation. *Development Policy Review*, 21(5-6), 683-696.
- D’Haese, M., & Van Huylenbroeck, G. (2005). The rise of supermarkets and changing expenditure patterns of poor rural households case study in the Transkei area, South Africa. *Food Policy*, 30(1), 97-113.
- Fafchamps, M., & Shilpi, F. (2003). The spatial division of labour in Nepal. *The Journal of Development Studies*, 39(6), 23-66.
- Ferré, C., Ferreira, F. H., & Lanjouw, P. (2012). Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, lhs007.
- Goldman, A. (2000). Supermarkets in China: the case of Shanghai. *The international review of retail, distribution and consumer research*, 10(1), 1-21.
- Gómez, M. I., Barrett, C. B., Raney, T., Pinstrip-Andersen, P., Meerman, J., Croppenstedt, A., Carishma, B., & Thompson, B. (2013). Post-green revolution food systems and the triple burden of malnutrition. *Food Policy*, 42, 129-138.

- Gómez, M. I., & Ricketts, K. D. (2013). Food value chain transformations in developing countries: Selected hypotheses on nutritional implications. *Food Policy*, 42, 139-150.
- Haggblade, S., Hazell, P., & Reardon, T. (2010). The rural non-farm economy: Prospects for growth and poverty reduction. *World Development*, 38(10), 1429-1441.
- Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate behavioral research*, 45(4), 627-660.
- Hino, H. (2014). Shopping at different food retail formats: Understanding cross-shopping behavior through retail format selective use patterns. *European Journal of Marketing*, 48(3/4), 674-698.
- Huang, J., & Bouis, H. (2001). Structural changes in the demand for food in Asia: empirical evidence from Taiwan. *Agricultural Economics*, 26(1), 57-69.
- Leser, C. E. V. (1963). Forms of Engel functions. *Econometrica: Journal of the Econometric Society*, 694-703.
- Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The manchester school*, 22(2), 139-191.
- Monteiro, C. A., Levy, R. B., Claro, R. M., Castro, I. R. R. D., & Cannon, G. (2010). A new classification of foods based on the extent and purpose of their processing. *Cadernos de saude publica*, 26(11), 2039-2049.
- Monteiro, C. A., Moubarac, J. C., Cannon, G., Ng, S. W., & Popkin, B. (2013). Ultra-processed products are becoming dominant in the global food system. *Obesity reviews*, 14(S2), 21-28.
- Nelson, G. C. (2002). Introduction to the special issue on spatial analysis for agricultural economists. *Agricultural Economics*, 27(3), 197-200.
- Neven, D., Reardon, T., Chege, J., & Wang, H. (2006). Supermarkets and consumers in Africa: the case of Nairobi, Kenya. *Journal of International Food & Agribusiness Marketing*, 18(1-2), 103-123.
- 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*, volume 11, 619-632.
- Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1), 121-133.
- Pingali, P. (2007). Westernization of Asian diets and the transformation of food systems: implications for research and policy. *Food policy*, 32(3), 281-298.

- Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition reviews*, 70(1), 3-21.
- Reardon, T., & Timmer, C. P. (2007). Transformation of markets for agricultural output in developing countries since 1950: How has thinking changed?. *Handbook of agricultural economics*, 3, 2807-2855.
- Rischke, R., Kimenju, S. C., 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(4), 920-927.
- Serber, G. A. F., & Wild, C. J. (1989). *Nonlinear Regression*. John Wiley & Sons, 435.
- Sharma, A. (2016). Urban Proximity and Spatial Pattern of Land Use and Development in Rural India. *The Journal of Development Studies*, 1-19.
- Smith, L. C. (Ed.). (2003). *The importance of women's status for child nutrition in developing countries* (Vol. 131). Intl Food Policy Res Inst.
- Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. *The Economic Journal*, 64(255), 511-527.
- Strydom, J. W. (2011). Retailing in disadvantaged communities: the outshopping phenomenon revisited. *Journal of Contemporary Management*, 8, 150-172.
- Subramanian, S., & Deaton, A. (1996). The demand for food and calories. *Journal of political economy*, 104(1), 133-162.
- Thiele, S., & Weiss, C. (2003). Consumer demand for food diversity: evidence for Germany. *Food Policy*, 28(2), 99-115.
- Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015). The rise of a middle class in East and Southern Africa: Implications for food system transformation. *Journal of International Development*, 27(5), 628-646.
- Turrell, G., & Kavanagh, A. M. (2006). Socio-economic pathways to diet: modelling the association between socio-economic position and food purchasing behaviour. *Public health nutrition*, 9(03), 375-383.
- Veeck, A., & Veeck, G. (2000). Consumer segmentation and changing food purchase patterns in Nanjing, PRC. *World Development*, 28(3), 457-471.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Working, H. (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, 38(221), 43-56.

World Bank. (2015, October 4). The international poverty line has just been raised to \$1.90 a day, but global poverty is basically unchanged. How is that even possible? [Blog post]. Retrieved from <http://blogs.worldbank.org/developmenttalk/international-poverty-line-has-just-been-raised-190-day-global-poverty-basically-unchanged-how-even>

ESSAY 3: CITY SIZE, SUPERMARKETS, AND PROCESSED FOODS: EVIDENCE FROM ZAMBIA

3.1 Introduction

The past two decades in developing Sub-Saharan Africa can be categorized as a time of significant lifestyle change driven by increasing incomes and rapid urbanization. Two of these lifestyle changes are the increased utilization of supermarkets and the increased consumption of processed food.

Retail modernization in developing countries dates back to the 1960s, when supermarkets in Latin America were gaining popularity in response to growth of household incomes (Moyer & Hollander, 1968; Harrison et al., 1974). However, it was not until after market liberalization, which led to the influx of foreign direct investment (FDI), that supermarkets' share of retail food in this part of the world rose rapidly during the 1990s (Reardon et al., 2004). Diffusion of supermarkets in developing regions is the product of the following supply and demand conditions (Reardon & Timmer, 2007): On the supply side; market liberalization facilitates the inflow of FDI, complementary infrastructure provides needed logistics, and large-scale food manufacturers provide food products to be sold via supermarkets. On the demand side; the combination of rising incomes and urbanization contribute to market size that can support the presence of supermarkets and an increase in household opportunity cost of time that strengthen the benefits these supermarkets offer. These conditions were present in Latin America in the 1990s and are becoming ever more present in developing Africa. In the 1990s there was little supermarket presence in Sub-Saharan Africa outside of South Africa, but currently the market

share of supermarkets in developing Africa, albeit small, is growing⁴⁰ (Neven & Reardon, 2004; Traill, 2006).

Consumer access to supermarkets provides at least three potential benefits. First, supermarkets offer households time-savings with the ability to purchase a variety of foods from a single location. Second, the cleanliness of supermarkets compared with traditional open-air markets is a benefit for households with higher health awareness that accompanies increased education (Turrell & Kavanagh, 2006). Third, supermarkets provide households with cost savings as they typically offer goods such as processed foods at cheaper prices than other retailers due to supermarkets' ability to procure goods in bulk (Neven et al., 2006; Minten et al., 2010).

Supermarkets offering processed foods at cheaper prices than other outlets could contribute to the increased consumption of processed food that is primarily attributed to economic development, characterized by rising incomes and rapid urbanization (Veeck & Veeck, 2000; Pingali, 2007; Popkin et al., 2012; Tschirley et al., 2015). The increased consumption of processed food that occurs with rising income and urbanization is similar to the relationship shown by Bennett's Law, which identifies income as a driver of changing patterns of commodity based food consumption. Increasing opportunity cost of time accentuates this pattern as it is related to income growth and has also been found to increase consumption of processed food, particularly when the opportunity cost of time is driven by the increased participation of females in nonfarm employment (Senauer et al., 1986; Kennedy & Reardon, 1994).

The concept of "urbanization" embodies multiple characteristics, each of which could have differing impacts on the pattern of processed food consumption. Urbanization increases

⁴⁰ Neven & Reardon (2004) estimate that supermarket share of food retail in urban Kenya was 20%. Traill (2006) projected that in 2015 supermarkets would capture 15% of Kenya's national food retail market, up from 10% in 2002.

population density, providing retailers with a concentrated demand for food that incentivizes the construction of supermarkets in urban areas (Battersby & Peyton, 2014). The rise of formal retail outlets in urban areas introduces a greater variety of foods and promotes the consumption of processed foods (Gómez & Ricketts, 2013). Urban areas also offer additional employment- or entrepreneurial opportunities and a generally faster pace of life (Bettencourt et al., 2007), both of which increase household opportunity cost of time, further affecting the demand for processed food.

Urbanization is the growth of the share of population in urban areas relative to rural areas. This growth occurs via migration from rural to urban areas and the expansion of urban areas. Recently the poverty and migration literature has begun to take a more disaggregated view of urbanization and explore the impacts of city size (determined by population) on income and poverty. Households in larger cities generally have higher incomes than households in smaller cities, while smaller cities provide greater opportunity for migrating rural households to escape from poverty (Ferré et al., 2012; Christiaensen et al., 2013; Berdegue et al., 2015). As city size affects income and the escape from poverty, it is reasonable to expect that city size also affects food consumption decisions of households. When examining consumer demand for food variety in Germany, Theile & Weiss (2003) find that households living in larger cities demand greater variety of food. This paper further considers the effects of city size on food consumption patterns.

Jointly considering increasing supermarket utilization, processed food consumption and the impact of city size exposes a gap in the literature. Literature exists regarding supermarket adoption, processed food consumption, and, more recently, the impact of supermarket adoption on processed food consumption (Asfaw, 2008; Rischke et al., 2015). Previous studies have also

examined how food consumption patterns vary across rural and urban areas (Popkin, 1999; Ecker & Qaim, 2011; Tschirley et al., 2015). To my knowledge, no previous study has unpacked the impact of urbanization on the food households in developing countries consume and from where they acquire it by examining how this impact varies by city size.

This paper will contribute to the literature by examining the effects of city size on supermarket use and processed food consumption when controlling for the effects of income and other household variables. It will do this using a unique household dataset covering four cities in Zambia. The direct and indirect effects of city size on these consumption decisions are tested using previously applied empirical models that have been adapted to this study.

The paper proceeds as follows: Section two presents the conceptual model used to articulate the decision process and hypotheses. Section three addresses the methodology, which includes description of data, definitions, general and empirical models, and estimation procedures. Section four reflects on the results and provides discussion. Lastly, section five provides conclusions.

3.2 Conceptual Framework

Rischke et al. (2015) use the concept of the food environment, describing the supply side facing consumers, combined with demand side factors, to motivate a discussion and generate hypotheses about where people source their food, what and how much they source, and how they eat. They do not consider city size. I adapt their framework to generate hypotheses regarding the effect of city size on the likelihood of consumers shopping in supermarkets and of purchasing processed foods when they do so.

Figure 3-1 summarizes my conceptual approach. City size affects characteristics of the urban environment that are exogenous to consumers' food choices; this urban environment influences the food environment that consumers face (the supply side) along with the factors that drive consumer demand. Together, the food environment and consumer demand factors influence the food choices that consumers make: where to shop (supermarkets or other outlet types) and what to buy (processed foods or others). The food consumed by households can be influenced directly by the food environment and consumer demand factors, and indirectly through the impact of these factors on supermarket shopping.

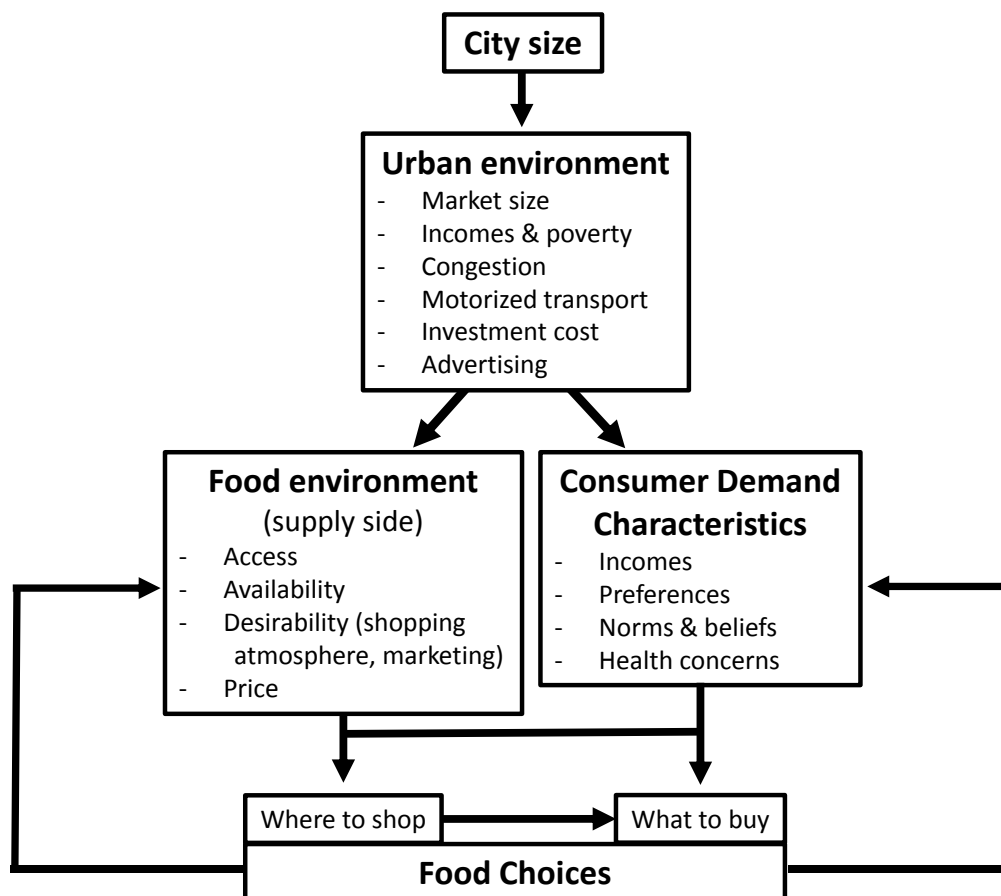


Figure 3-1: Conceptual approach linking city size to consumer food choices

The concept of food environment applies a multi-dimensional lens to the factors that describe the supply side of retail food systems. Given that supermarkets in ESA compete with a wide range of other outlet types (large and small supermarkets, traditional shops, open air markets, and often thousands of dispersed street vendors), the factors need to be thought of with respect to supermarkets per se, and to supermarkets relative to other outlet types. In Rischke et al. (2015), food environment considers access, availability, desirability, and prices. Access refers to the cost of gaining physical access to foods, such as the density of outlets, and the time and money cost of reaching them. This concept is not to be confused with that of economic access. Availability identifies the types and variety of foods (including processed foods) that supermarkets and other outlets offer. Desirability recognizes that the food environment affects household desire for products beyond household desire for product characteristics via shopping atmosphere and marketing. Prices vary across products and across retailers (Gómez & Ricketts 2013). Shopping atmosphere is created by the characteristics of the physical shopping space (size, design, internal arrangement), and by promotions and product placements. In urban ESA, supermarkets are unique in the extent to which they purposively invest in shopping atmosphere. Marketing affects how product characteristics relate to credence attributes created through branding, labeling, and advertising (e.g. health claims). Note that these marketing investments are always made first in packaged processed foods.

The food environment households experience is both a result of the retail food system available to households and the retailers that households select to use. The variation between retailers regarding availability, desirability and prices of food are accentuated when households select to shop at specific retailers – such as, the desirability of certain foods being highlighted by the promotions and product placements that are only experienced when shopping at a particular

retailer. Therefore, the decision of where to buy contributes to household decisions of what to buy.

The demand side in this approach considers, in addition to consumer incomes, their individual and household preferences (e.g. for taste, energy density, and convenience); their social and individual norms and beliefs regarding the shopping environment, beauty and body shapes, and lifestyle; and their health concerns including foods' nutritional content and perceived safety.

It is clear factors that influence the created food environment (the supply side), such as branding, labeling, advertising, packaging, and store layout, also affect the demand side through their impact on the credence that consumers lend to overt and implicit claims about products.

As consumers make their food choices, learning drives feedback loops to both the supply side and the demand side: retailers learn by observing consumer decisions, and adjust their approaches to stocking, marketing, and store layout (and over a longer time horizon, they affect access through decisions on store location), and consumers learn about products and change their assessment of their own tastes and preferences over time.

In rough order of certainty, I posit how city size affects aspects of the exogenous urban environment that drives both the supply side (the food environment) and consumer demand factors. I suggest the following effects as city size rises:

1. The size of the market increases more than proportionally to city size: this follows directly from a positive relationship between city size and mean consumer incomes (see next point). Increased size of market attracts investment in supermarkets, improves physical access to them, and thus favors shopping in them. Increased market size should also promote the

availability of a wider array of foods, including processed foods; all else equal, this should spur the consumption of processed foods.

2. Mean incomes rise and poverty rates fall (Ferré et al., 2012). These are robust empirical patterns across Africa (World Bank, 2009; Christiaensen et al., 2013). Both should increase shopping in supermarkets and also have direct positive impacts on the purchase of processed foods, since these foods tend to have higher demand elasticities than do unprocessed foods (Tschirley et al., 2015).
3. Congestion rises. This means that the cost of moving a given distance by motorized transport rises. However, the cost of moving by foot or bicycle – both common forms of transport that are negatively associated with household incomes and with shopping in supermarkets – is probably insensitive to city size. Congestion thus asymmetrically affects physical access to food, reducing the access advantage conveyed by motorized transport thus reducing the probability of shopping in a supermarket in larger cities. Alternatively, due to the increased cost of transportation, the costs of cross-shopping⁴¹ increases and the supermarket benefit of offering a wide variety of food items in one location is of greater value to households, thus increasing the probability of shopping in a supermarket in a larger city. Since congestion increases the time cost of a given activity and thus increases the value of time, it should directly drive demand for processed food and the convenience it offers.
4. Access to motorized transport – both public and private – increases. This effect follows from higher consumer incomes and absolute amount of purchasing power in large cities (the income elasticity of demand for private motorized transport is certainly positive, and public

⁴¹ The practice of cross-shopping is where households shop at various locations. Cross-shopping is common in China (Goldman 2000), Israel (Hino 2014) and in South Africa, where households travel to locations away from their area of residence to acquire food (D'Haese & Van Huylenbroeck 2005, Strydom 2011).

transport availability should respond positively to total purchasing power). Access to motorized transport has an asymmetric effect on physical access to food: motorized transport improves physical access for outlets such as supermarkets that are larger and fewer in numbers, but should have little or no effect on access to other outlet types, which have developed historically to serve largely foot-bound consumers. Access to transport thus trends in favor of supermarket shopping in large cities, in part because a key advantage of shopping in such outlets is the ability to buy in bulk. Its effect on consumption of processed foods, independent of shopping in supermarkets, is not clear.

5. The cost of infrastructural investment rises. This follows from higher congestion, higher wages, and (if market mechanisms operate in any way in the land market) higher costs of land. Higher investment cost probably works asymmetrically against investment in supermarkets as city size rises: open air markets often arise spontaneously (informally) with no explicit cost of land, and street vendors do not typically pay for the space they occupy. These costs also raise food prices, reducing real incomes of consumers; any effect on processed foods would be only through the supermarket effect.
6. Consumers are more exposed to modern advertising. This is a straightforward supply side response of retailers and food processing companies to the larger market size in large cities, combined with economies of scale (more eyes see a given billboard; tv and radio advertising reach more people; the marginal return to advertising should be higher because incomes are higher). By design, such advertising increases the likelihood that consumers will pursue a more “modern” shopping experience, thus favoring use of supermarkets and (even independent of such shopping) consumption of processed foods.

Table 3-1 summarizes my expectations regarding the impact of city size on food choices, working through the urban environment, the food environment, and consumer demand factors. Of the six urban environment factors, four drive a positive relationship between city size and the use of supermarkets: market size, mean incomes, access to motorized transport, and exposure to advertising. The cost of infrastructural investment drives a negative relationship and congestion has mixed effects. Four factors also drive a positive relationship between city size and processed food demand: market size, mean incomes, congestion, and advertising. Notably, I don't expect negative direct effects on processed food demand from any of the six factors.

Table 3-1: Impact of city size on food choices: urban environment, food environment, and demand factors

Change in exogenous urban environment factor as city size rises	Impact on supply (food environment) and demand factors		Impact on food choices	
	Food environment	Demand factors	Choice of supermarkets	Processed food demand
Disproportionate increase in market size	Increased access to supermarkets; increased availability of processed (and other) foods		Positive	Positive
Increased incomes, lower poverty		Higher incomes; operation of Bennet's Law; rising health concerns; more "modern" norms & beliefs	Positive	Positive
Higher congestion	Asymmetric impact on physical access – reduces advantage conveyed by motorized transport while also increasing the benefit of being able to purchase all food from one location	Higher opportunity cost of time	Positive or Negative	Positive
Greater access to motorized transport	Asymmetric impact on physical access – increases it for supermarkets but little or no effect for other outlet types		Positive	Only through supermarkets
Higher investment cost	Reduced physical access to supermarkets		Negative	Only through supermarkets
Greater exposure to advertising	Increased desirability of food, especially of packaged, processed foods	More "modern" norms & beliefs around food	Positive	Positive

This framework thus predicts a strongly positive effect of city size on processed food consumption: all direct factors are positive; any negative effects are indirect through a negative effect on use of supermarkets and a hypothesized positive effect of supermarkets on processed food demand; and the possible indirect negative effects of congestion are offset by the positive direct effects.

The net effect of city size on use of supermarkets depends on the size effect of each factor, which varies on a host of factors specific to any given locale. Note first that there is clearly a threshold effect: supermarkets do not exist in rural areas or small towns in Africa; urban settlements have to reach a certain size before they will attract a supermarket. Once that threshold is reached, however, I see no way, a priori, to predict the relationship between city size and supermarket use⁴². Notably, the framework does suggest the possibility of city size negatively affecting the use of supermarkets, depending on the impact on investment costs and household response to the level and cost of congestion.

3.3 Methods

3.3.1 Data

To estimate the effects of city size on supermarket use and processed food consumption I use data from the 2007/2008 Zambia Urban Consumption Survey (UCS), which were gathered via the joint efforts of the Zambian Central Statistics Office and the Zambia Food Security Research Project. These data include a sample of 2,160 urban households across four cities: Lusaka, Kitwe, Kasama and Mansa (Table 3-2). To account for potential seasonal variation in consumption, data were collected in two rounds: August 2007 and February 2008. August is soon after primary harvest while February is traditionally the height of the lean season. During the second round of data collection, 1,865 households were successfully re-interviewed. The standard enumeration areas were selected in such a way to collect data from households that represent the neighborhoods across multiple socioeconomic strata. See Hichaambwa et al. (2009) for more sampling details of this survey.

⁴² Note that all the cities in this sample have a supermarket.

Table 3-2: City populations

City	Census Year		Annual Growth %
	2000	2010	
Lusaka	1,084,703	1,747,152	4.88%
Kitwe	363,734	501,360	3.26%
Kasama	74,243	101,845	3.21%
Mansa	41,059	78,153	6.65%
Zambia National	10,100,981	13,216,985	2.73%
Zambia Urban	3,515,393	5,118,243	3.83%
Urban Share	34.8%	38.7%	1.07%

Sources: city data obtained from <http://www.citypopulation.de/Zambia-Cities.html>, national data obtained from the United Nations Population Division, Department of Economic and Social Affairs, 2014 Revision of the World Urbanization Prospects.

Lusaka, Kitwe, Kasama and Mansa were selected to represent multiple city types in Zambia. Lusaka is the nation's capital and is three times larger than the next largest city, Kitwe. Kitwe is the largest city in the Copperbelt, where the economic activity is driven by copper mining. Kasama and Mansa are smaller, ranking as the tenth and twelfth largest cities in Zambia per 2010 census data⁴³. Like Lusaka and Kitwe, Kasama and Mansa are district capitals, but these cities have a higher share of informal employment than their larger counterparts. These four cities also vary in per adult equivalent average expenditure⁴⁴, with average Lusaka expenditures about double those in the tertiary cities (Table 3-3). I recognize that four cities comprise a limited set of cities to formulate strong assertions regarding patterns driven by city size, but given the variation in the city populations I feel that this analysis will provide a basis for understanding the effects of city size on food consumption patterns.

⁴³ <http://www.citypopulation.de/Zambia-Cities.html>

⁴⁴ Expenditure data were converted into USD using historic exchange rates obtained from XE.com, and conversion information for constant 2011 international dollars in purchasing power parity from worldbank.org. Data includes values for food from own production, gifts of food and imputed rent. Adult equivalents (AE) are calculated as one AE for each household member fifteen years of age or older, three quarters of an AE for each child aged five to fourteen, and a half an AE for each child younger than five years old.

Table 3-3: Total daily household expenditure per adult equivalent

	All	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Lusaka	8.94	2.69	4.86	7.30	11.14	25.63
Kitwe	6.98	2.06	3.64	5.19	7.93	20.45
Kasama	4.29	1.28	2.17	3.18	4.58	11.51
Mansa	4.51	1.22	2.29	3.53	5.40	12.03

Source: Zambia Urban Consumption Survey 2007/2008

Note: Quintiles represent total expenditure quintiles, by city.

Values presented in constant 2011 USD, PPP adjusted.

Zambia is part of the final wave of the “supermarket revolution”, with a growing presence of supermarkets across the country (Reardon et al. 2009). Each of these cities has at least one large chain supermarket and smaller non-chain supermarkets. Data collected in this survey include distance from the household to each of the retailers from which it acquired food (Table 3-4). Average distances to supermarkets vary across cities and expenditure quintiles. Supermarkets are located closer to wealthy neighborhoods than to less affluent neighborhoods, consistent with patterns observed in South Africa (Battersby & Peyton, 2014).

Table 3-4: Kilometers to supermarket relative to closest non-supermarket food retailer to the household

	All	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Lusaka	3.2	3.5	3.4	3.7	3.1	1.8
Kitwe	7.1	8.4	7.6	7.2	6.8	4.3
Kasama	3.3	4.6	3.9	3.4	2.5	1.8
Mansa	1.8	2.5	2.0	1.9	1.2	1.0

Source: Zambia Urban Consumption Survey 2007/2008

Data collected by this survey include household demographics (health, income and education), asset ownership, general expenditures from the previous six months and detailed information on food consumption. The food consumption module collected information on 68 commonly consumed food items from the last 30 days that were either purchased, received as gifts, or from own production. Data on food consumed away from home were collected in a separate module using 24 hour recall. Table 3-5 identifies the share of market value by

acquisition type. Data regarding the source and distance to household were recorded for all purchased food items. Household level prices of food purchases were not able to be computed due to non-standard units.

Table 3-5: Shares of market value by acquisition type

	Own Production	Gifts	Supermarket Purchases	Non-Supermarket Purchases	Food Away From Home
Weighted Average (all cities)	2.5	1.9	7.6	73.4	14.7
Lusaka	1.6	1.8	7.9	72.7	16.0
Kitwe	4.0	1.6	5.8	77.7	10.8
Kasama & Mansa Average	9.7	3.8	9.6	66.7	10.2
Kasama	10.4	4.4	8.7	67.7	8.7
Mansa	8.2	2.4	11.4	64.7	13.3

Source: Zambia Urban Consumption Survey 2007/2008

Note: Food away from home is not considered to be non-supermarket purchase, but is segregated out in this table to represent its market share.

3.3.2 Definitions

Key definitions regarding supermarkets, processed food and city size are specified here. First, *supermarket* is defined as a self-service retail outlet with a variety of goods that is typically characterized by multiple cashiers. The definition used for supermarkets both large and small supermarkets that are both independent or chains. *Supermarket share* refers to the household of total household food consumption that was obtained in a supermarket. Second, *processed food* is defined as food that prior to household acquisition has been transformed from its original state, beyond removal from the plant and (for non-perishables) drying. Processed food therefore identifies food embedded with the labor provided by humans or machines prior to acquisition, reducing household time and energy required for preparation. *Processed share* refers to the household share of processed food in total household food consumption. The processed food aggregate is further observed using sub-aggregates of low- and high processed foods. High processed foods are processed foods that satisfy at least two of the three conditions: multiple ingredients; physical change induced by heating, freezing, extrusion, or chemical processes; and

packaging more complex than simple paper or plastic. Although the processed sub-aggregates are constructed based on implications on the structure of the food system they closely mirror the processed food aggregates considered nutrition based literature⁴⁵ (Monteiro et al., 2010). Third, *city size* classifies cities by population: primary cities have populations greater than one million (Lusaka), secondary cities have populations below one million and above than one hundred thousand (Kitwe), and tertiary cities have populations below one hundred thousand (Kasama and Mansa).

3.3.3 General Model

Our model estimates how consumption decisions are affected by household variables, including variables that proxy for household income. An Engel's Curve Model is an appropriate general model to use given that it estimates the relationship between shares of consumption against household income, allowing for estimation of the effects of additional determinants of consumption such as city size. The functional form of Engel Curves proposed by Banks et al. (1997) is an appropriate form for this analysis as this functional form allows for the effect of total expenditure on the shares of food aggregates to vary at different levels of expenditure.

3.3.4 Empirical Model

Our conceptual model posits that city size influences the urban environment, which affects both the supply side (the food environment) and the demand side of the consumer's food consumption decisions. In empirically specifying the model, I search for variables that capture as many of the supply- and demand side factors as possible. I also include dummy variables for

⁴⁵ Figure 3-A-1 shows a sample of the food items included in the Zambia data and how they map into both the Monteiro et al. 2010 aggregates and the processing sub-aggregates used here.

each city which, due to the large size differences, can be argued to primarily reflect city size. To the extent that variables are missing, or imperfectly capture the concepts in the conceptual model, I use the conceptual model to interpret the meaning of the city size variables and generate defensible conclusions.

The dataset captures consumer demand characteristics relatively well (Figure 3-1), with variables for household per capita expenditure along with demographic variables that proxy for preferences, norms & beliefs, and health concerns.

Data are less complete on the food environment. The most representative variable that I have regarding food environment is *relative distance to supermarket*, which I compute as the distance from household to nearest supermarket minus the distance from household to nearest non-supermarket food retailer. This variable relates to physical access yet does not fully capture it. In particular, because it is based only on distance it is unable to capture the effects of congestion in the urban environment. Logically, greater congestion should increase the real costs related to distance by increasing the time needed to travel a given distance when using motorized transport. I thus expect that relative distance to supermarket will have a larger (negative) effect on supermarket share in the larger cities (which I expect to be more congested).

Measures of availability and food desirability are not available. Systematic differences across cities will be captured by city dummy variables in pooled regressions. One such systematic difference is how city size affects the desirability of food items via heterogeneous prices. The dataset does not capture household level prices due to the high prevalence of non-standard units. I do know that general price levels are higher in larger cities than in smaller cities, as indicated by a price index⁴⁶ calculated from select community level prices captured within the

⁴⁶ Expenditure weighted price index comprised of the 10 most consumed items during each round of data collection: commercial maize meal, bread, chicken, beef, cooking oil, rape, tomato, and sugar in both rounds with rice and dried fish in round 1 and

data (Table 3-A-1). However, certain processed foods (i.e. bread and commercial maize meal) are found to have lower average prices in primary cities. The lower prices of processed foods are expected to contribute to greater consumption of processed foods by households in primary cities, an effect expected to be captured by the city dummy variables.

Examining the six urban environment factors from Table 3-1 of the conceptual model, I note that I am able to directly control for two of them: income and access to motorized transport. Both are hypothesized to have positive effects on use of supermarkets. I am left with four factors that need to be captured by city size dummies – market size, congestion, investment cost, and advertising. Two of these, congestion and investment cost, are hypothesized to have negative effects on the use of supermarkets. It would thus not be surprising if I find that increasing city size decreases the use of supermarkets.

The effects of city size and the other selected variables on supermarket share and on processed share are estimated by the following Engel's Curve Model:

$$D_{Dependent\ Variable} = f(Ac^*, Av^*, D^*, P^*, Y^*, Z^*, W^*, H^*) \quad (3.1)$$

The dependent variables are supermarket share and processed share (Table 3-6). Processed share is further disaggregated into shares of low processed and high processed food. Foods are high processed if they satisfy at least two of the following three conditions: multiple ingredients; physical change induced by heating, freezing, extrusion, or chemical processes (i.e., more than simple physical transformation); and packaging more complex than simple paper or plastic.

kapenta and beans in round 2. Round 1 Index: Lusaka = 100, Kitwe = 108.7, Kasama = 92.0, Mansa = 99.1. Round 2 Index: Lusaka = 100, Kitwe = 89.5, Kasama = 76.1, Mansa = 81.5. Overall mean by city Lusaka = 100, Kitwe = 100.9, Kasama = 85.6, Mansa = 94.1..

Table 3-6: Dependent variables

Supermarket Shares	
Food acquired from supermarkets such as Shoprite or Melissa	
Processed Food Shares	
Any Processed Food	Any food that is acquired in a processed form
Own Production	Food that was produced by the household
Unprocessed	Food acquired as unprocessed, e.g. maize grain
Low Processed	Food acquired as low processed, e.g. maize flour
High Processed	Food acquired as high processed, e.g. bread

Independent variables that are identified in Table 3-7 are categorized into three groups: variables that are only determinants for supermarket share, variables that are only determinants of processed share, and variables that are shared determinants for both supermarket share and processed share.

Table 3-7: Key for determinants of demand in equation 3.1

Ac*	Access	Relative distance to supermarket, Own a car, Own a motorcycle, Own a bicycle, Use public transportation, City size	Y*	Disposable Income	Daily total expenditure per adult equivalent, City size
Av*	Availability	Supermarket share, Lean season, City size	Z*	Individual and Household Preferences	Opportunity cost of time (Percentage working age females in nonfarm labor, Dependency ratio), Adult equivalents, Own farmland, Own a stove, Own a hotplate, Maximum level of education, Age of household head, Female head of household, City size
D*	Desirability	Supermarket share, Lean season, City size	W*	Social and Individual Norms and Beliefs	Maximum level of education, Own a phone, City size
P*	Prices	City size	H*	Health Concerns	Maximum level of education, Female head of household, City size

Note: Supermarket share is only included for processing share dependent variables

Relative distance to supermarket is only included in the supermarket share model. Distance to supermarket is shown to be negatively correlated with the use of supermarkets (Gómez et al., 2013), and relative distance is a refinement of this measure in an effort to highlight the opportunity costs of transport to a specific retail location (Goldman, 2002 & 2005). Any effect of relative distance to supermarket on processed share would only occur via the use of

a supermarket. This variable is not intended to capture the effect of distance to household per se, rather it is to capture the effect of the marginal cost of transportation required to access a supermarket versus a non-supermarket retailer. It is calculated as the logarithmic value of the difference in kilometers from household residence to the nearest supermarket where the household purchased food and the distance to the nearest non-supermarket where the household purchased food. If a household did not acquire food from a supermarket, the median relative distance from household to supermarket among all households in the enumeration area is used. If the relative distance is negative⁴⁷, the distance measure is replaced with 0.001 to allow for the calculation of a logarithmic value.

Interaction variables of relative distance to supermarket multiplied by city dummy variables are included to allow for varying effects of relative distance to supermarket on supermarket share across the cities. Interaction variables for Lusaka, Kitwe and Kasama are included in the model, with the relative distance to supermarket variable capturing that effect in Mansa. I expect stronger negative effects of relative distance to supermarkets in larger cities due to increased congestion that would increase the cost of motorized transport.

Supermarket share is included as an independent variable in the processed share model. Supermarket share is likely to be endogenous to processed share due to self-selection. Potential endogeneity is not expected to be the result of simultaneity as processed food is also available from non-supermarkets and unprocessed food is available from supermarkets. Treatment of the potential endogeneity is addressed in the estimation section of the paper. Supermarket share is included as a determinant of the share of processed food consumed, as supermarkets are expected to increase the share of processed food consumption (Asfaw, 2008; Rischke et al.,

⁴⁷ Negative relative distance to supermarket occurs in less than half of one percent of observations.

2015). Shopping at a supermarket is expected to increase household access to a greater variety of food, including processed foods, to expose households to greater marketing efforts that increase the desirability of processed food, and to offer processed foods at reduced prices.

Interaction variables of supermarket share multiplied by city dummy variables are included to allow for varying effects of supermarket share on processed food consumption across the cities. Interaction variables for Lusaka, Kitwe and Kasama are included in the model, with the supermarket share variable capturing the effects of supermarket use in Mansa.

The following variables are included in both supermarket share and processed share models.

City dummy variables are included for Lusaka, Kitwe and Kasama, while the effects of a household residing in Mansa are captured in the intercept. These variables capture systematic differences across cities in the pooled regressions, identifying the partial direct effects of cities on the households' consumption decisions. A mediation model as described by Hayes & Preacher (2010) will be described in the estimation method of this paper and is used to identify indirect effects of city size on the consumption decisions.

Total daily household expenditure per AE is included in the model serving as a proxy for income, a key factor of consumer demand. It enters the estimation in both the logarithmic and squared logarithmic forms to remain consistent with the Banks et al. (1997) form of the Engel Curve model. The marginal effects of total expenditure are expected to vary across income, but generally increase with rising total expenditure for both supermarket and processed shares.

In addition to city size, the following variables are included to capture the effects of individual and household preferences: percentage working age females in nonfarm labor,

dependency ratio, adult equivalents, own farmland, own a gas or electric stove, own a hotplate, maximum level of education, age of household head, and female head of household.

A rising *percentage of working age females employed in nonfarm labor* is expected to increase the households' opportunity cost of time, resulting in increased demand for processed food (Senauer et al., 1986; Kennedy & Reardon, 1994); this increased opportunity cost of time is also expected to impact supermarket retail demand due to the time saving motivation of acquiring all foods from a single location. The calculation for this variable includes the number of working age⁴⁸ women who are employed in nonfarm labor divided by the number of working age women in the household. The *dependency ratio*, defined as a percentage of non-working age members of the household, represents potential increasing opportunity cost of time for those providing the support to these household dependents as well as a motivation to provide more healthy food options. *Household adult equivalents* account for the effect of household size on consumption decisions. *Household ownership of a farm* increases the likelihood that the household would consume food from own production, reducing the household demand to purchase food from supermarkets and, if consuming food from its own production, directly reducing the share of acquired processed food. *Household ownership of a stove (gas or electric)* and *household ownership of a hotplate* both reduce the preparation costs of unprocessed food, which leads to an expectation of reduced processed share. At the same time, ownership of these appliances indicates a willingness of the household to invest in time saving capital, which implies a higher opportunity cost of time that would lead to increased supermarket and processed shares. The *maximum years of completed education* by a member of the household represents the household capacity for more educated decisions. Higher education is associated with healthier

⁴⁸ Ages 15-64

food consumption decisions (Turrell & Kavanagh, 2006), potentially influencing the shopping decisions of households. *Age of the head of household* is included as elderly households are less likely to transition away from the more traditional diet and food acquisition patterns. *Female head of households* are likely to consume less food from supermarkets and consume less processed food, as females commonly place a greater emphasis on health and deferred utility (Smith, 2003).

Social and Individual Norms and Beliefs are captured with the variable of ownership of a mobile phone and the aforementioned city size and maximum years of completed education variables. *Ownership of a mobile phone* could indicate a household's openness to modernization and therefore willingness to purchase processed food and shop at supermarkets.

Variables that capture level of household health concerns are maximum level of education, female head of household and dependency ratio; each are described above.

In addition to city size and supermarket share, five variables are included to capture the effects of food environment. *Ownership of a car*, ownership of a motorcycle, *ownership of a bicycle* and *use of public transportation* directly affect access to markets. Household ownership of a car would reduce the time required in transportation to the supermarkets and correspondingly, the opportunity costs involved is purchasing food from a supermarket as well as increase household ability to buy and transport food purchased in bulk. Ownership of a motorcycle is expected to have a similar effect on transportation costs as ownership of a car, but offering less benefit on the ability to shop in bulk. Ownership of a bicycle also reduces the transportation costs associated with small purchases, but would likely have little effect on household ability to shop in bulk. The use of public transportation allows households to have improved access to markets without car, motorcycle or bicycle ownership. Each of these

transportation variables additionally indicate a reduced transportation cost of households bringing unprocessed food to processing centers, thereby reducing household cost of processing unprocessed food. The net effect of these variables is expected to be positive for supermarket and processed food shares.

The final variable that captures the effects of the food environment factors, availability and desirability, is the *lean season dummy variable*, where a value of one indicates the data were collected during the lean season. This variable captures differences in consumption decisions driven by the season when the consumption occurs. The lean season is typically defined by reduced stocks of food and higher food prices directly contributing to the availability and desirability of the food environment.

Table 3-8: Household characteristics of the sample

Independent Variable	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	8.1 (5.4)	8.9 (6.2)	7.0 (4.6)	4.3 (3.1)	4.5 (3.2)
Percentage Working Age Females Employed in Nonfarm Labor	34.2 (25.0)	34.7 (25.0)	32.5 (16.7)	34.9 (0.0)	34.9 (25.0)
Dependency Ratio (%)	39.8 (40.0)	39.3 (40.0)	39.7 (40.0)	44.9 (50.0)	42.6 (42.9)
Adult Equivalents	5.7 (5.4)	5.5 (5.1)	6.1 (5.9)	5.7 (5.4)	6.0 (5.9)
Relative Distance to Supermarket (km)	4.1 (3.0)	3.2 (2.8)	7.1 (7.5)	3.3 (2.5)	1.8 (1.5)
Percentage Own a Car	11.4 (0.0)	12.9 (0.0)	9.1 (0.0)	5.6 (0.0)	7.0 (0.0)
Percentage Own a Motorcycle	0.5 (0.0)	0.5 (0.0)	0.6 (0.0)	0.6 (0.0)	1.1 (0.0)
Percentage Who Use Public Transportation	71.3 (100.0)	73.3 (100.0)	77.1 (100.0)	40.9 (0.0)	36.9 (0.0)
Percentage Own a Bicycle	20.6 (0.0)	15.9 (0.0)	20.9 (0.0)	57.3 (100.0)	54.7 (100.0)
Percentage Own a Farm	14.6 (0.0)	8.7 (0.0)	18.7 (0.0)	50.8 (100.0)	46.3 (0.0)
Percentage Own a Stove (Gas or Electric)	34.4 (0.0)	34.0 (0.0)	40.5 (0.0)	20.9 (0.0)	21.9 (0.0)
Percentage Own a Hotplate	24.9 (0.0)	30.7 (0.0)	13.6 (0.0)	8.2 (0.0)	15.2 (0.0)
Percentage Own a Mobile Phone	75.1 (100.0)	77.2 (100.0)	74.9 (100.0)	56.4 (100.0)	63.4 (100.0)
Highest Completed Grade of Education in Household	11.3 (12.0)	11.4 (12.0)	11.1 (12.0)	10.4 (10.0)	10.8 (12.0)
Age of Household Head	43.7 (42.0)	42.9 (41.0)	45.6 (45.0)	44.3 (42.0)	45.3 (44.0)
Female Head of Household	19.0 (0.0)	20.0 (0.0)	16.6 (0.0)	17.2 (0.0)	20.9 (0.0)
Weighted Percentage of Sample	100	67.8	23.6	5.8	2.9
Number of Household Observations	4025	1330	1352	661	682

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

3.3.5 Estimation Method

The estimation of the empirical models is completed in two steps. The first step estimates the partial effects of all determinant variables on consumption decisions, including the direct effects of city size. The second step estimates the indirect effects that city size has on consumption through the other independent variables.

Estimating direct effects: A nested set of econometric issues needs to be dealt with in the estimation: the likely endogeneity of supermarket share, the proper estimator for share variables

bounded between zero and one, and methods for dealing with endogeneity depending on the estimator chosen⁴⁹. I discuss each in sequence in the next paragraphs.

Supermarket share is likely endogenous as an explanatory variable of processed share due to unobserved systematic differences between households that frequently shop at supermarkets and those that do not – differences that potentially affect household decisions regarding processed food consumption. To take one example, supermarket shoppers may have higher average incomes, thus higher opportunity costs of time, and thus a higher likelihood of consuming processed foods; supermarket shoppers could also be more favorably disposed towards “modern” consumption habits and for this reason choose processed foods.

Endogeneity would typically be dealt with by use of two stage least squares (2SLS). Relative distance to supermarket would be used as an instrumental variable (IV) as relative distance to supermarket has a clear impact on the costs of shopping at a supermarket, but should not independently affect the share of processed food consumption except through the purchasing of food from a supermarket. This approach has been applied previously by Rischke et al. (2015). The first stage regression would estimate supermarket share with equation 3.2.

$$S_i = \theta D_i + \varphi C_i + \tau Y_i + \rho X_i + \omega_i \quad (3.2)$$

where S_i is the share of food value purchased from supermarkets in total food value consumed by households; D_i is the log of the household net distance to supermarket that captures much of the effect of the access to market factor (Ac*) of food environment; C_i are city dummy variables for Lusaka, Kitwe and Kasama that capture direct effects of city size that affect multiple factors of food environment and consumer demand; Y_i are the logarithmic and squared logarithmic forms of daily household total expenditure per adult equivalent, which proxies for the effect of

⁴⁹ OLS regression results are included in the appendix as a robustness check to the estimations that control for potential endogeneity.

income (Y^*) on consumption decisions; X_i are other demographic and seasonal dummy variables identified in the general equation that affect multiple factors of food environment and consumer demand; θ , φ , τ and ρ are estimated coefficients; and ω_i is the error term. An F-test of the significance of distance to supermarket on SM share results in a value greater than ten; suggesting that distance to supermarket is a strong IV (Staiger & Stock 1994).

The second stage in 2SLS estimates processed share continues to be based on equation 1 while incorporating household estimates of supermarket share from the first stage regression, as shown in the following equation:

$$F_i = \gamma \hat{S}_i + \alpha C_i + \vartheta(C_i * \hat{S}_i) + \beta Y_i + \delta X_i + \varepsilon_i \quad (3.3)$$

where F_i is the share of processed food value in total household food consumption; $\hat{S}_i$ is the estimated share of food value purchased from supermarkets in total food value consumed by households (from equation 2); $C_i * \hat{S}_i$ are dummy variables of Lusaka, Kitwe and Kasama interacted with the estimated supermarket share; C_i , Y_i , and X_i represent the same variables as indicated in equation 2; γ , α , ϑ , β and δ are estimated coefficients; and ε_i is the error term.

2SLS presents a weakness in this case, however: the bounded nature of the variables makes the fractional probit (FP) the best estimator, and the non-linearity of the FP model requires the use of a control function to account for endogeneity, not 2SLS. The shares of processed food and of supermarket purchases in total value of consumed food are never greater than one and never lower than zero. Given the characteristics of the dependent variables and the assumption of asymptotically normal standard errors, the appropriate method of estimation is the FP model (Papke & Wooldridge 1996 & 2008).

Due to the nonlinearity of the FP a control function is used to control for endogeneity (Smith & Blundell, 1986). A control function uses the residuals of the first stage regression as an

explanatory variable in the second stage, together with actual values of the endogenous variable. Insufficient methodology pertaining to predicting residuals from the FP model in the first stage regression requires the estimation of the first stage with OLS (eq. 4) and the second stage using FP (eq. 5) to apply the control function that accounts for potential endogeneity (Wooldridge, 2010).

$$S_i = \theta D_i + \varphi C_i + \tau Y_i + \rho X_i + \omega_i \quad (3.4)$$

$$F_i = \gamma S_i + \alpha C_i + \vartheta(C_i * S_i) + \beta Y_i + \delta X_i + \pi \widehat{\omega}_i + \varepsilon_i \quad (3.5)$$

Variables included in equations 4 and 5 mirror those of equations 2 and 3 with the following exceptions: S_i is included in the processed share regression as observed at the household level and the predicted errors from 4 are included in 5. Regarding discussion specific to supermarket use, I maintain the use of the FP model in the estimation of equation 4 as over two thirds of the population consumes zero percent of their food from supermarkets. The OLS estimates used in the first stage of processed share estimation are reported in the appendix (Table 3-A-2). For an initial robustness check the model is also estimated using 2SLS (Tables 3-A-8, 3-A-9, & 3-A-10), and these estimates can be found in the appendix with only minor differences from the FP estimates.

As household expenditure per adult equivalent enters the estimation in both logarithmic and squared logarithmic forms (Y_i) the estimates of the average partial effects of expenditure are calculated as follows:

$$\frac{\partial}{\partial(\text{expenditure})} = \left[a \frac{1}{\text{expenditure}} + 2b \frac{\ln(\text{expenditure})}{\text{expenditure}} \right] * E\phi[X\beta] \quad (3.6)$$

where a equals the marginal effects of log expenditure and b equals the marginal effects of squared log expenditure from equations 4 and 5 for supermarket share or processed share

respectively, and is equal to the average of the normal densities of the predicted values of the dependent variable.

A second robustness check is to take advantage of the two rounds of panel data by using a correlated random effects (CRE) model to control for unobserved household heterogeneity. As in 4 and 5, the first stage is estimated with OLS and the second stage with FP and a control function to account for potential endogeneity:

$$S_i = \theta D_i + \varphi C_i + \tau Y_i + \rho X_i + \sigma \bar{Z}_i + \omega_i \quad (3.7)$$

$$F_i = \gamma S_i + \alpha C_i + \vartheta(C_i * S_i) + \beta Y_i + \delta X_i + \mu \bar{Z}_i + \pi \widehat{\omega}_i + \varepsilon_i \quad (3.8)$$

Variables included in equations 7 and 8 mirror those of equations 4 and 5 with the addition of $\bar{Z}_i$, which represents the household averages of the time variant determinant variables. Due to low variance in the Z_i between the two rounds, results of the CRE are very close to those from 4 and 5; the CRE output can be found in the appendix (Tables 3-A-11 – 3-A-15).

Finally, in addition to estimating the models with city dummy variables, I drop the city dummy variables and repeat all of the estimations separately for each city.

Estimating indirect effects: The above estimation methods calculate the direct effects of cities on the consumption decisions of supermarket share and processed share. The estimation of indirect effects is completed by treating the non-city determinant variables as mediator variables. A mediator variable acts as a conduit for a partial effect of a predictor variable (city variables). The indirect effect of the predictor variable is calculated by multiplying the estimated effect of the mediator⁵⁰ on the dependent variable by an estimated effect of the predictor variable on the mediator:

⁵⁰ When total expenditure is the mediator variable, it should be noted that the marginal effect of total expenditure on the dependent variable follows from equation 6.

$$\text{indirect effect} = \frac{\partial(\text{mediator variable})}{\partial(\text{city variable})} * \frac{\partial(\text{dependent variable})}{\partial(\text{mediator variable})} \quad (3.9)$$

Total effects of the predictor variable on the dependent variable are the summation of significant direct and indirect effects. OLS is used to estimate city indirect effects on the independent variables from the model. Bootstrapping is used to obtain asymptotic approximations of the statistical significance of the indirect effects (Wooldridge, 2010). This estimation of indirect effects follows the methodology described in Hayes & Preacher (2010).

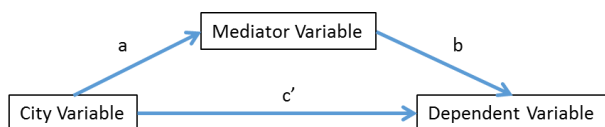


Figure 3-2: Estimation of indirect effects

3.4 Results

3.4.1 Descriptive Statistics

Supermarket share and processed share vary across city size (Table 3-9). Households from tertiary cities choose to purchase at least some of their food from supermarkets at a greater rate than households from larger cities⁵¹, and also obtain a greater share of their food from supermarkets⁵². These initial relationships between supermarket utilization and city size follow the potential negative relationship highlighted in the conceptual model. The primary and secondary cities have average incomes fifty to one hundred percent higher than the tertiary cities, which is expected to increase supermarket use, but this alone does not result in a greater share of supermarket adoption. Higher congestion and increased opportunity costs are effects of city size on urban environment that could lead to this relationship.

⁵¹ 46% versus 32% and 35%

⁵² 7% versus 5% and 4%

Table 3.9: Household weighted averages of consumption shares

	Full Sample Population				Subset of Households that Bought Some Food from a Supermarket		
	Share of food from supermarkets in total food consumption	Share of processed food in total food	Share of food from supermarkets that is processed	Percentage of households that shop at supermarkets	Share of food from supermarkets in total food consumption	Share of processed food in total food	Share of processed food from supermarkets in total processed food
Weighted Average	5.1	65.2	86.8	33.7	15.1	69.4	19.3
(all cities)	(0.0)	(67.1)	(100)	(0.0)	(8.9)	(70.6)	(12.3)
Lusaka	5.4	67.0	87.3	31.9	17.0	70.4	21.5
	(0.0)	(68.3)	(100)	(0.0)	(11.0)	(71.3)	(14.2)
Kitwe	3.5	64.8	85.8	34.5	10.2	70.2	12.7
	(0.0)	(67.1)	(100)	(0.0)	(6.4)	(70.8)	(8.3)
Kasama & Mansa	6.7	52.4	86.5	46.1	14.6	62.0	20.7
Average	(0.0)	(53.9)	(100)	(0.0)	(11.2)	(64.3)	(16.6)
Kasama	5.6	50.1	85.7	41.1	13.7	60.0	19.7
	(0.0)	(50.7)	(100)	(0.0)	(9.8)	(61.4)	(14.6)
Mansa	8.9	57.2	87.6	56.2	15.9	64.8	22.1
	(3.1)	(61.9)	(100)	(100)	(13.0)	(67.0)	(18.0)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

The hypothesis linking city size to greater consumption of processed food is consistent with the results in Table 3-9, with average processed shares more than ten percent greater in larger cities. Over eighty-five percent of food purchased from supermarkets is processed, which is greater than the average processed share in any city, so it could reasonably be inferred that purchasing food from supermarkets will increase processed share. Households that purchase a portion of their food from supermarkets have average processed shares greater than the entire sample of households; this is consistent with theory based on supermarkets increasing the desirability of processed food by pricing it at lower prices than non-supermarket retailers. However, given that households from small cities have larger supermarket shares and households from larger cities have larger processed shares, regression analysis is required to determine if supermarkets increase the share of consumed processed food.

Table 3-9 also shows that households that purchase a portion of their food from supermarkets continue to purchase approximately eighty percent of their processed food, by value, from retailers other than supermarkets, highlighting that processed food remains primarily sourced from non-supermarkets retailers even to households that shop at supermarkets. Non-supermarket retailers are the dominant source of food in urban Zambia (over eighty percent), with supermarkets providing a smaller yet nontrivial source of all food – primarily driven by high processed food and low processed maize (Table 3-10).

Table 3-10: Household weighted averages of sources of food aggregates

	All Food	All Processed	Processed High	Processed Low	Maize - Low Processed	Non-maize Low Processed
Supermarkets	5.1	6.5	7.0	6.5	8.4	5.9
Non-supermarkets	81.4	81.3	76.0	91.3	88.4	92.5
Food Away From Home	7.8	10.4	16.0	0.0	0.0	0.0
Gifts	2.1	1.7	1.0	2.2	3.1	1.6
Own Production	3.6	0.0	0.0	0.0	0.0	0.0
Total	100	100	100	100	100	100

Source: Zambia Urban Consumption Survey 2007/2008

Differences exist in the consumption patterns of low and high processed food across the cities within this sample (Table 3-11). Lusaka, the largest city in Zambia, is the only city where the average value of consumed high processed is greater than low processed food. The shares of processed food aggregates from supermarkets also vary such that the supermarkets in larger cities source considerable more high processed food than low processed food, relative to supermarkets in tertiary cities. The ratio of consumed value of high processed food to low processed food is greater for households that purchase some food from supermarkets relative to the total population, in all cities other than Mansa, suggesting that supermarkets have stronger positive effects on the consumption of high processed food than low processed food. The affect that supermarket use has on processed food consumption will be further addressed within the regression results.

Table 3-11: Household weighted averages of consumption shares by low and high processing levels

	Full Sample Population				Subset of Households that Bought Some Food from a Supermarket			
	Share of low processed food in total food	Share of high processed food in total food	Share of food from supermarkets that is low processed	Share of food from supermarkets that is high processed	Share of low processed food in total food	Share of high processed food in total food	Share of low processed food from supermarkets in total low processed food	Share of high processed food from supermarkets in total high processed food
Weighted Average (all cities)	33.5 (33.6)	31.8 (29.3)	36.6 (28.3)	50.3 (47.6)	33.7 (34.1)	35.7 (33.2)	19.3 (9.4)	20.8 (11.5)
Lusaka	33.4 (33.4)	33.6 (31.7)	38.3 (29.9)	48.9 (46.2)	33.0 (32.8)	37.4 (34.1)	22.1 (11.2)	22.7 (12.5)
Kitwe	35.2 (35.9)	29.6 (27.4)	28.4 (6.1)	57.4 (63.7)	35.5 (36.0)	34.7 (32.2)	11.0 (1.3)	15.9 (9.5)
Kasama & Mansa Average	29.3 (29.3)	23.1 (19.1)	43.6 (44.4)	42.8 (37.7)	33.6 (33.8)	28.3 (25.6)	20.9 (16.7)	20.8 (12.6)
Kasama	28.7 (28.7)	21.4 (17.8)	41.1 (41.2)	44.5 (41.4)	32.8 (32.9)	27.3 (24.3)	19.2 (15.5)	20.4 (11.2)
Mansa	30.5 (31.3)	26.7 (23.1)	47.3 (48.4)	40.3 (34.5)	34.9 (35.6)	29.9 (26.5)	23.5 (17.6)	21.3 (14.0)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the population weighted average. Parentheses indicate the population weighted median.

3.4.4 Regression Results

Direct effects of determinant variables on supermarket share: Table 3-12 presents the regression results that address how city size affects supermarket share. Five observations are noted here. First, city size, independent of household expenditure and other drivers of consumption, has a negative direct effect on supermarket share, as evidenced by the negative average partial effects of Lusaka and Kitwe dummy variables on supermarket share in the pooled regression. This finding is consistent with descriptive statistics that identify tertiary cities as having greater shares of consumption from supermarkets. The conceptual framework highlighted this potential result, indicating that increased city size could contribute to reduced use of supermarkets due to household response to congestion and the asymmetric effects of investment costs.

Table 3-12: Fractional probit estimates of the average partial effects of the determinants of household share of supermarket purchases in total food consumption

Variable	Share Supermarket in Total Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.004** (0.000)	0.005** (0.000)	0.003** (0.000)	0.012** (0.000)	0.011** (0.000)
Lusaka dummy variable	-0.075** (0.000)				
Kitwe dummy variable	-0.049** (0.000)				
Kasama dummy variable	-0.017** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	-0.007 (0.201)	-0.008 (0.290)	-0.005 (0.308)	-0.004 (0.673)	-0.014 (0.314)
Dependency Ratio	0.020 (0.108)	0.030 (0.077)	-0.007 (0.499)	0.013 (0.440)	0.008 (0.752)
Adult Equivalents	0.002 (0.262)	0.004 (0.164)	-0.001 (0.446)	-0.003 (0.266)	0.005 (0.057)
Relative Distance to Supermarket in Lusaka (km) (log)	-0.008** (0.000)	-0.011** (0.000)			
Relative Distance to Supermarket in Kitwe (km) (log)	-0.005** (0.000)		-0.005** (0.000)		
Relative Distance to Supermarket in Kasama (km) (log)	-0.006** (0.005)			-0.002 (0.347)	
Relative Distance to Supermarket in Mansa (km) (log)	-0.008* (0.024)				-0.007** (0.007)
Own a Car	0.019* (0.016)	0.023* (0.030)	0.002 (0.670)	0.041* (0.033)	-0.009 (0.486)
Own a Motorcycle	-0.017 (0.154)	-0.025 (0.100)	0.005 (0.730)	0.015 (0.409)	-0.021 (0.428)
Use Public Transit	-0.005 (0.347)	-0.007 (0.333)	0.001 (0.723)	0.001 (0.855)	-0.010 (0.230)
Own a Bicycle	-0.002 (0.678)	-0.003 (0.692)	-0.001 (0.792)	-0.009 (0.236)	0.010 (0.400)
Own Farmland	-0.002 (0.680)	-0.001** (0.882)	0.001 (0.845)	-0.004 (0.563)	-0.026* (0.011)
Own a Stove (Gas or Electric)	0.015* (0.040)	0.008** (0.454)	0.025** (0.000)	0.018 (0.091)	0.044** (0.000)
Own a Hotplate	0.007 (0.290)	0.004 (0.652)	0.019** (0.006)	0.003 (0.742)	0.029* (0.020)
Own a Mobile Phone	0.009 (0.194)	0.010 (0.343)	0.001 (0.814)	0.009 (0.392)	0.040** (0.004)
Highest Completed Grade of Education in Household	0.007** (0.000)	0.007** (0.000)	0.005** (0.000)	0.009** (0.000)	0.003 (0.154)
Age of Household Head	0.000 (0.228)	0.000 (0.265)	0.000 (0.856)	0.000 (0.284)	-0.001* (0.024)
Female Head of Household	0.014 (0.074)	0.015 (0.148)	0.016 (0.055)	-0.006 (0.583)	0.001 (0.961)
Lean Season dummy variable	-0.007* (0.024)	-0.007 (0.123)	-0.008** (0.007)	-0.002 (0.737)	-0.009 (0.189)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Second, income has a positive average partial effect on supermarket use at all expenditure levels in all cities with magnitudes twice as large in tertiary cities than in larger cities. The direct effects of total expenditure are shown across the income spectrum and by city in the left panel of Figure 3-3, which are calculated from the same estimation used to generate Table 3-12. The

difference in average partial effects of expenditure along with the direct effect of city size on supermarket share provide justification for households within tertiary cities having larger supermarket shares than households within primary and secondary cities in spite of higher⁵³ total expenditure in primary and secondary cities.

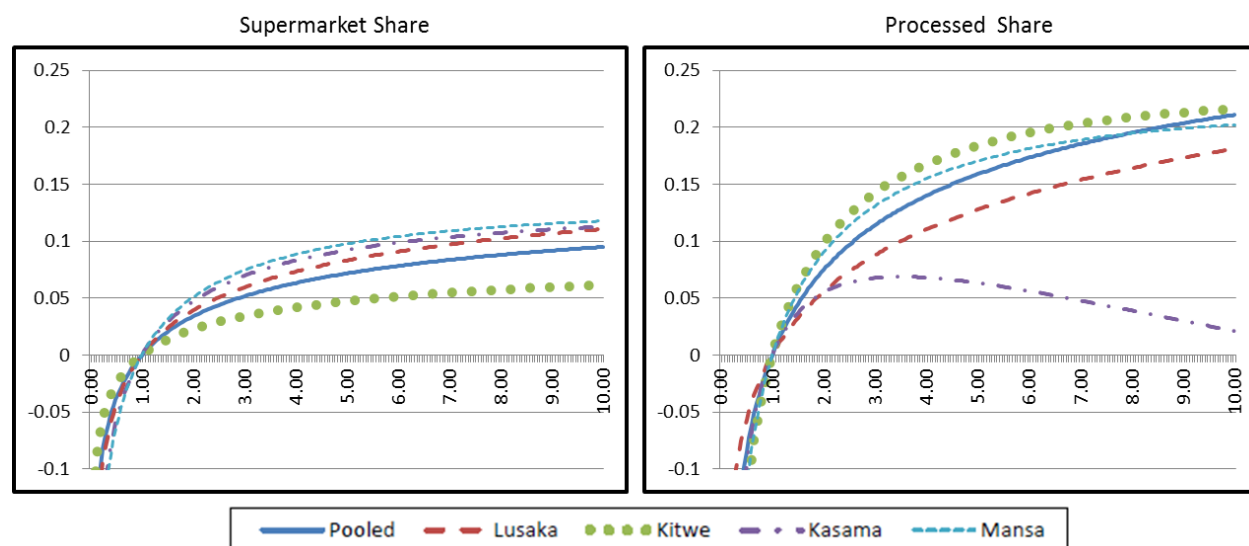


Figure 3-3: Estimated effects of expenditure on supermarket use and processed food consumption

Third, relative distance to supermarkets has a larger – nearly two times larger – negative effect on supermarket share in the primary city compared to all other cities. This identifies greater sensitivity to the distance that is required to travel to supermarkets in larger cities, consistent with the hypothesis that larger cities have greater congestion that increases transportation costs. Additionally, due to a greater presence of non-supermarket retailers, the sensitivity to distance could result from nearby non-supermarket retailers that would reduce household willingness to expend the additional transportation costs to shop at a supermarket.

⁵³ The average expenditures in the primary city are two times those of the tertiary cities, and the average expenditures in secondary cities are more than fifty percent greater than those in tertiary cities.

Fourth, education has a strong positive effect on supermarket share with no evidence that the effect varies systematically with city size. The direction of the average partial effect is consistent with the expectation that more educated households will place a premium on the health benefits associated with the cleanliness of supermarkets. The consistency of the effect of education across city size shows the transferability of education across demographics.

Fifth, although expected to affect supermarket share, the percentage of working age females employed in nonfarm labor does not have a statistically significant effect. This variable is a proxy for opportunity cost of time, which is also controlled for by total household expenditure. The effect of the additional income earned by female nonfarm employment would be captured by the expenditure variables, but this finding shows that the time that females allocate towards nonfarm employment does not statistically affect supermarket share.

Direct effects of determinant variables on processed share: Table 3-13 addresses city size effects on processed food consumption with seven observations out of note. First, city size positively affects the consumption of processed food when controlling for other determinant variables, as indicated by the positive and significant average partial effects of primary and secondary city dummy variables. This is consistent with the primary hypothesis regarding processed share, which following the conceptual framework highlighted only positive expected effects of city size on urban environment and ultimately processed food consumption.

Table 3-13: Second stage of control function regression; estimating the share of processed food in total food consumption

Variable	Share of Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.010** (0.000)	0.009** (0.000)	0.011** (0.000)	0.003 (0.983)	0.019 (0.770)
Household shops at a supermarket in Lusaka	0.007 (0.945)	0.021 (0.855)			
Household shops at a supermarket in Kitwe	-0.091 (0.425)		0.337 (0.150)		
Household shops at a supermarket in Kasama	0.168 (0.196)			2.867** (0.006)	
Household shops at a supermarket in Mansa	0.411** (0.001)				0.843* (0.021)
Lusaka dummy variable	0.035** (0.010)				
Kitwe dummy variable	0.037** (0.003)				
Kasama dummy variable	-0.049** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.006 (0.375)	-0.002 (0.849)	0.021 (0.051)	0.018 (0.266)	0.046* (0.013)
Dependency Ratio	-0.006 (0.683)	-0.005 (0.801)	0.015 (0.454)	-0.048 (0.212)	-0.044 (0.242)
Adult Equivalents	0.006** (0.001)	0.006* (0.019)	0.005* (0.023)	0.015* (0.013)	0.007 (0.074)
Own a Car	-0.016 (0.116)	-0.015 (0.257)	-0.004 (0.786)	-0.210** (0.004)	-0.020 (0.327)
Own a Motorcycle	0.002 (0.930)	0.030 (0.416)	-0.067* (0.019)	-0.100 (0.055)	-0.033 (0.266)
Use Public Transit	0.007 (0.317)	0.013 (0.158)	0.003 (0.758)	-0.013 (0.357)	-0.006 (0.700)
Own a Bicycle	-0.004 (0.522)	-0.009 (0.372)	0.013 (0.235)	0.015 (0.344)	-0.046** (0.002)
Own Farmland	-0.037** (0.000)	-0.042** (0.004)	-0.035** (0.002)	-0.023 (0.142)	-0.033 (0.052)
Own a Stove (Gas or Electric)	-0.005 (0.591)	-0.014 (0.257)	0.018 (0.141)	-0.117* (0.016)	-0.028 (0.386)
Own a Hotplate	-0.014 (0.073)	-0.018 (0.055)	-0.003 (0.838)	-0.051 (0.053)	0.013 (0.585)
Own a Mobile Phone	0.014 (0.108)	0.012 (0.313)	0.017 (0.199)	0.038 (0.054)	0.010 (0.651)
Highest Completed Grade of Education in Household	0.001 (0.692)	0.002 (0.467)	-0.002 (0.461)	-0.020* (0.022)	-0.005 (0.127)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.001 (0.071)	-0.002** (0.000)	-0.002* (0.014)
Female Head of Household	-0.012 (0.169)	-0.007 (0.506)	-0.022 (0.080)	-0.009 (0.716)	-0.032 (0.092)
Lean Season dummy variable	-0.006 (0.253)	-0.006 (0.445)	-0.011 (0.136)	0.000 (0.989)	0.036** (0.006)
Endogenous Error from 1st Stage Regression	-0.183 (0.092)	-0.132 (0.272)	-0.410 (0.085)	-2.639* (0.011)	-0.712* (0.046)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Second, shopping in supermarkets significantly increases processed share within tertiary cities, but not within primary or secondary cities. Relative to tertiary cities, this finding is consistent with the Rischke et al. (2015) study that found a positive and significant effect of

supermarket share on processed food consumption with their data sample of three tertiary cities within Kenya. The positive, though of low statistical significance, average partial effects of supermarket share on processed share in the primary and secondary cities limit the ability to expand this relationship beyond smaller cities (Table 3-13).

Third, total household expenditure has a positive effect on processed share in primary and secondary cities, but not within tertiary cities. The right panel of Figure 3-3 shows the average partial effects of income on processed food by city, regardless of the statistical significance, and this shows positive effects at all levels of expenditure for all cities other than Kasama. The overall pattern of expenditure on processed food consumption is as hypothesized.

The descriptive statistics show that processed food is available to be acquired from supermarkets and non-supermarkets alike in each city in the survey. However, these results show that the consumption of processed food in tertiary cities is access driven by the presence and utilization of supermarket retailers, whereas the consumption of processed food in primary and secondary cities is income driven – responding to changes in total expenditure.

Fourth, ownership of farmland negatively affects the consumption of processed food, with little evidence of systematic variance in the effect across city size. Although the average partial effects do not systematically vary across city size, the impact of these effects are more pertinent to households within tertiary cities as more than forty percent of tertiary households own some farmland, relative to nine percent of households within primary cities.

Fifth, age of household head negatively affects the consumption of processed food, without evidence of systematic variance in the affect across city size. This effect is expected as elderly households are less likely to adopt the less traditional consumption patterns of consuming processed food.

Sixth, education is found to have a statistically significant negative effect in only one of the cities. This result is minimally consistent with the expectation of increased education leading to healthier consumption decisions given that processed foods are generally considered less healthy than unprocessed foods.

Seventh, the percentage of working age females employed in nonfarm employment has positive and significant effects on processed share for households in Kitwe and Mansa, in accord with the findings of Senauer et al. (1986) and Kennedy & Reardon (1994). There is inadequate evidence to identify a pattern regarding city size specific effects of participation in nonfarm employment.

The consumption patterns of processed food are further analyzed by separately considering the drivers of the consumption for low and high processed food. Five observations are highlighted from Tables 3-A-3 and 3-A-4 that are located within the appendix.

First, similar to the pattern observed regarding the processed share of households within tertiary cities, supermarket share positively effects the consumption of low processed food by households within tertiary cities. The positive average partial effects of supermarket share on processed food and on low processed food shows that the increased processed food consumption in tertiary cities due to supermarkets is primarily driven by increased consumption of low processed food such as maize meal.

Second, the consumption of highly processed food is not directly driven by the use of supermarket as supermarket share was not statistically significant for high processed food. This finding was not expected following the patterns observed in the descriptive statistics, which showed households that use supermarkets have greater consumption of high processed food relative to low processed food than those who do not use supermarkets (Table 3-11).

Third, households within larger cities transition from low processed food to high processed food as household total expenditures rise. This is shown in Tables 3-A-3 and 3-A-4 by the negative average partial effects of total expenditure on low processed food and positive average partial effects of total expenditure on high processed food. The larger absolute value of the average partial effects on high processed food shows that demand for highly processed food is driving the rise in consumption of processed food within primary and secondary cities.

Fourth, the average partial effect of expenditure is not significant on low or high processed food for households within tertiary cities, further reinforcing the idea that consumption of processed food within tertiary cities is access driven more than income driven.

Fifth, seasonal variability exists regarding the level of processed food consumed by households. During the lean season there is an increase in the consumption of low processed food and relative decrease in the consumption of high processed food. This is likely driven by increased consumption of purchased maize meal during the lean season, as maize grain becomes scarce on the market at this time and households therefore must acquire it as maize meal.

Indirect effects of determinant variables on supermarket share and processed share: The effects of city size on supermarket and processed share also include indirect effects that are driven by the effects that city size has on other determinant variables. Table 3-14 highlights the indirect effects of city size on supermarket share and processed share via the mediator variables of relative distance to supermarkets, total daily expenditure per adult equivalent and percentage of working age females employed in nonfarm labor. The indirect effects of city size are not statistically significant via the percentage of working age females employed in nonfarm labor, while the indirect effects via relative distance to supermarkets and total daily expenditure per adult equivalent are statistically significant and are noted below.

Table 3-14: City indirect effects via specific explanatory variables

	Supermarket share			Processed Share	
	Distance to Supermarket minus distance to nearest retailer (log)	Total Daily Household Expenditure per Adult Equivalent	Percentage Working Age Females Employed in Nonfarm Labor	Total Daily Household Expenditure per Adult Equivalent	Percentage Working Age Females Employed in Nonfarm Labor
Lusaka	-0.004* (0.025)	0.012** (0.000)	0.000 (0.808)	0.027** (0.000)	0.000 (0.851)
Kitwe	-0.011** (0.001)	0.009** (0.001)	0.000 (0.912)	0.019** (0.000)	0.000 (0.922)
Kasama	-0.005 (0.064)	0.000 (0.886)	0.000 (0.940)	0.001 (0.885)	0.000 (0.955)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

The indirect effect of city size on supermarket share via relative distance to supermarket is in agreement with the direct effect: negative and statistically significant. The conceptual model posits that as city size increases the market size will also increase, incentivizing investment in supermarkets and non-supermarkets alike. The increased investment in non-supermarkets increases the likelihood that a non-supermarket would be located near the household, therefore increasing the relative distance to a supermarket. This combined with the previously noted negative impact of relative distance to supermarkets drive a negative indirect effect of city size on supermarket share via relative distance to supermarket; strengthening the total negative effect of city size on supermarket share.

Indirect Effect: City Size $\uparrow \rightarrow$ relative distance to supermarket $\uparrow \rightarrow$ % supermarket $\downarrow$

Direct Effect: City Size $\uparrow \rightarrow$ % supermarket $\downarrow$

Total Effect: City Size $\uparrow \rightarrow$ % supermarket $\downarrow \downarrow$

The indirect effect of city size on supermarket share via total income is positive and significant for primary and secondary cities when calculated at the average city level of total income. Increasing city size provides increased earning opportunities for households, positively increasing consumer demand for normal and luxury goods. This indirect effect is in agreement

with the common expectation that increased city size would positively affect supermarket share, a luxury in developing countries. The direct effects captured by the primary and secondary city dummy variables are negative and of greater absolute value than the indirect effects, thereby dominating the indirect effects and resulting in negative total effects of primary and secondary cities on supermarket share.

Indirect Effect: City Size $\uparrow \rightarrow$ hh income $\uparrow \rightarrow$ % supermarket $\uparrow$

Direct Effect: City Size $\uparrow \rightarrow$ % supermarket $\downarrow$

Total Effect: City Size $\uparrow \rightarrow$ % supermarket $\downarrow$

The indirect effect of city size on processed share via total income has a positive effect when calculated at the average city level of total income in primary and secondary cities. These positive indirect effects strengthen the effects that city size has on processed share beyond that observed by the direct effects.

Indirect Effect: City Size $\uparrow \rightarrow$ hh income $\uparrow \rightarrow$ % processed $\uparrow$

Direct Effect: City Size $\uparrow \rightarrow$ % processed $\uparrow$

Total Effect: City Size $\uparrow \rightarrow$ % processed $\uparrow \uparrow$

3.5 Conclusions

Analyzing the urban household data from Zambia, one observes that household consumption patterns pertaining to the share of food purchased from supermarkets in total food consumption and the share of processed food in total food consumption do vary across city size and that city size is a significant determinant of these consumption patterns.

Conventional wisdom indicates that the share of food from supermarkets in total food consumption would be greater for households in larger cities than for households in smaller

cities. This analysis contradicts the conventional wisdom in finding that households within primary and secondary cities shop at supermarkets less often and have lower average supermarket shares than households from tertiary cities. The conceptual framework recognizes that beyond the threshold of city size required to support a supermarket, a greater city size could in fact reduce the use of supermarkets driven by the urban environment affecting the food environment such that the absolute or relative access to supermarkets is reduced. Large cities do have a positive indirect effect on supermarket share through household total expenditure, a statistically significant mediator variable, but again the negative direct effect of city size continues to dominate the total effect.

Processed food consumption patterns follow conventional wisdom across city size, where households from larger cities consume a greater average share of processed food in their diet than households in smaller cities.

This analysis affirms previous studies that indicate that shopping at supermarkets has a positive effect on the share of processed food consumption in tertiary cities. This study observes the same directional effect of supermarket share on processed share in primary and secondary cities, but these average partial effects lack the statistical significance to definitively confirm the pattern.

Finally, consistent with the general finding that supermarkets increase processed food consumption, a subsample of households from all city sizes that purchase at least some of their food from supermarkets consume greater shares of processed food than the entire data sample. This same subsample of households continues to purchase approximately eighty percent of their processed food value from non-supermarket retailers. This later finding is a reminder that

although supermarket use is shown to have a positive effect on the consumption of processed food, the majority of processed food continues to be sourced from non-supermarket retailers.

APPENDIX

APPENDIX

Current Aggregates:	Unprocessed	Low Processed		High Processed	
Monteiro Aggregates:	Unprocessed or Minimally Processed	Processed Culinary	or Ingredients	Ultra-Processed	
	Whole grains	Rice	Sugar	Vegetable oils	Breads
	Fresh fruits	Butchered meat	Spices	Vegetable fats	Biscuits
	Roots & tubers	Dried fish	Wheat flour	Pasta	Soft drinks
	Fresh vegetables		Maize meal	<i>Pasteurized milk*</i>	Cheeses
	Fresh fish		Other flours		Prepared food away from home
	Eggs				

Figure 3-A-1: Processing level aggregates compared with aggregates from Monteiro et al. 2010

Note: Pasteurized milk is considered to be unprocessed or minimally processed by Monteiro et al. 2010 and high processed in this paper.

Table 3-A-1: Prices of commonly consumed items

Products	Prices - Round 1				Prices - Round 2				Average Prices			
	Lusaka	Kitwe	Kasama	Mansa	Lusaka	Kitwe	Kasama	Mansa	Lusaka	Kitwe	Kasama	Mansa
Maize meal	1,633	1,840	1,728	1,812	1,497	1,349	2,002	2,499	1,565	1,594	1,865	2,156
Bread	3,801	4,664	4,117	3,964	4,509	6,117	5,895	5,232	4,155	5,391	5,006	4,598
Chicken and other poultry	12,010	11,233	10,260	9,639	14,041	14,236	11,792	10,490	13,025	12,735	11,026	10,065
Beef and other red meat	15,402	16,308	12,560	15,983	16,766	16,551	12,931	16,317	16,084	16,429	12,746	16,150
Cooking oil	8,214	8,827	7,730	9,211	9,995	9,556	8,340	9,648	9,105	9,192	8,035	9,430
Rape	1,911	1,730	1,563	1,013	3,236	2,268	1,271	1,569	2,574	1,999	1,417	1,291
Tomato	1,883	1,895	1,598	1,458	5,628	5,511	2,529	3,706	3,756	3,703	2,064	2,582
Sugar	4,340	4,372	4,427	4,314	4,809	4,138	4,529	4,945	4,575	4,255	4,478	4,629
Rice	4,079	6,088	4,644	4,403	4,421	4,881	4,053	4,091	4,250	5,485	4,349	4,247
Dried fish	20,824	27,068	21,745	23,547	29,376	17,852	16,605	23,722	25,100	22,460	19,175	23,634
Kapenta	34,220	28,112	18,132	14,274	37,068	19,630	18,509	17,916	35,644	23,871	18,320	16,095
Beans	6,406	4,967	3,938	3,050	10,846	7,624	5,464	5,953	8,626	6,295	4,701	4,502
Price Index	100	108.7	92.0	99.1	100	89.5	76.1	81.5	100	100.9	85.6	94.1

Note: Price indices are weighted by the consumption shares of the top ten items consumed specific to the round in which it was calculated.

Prices are presented in Zambian Kwacha.

Table 3-A-2: First stage of control function regression; OLS estimates of the determinants of the share of supermarket purchases in total food consumption

Variable	Share Supermarket in Total Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.004** (0.000)	0.004** (0.000)	0.003** (0.000)	0.008** (0.000)	0.007** (0.000)
Lusaka dummy variable	-0.051** (0.000)				
Kitwe dummy variable	-0.052** (0.000)				
Kasama dummy variable	-0.019** (0.005)				
Percentage Working Age Females Employed in Nonfarm Labor	-0.004 (0.422)	-0.006 (0.406)	-0.003 (0.516)	0.002 (0.800)	-0.015 (0.232)
Dependency Ratio	0.025 (0.053)	0.038* (0.034)	-0.006 (0.599)	0.014 (0.443)	0.015 (0.565)
Adult Equivalents	0.001 (0.639)	0.002 (0.530)	-0.001 (0.235)	-0.004 (0.075)	0.003 (0.285)
Relative Distance to Supermarket in Lusaka (km) (log)	-0.020** (0.000)	-0.021** (0.000)			
Relative Distance to Supermarket in Kitwe (km) (log)	-0.011* (0.011)		-0.008** (0.000)		
Relative Distance to Supermarket in Kasama (km) (log)	-0.014 (0.045)			-0.004 (0.437)	
Relative Distance to Supermarket in Mansa (km) (log)	-0.014** (0.002)				-0.008* (0.030)
Own a Car	0.047** (0.001)	0.051** (0.007)	0.019 (0.158)	0.082* (0.020)	-0.007 (0.786)
Own a Motorcycle	-0.032 (0.321)	-0.059 (0.209)	0.006 (0.759)	0.040 (0.436)	-0.026 (0.642)
Use Public Transit	-0.007 (0.174)	-0.011 (0.156)	0.002 (0.705)	0.002 (0.838)	-0.007 (0.454)
Own a Bicycle	0.002 (0.705)	0.003 (0.801)	-0.002 (0.754)	-0.009 (0.237)	0.006 (0.626)
Own Farmland	0.001 (0.935)	0.006 (0.676)	0.002 (0.838)	-0.005 (0.469)	-0.022* (0.028)
Own a Stove (Gas or Electric)	0.018* (0.018)	0.013 (0.287)	0.020** (0.000)	0.042* (0.019)	0.079** (0.000)
Own a Hotplate	0.003 (0.690)	-0.001 (0.886)	0.007 (0.229)	0.015 (0.354)	0.047** (0.006)
Own a Mobile Phone	-0.002 (0.637)	-0.002 (0.769)	-0.004 (0.198)	-0.001 (0.894)	0.027* (0.031)
Highest Completed Grade of Education in Household	0.006** (0.000)	0.006** (0.000)	0.004** (0.000)	0.009** (0.000)	0.002 (0.241)
Age of Household Head	0.000 (0.388)	0.000 (0.367)	0.000 (0.909)	0.000 (0.828)	-0.001* (0.018)
Female Head of Household	0.013 (0.104)	0.012 (0.260)	0.015 (0.076)	-0.008 (0.533)	-0.001 (0.939)
Lean Season dummy variable	-0.007* (0.033)	-0.007 (0.149)	-0.008** (0.010)	-0.001 (0.893)	-0.009 (0.229)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-3: Second stage of control function regression; estimating the share of low processed food in total food consumption

Variable	Share of Low Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	-0.004** (0.000)	-0.005** (0.000)	-0.002 (0.212)	-0.020 (0.848)	-0.005 (0.916)
Household shops at a supermarket in Lusaka	-0.091 (0.425)	-0.062 (0.609)			
Household shops at a supermarket in Kitwe	-0.195 (0.124)		0.127 (0.634)		
Household shops at a supermarket in Kasama	-0.055 (0.656)			2.816** (0.002)	
Household shops at a supermarket in Mansa	0.282** (0.010)				0.834** (0.010)
Lusaka dummy variable	0.042** (0.001)				
Kitwe dummy variable	0.057** (0.000)				
Kasama dummy variable	-0.008 (0.426)				
Percentage Working Age Females Employed in Nonfarm Labor	0.000 (0.999)	-0.002 (0.828)	-0.003 (0.798)	0.022 (0.098)	0.042** (0.006)
Dependency Ratio	0.016 (0.264)	0.024 (0.222)	0.009 (0.662)	-0.048 (0.166)	-0.049 (0.090)
Adult Equivalents	-0.002 (0.090)	-0.004* (0.046)	-0.001 (0.506)	0.013** (0.008)	0.005 (0.163)
Own a Car	0.038** (0.000)	0.045** (0.001)	0.024 (0.136)	-0.167** (0.000)	0.037* (0.031)
Own a Motorcycle	-0.018 (0.431)	-0.017 (0.521)	-0.019 (0.684)	-0.071 (0.243)	-0.050 (0.252)
Use Public Transit	0.010 (0.142)	0.020* (0.028)	0.003 (0.734)	-0.032** (0.003)	0.017 (0.132)
Own a Bicycle	0.013 (0.070)	0.015 (0.174)	0.013 (0.183)	0.030 (0.053)	0.001 (0.909)
Own Farmland	-0.025** (0.000)	-0.012 (0.305)	-0.046** (0.000)	0.002 (0.878)	-0.026* (0.048)
Own a Stove (Gas or Electric)	0.023** (0.011)	0.017 (0.168)	0.039** (0.001)	-0.086* (0.019)	-0.043 (0.100)
Own a Hotplate	0.001 (0.884)	0.000 (0.974)	0.010 (0.441)	-0.038* (0.037)	-0.022 (0.238)
Own a Mobile Phone	0.009 (0.270)	0.009 (0.458)	-0.001 (0.941)	0.030 (0.078)	0.026 (0.127)
Highest Completed Grade of Education in Household	0.002 (0.240)	0.001 (0.496)	0.004 (0.057)	-0.020* (0.016)	0.000 (0.926)
Age of Household Head	0.000 (0.179)	0.000 (0.449)	0.001** (0.000)	-0.001 (0.057)	0.000 (0.554)
Female Head of Household	-0.002 (0.843)	-0.001 (0.961)	-0.009 (0.504)	0.008 (0.698)	0.006 (0.654)
Lean Season dummy variable	0.037** (0.000)	0.037** (0.000)	0.033** (0.000)	0.037** (0.000)	0.071** (0.000)
Endogenous Error from 1st Stage Regression	0.102 (0.376)	0.161 (0.203)	-0.053 (0.847)	-2.639** (0.003)	-0.545 (0.089)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-4: Second stage of control function regression; estimating the share of high processed food in total food consumption

Variable	Share of High Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.014** (0.000)	0.013** (0.000)	0.014** (0.000)	0.025 (0.454)	0.027 (0.249)
Household shops at a supermarket in Lusaka	0.098 (0.442)	0.082 (0.565)			
Household shops at a supermarket in Kitwe	0.096 (0.486)		0.205 (0.484)		
Household shops at a supermarket in Kasama	0.260 (0.047)			-0.148 (0.832)	
Household shops at a supermarket in Mansa	0.090 (0.518)				-0.145 (0.768)
Lusaka dummy variable	-0.005 (0.726)				
Kitwe dummy variable	-0.014 (0.342)				
Kasama dummy variable	-0.050** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.008 (0.310)	0.002 (0.867)	0.024* (0.024)	-0.002 (0.845)	0.003 (0.882)
Dependency Ratio	-0.022 (0.167)	-0.029 (0.193)	0.003 (0.884)	0.002 (0.959)	0.000 (0.996)
Adult Equivalents	0.008** (0.000)	0.010** (0.000)	0.007** (0.005)	0.002 (0.621)	0.003 (0.446)
Own a Car	-0.046** (0.000)	-0.052** (0.000)	-0.023 (0.173)	-0.007 (0.898)	-0.050** (0.005)
Own a Motorcycle	0.020 (0.536)	0.046 (0.363)	-0.039 (0.262)	-0.016 (0.800)	0.017 (0.701)
Use Public Transit	-0.003 (0.695)	-0.006 (0.552)	0.000 (0.970)	0.018 (0.091)	-0.022 (0.143)
Own a Bicycle	-0.018* (0.017)	-0.025* (0.018)	0.000 (0.968)	-0.016 (0.215)	-0.046** (0.005)
Own Farmland	-0.011 (0.190)	-0.026 (0.093)	0.011 (0.371)	-0.028 (0.019)	-0.011 (0.549)
Own a Stove (Gas or Electric)	-0.028** (0.002)	-0.031* (0.017)	-0.022 (0.123)	-0.017 (0.609)	0.027 (0.542)
Own a Hotplate	-0.016 (0.052)	-0.019 (0.070)	-0.013 (0.321)	-0.010 (0.605)	0.036 (0.234)
Own a Mobile Phone	0.006 (0.486)	0.005 (0.710)	0.019 (0.143)	0.010 (0.486)	-0.015 (0.510)
Highest Completed Grade of Education in Household	-0.001 (0.598)	0.001 (0.744)	-0.005* (0.043)	0.001 (0.846)	-0.004 (0.194)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.001** (0.004)	-0.002** (0.003)
Female Head of Household	-0.009 (0.310)	-0.006 (0.611)	-0.012 (0.386)	-0.020 (0.138)	-0.039* (0.031)
Lean Season dummy variable	-0.042** (0.000)	-0.041** (0.000)	-0.043** (0.000)	-0.037** (0.000)	-0.036** (0.005)
Endogenous Error from 1st Stage Regression	-0.262* (0.047)	-0.271 (0.072)	-0.326 (0.272)	0.174 (0.805)	-0.031 (0.950)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-5: Estimates of the marginal effects of the determinants of processed food in total food consumption (OLS)

Variable	Share of Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.011** (0.000)	0.009** (0.000)	0.013** (0.000)	0.024** (0.000)	0.025** (0.000)
Household shops at a supermarket in Lusaka	-0.129** (0.000)	-0.090** (0.000)			
Household shops at a supermarket in Kitwe	-0.214** (0.000)		-0.035 (0.316)		
Household shops at a supermarket in Kasama	0.058 (0.314)			0.151* (0.011)	
Household shops at a supermarket in Mansa	0.132** (0.000)				0.059 (0.272)
Lusaka dummy variable	0.008 (0.273)				
Kitwe dummy variable	0.011 (0.116)				
Kasama dummy variable	-0.057** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.009 (0.072)	-0.006 (0.458)	0.007 (0.385)	0.011 (0.399)	0.030** (0.049)
Dependency Ratio	-0.010 (0.303)	-0.005 (0.756)	0.017 (0.286)	-0.024 (0.386)	-0.044 (0.138)
Adult Equivalents	0.008** (0.000)	0.006** (0.001)	0.006** (0.001)	0.006 (0.070)	0.013** (0.000)
Own a Car	-0.009 (0.238)	-0.001 (0.958)	-0.008 (0.523)	-0.016 (0.582)	-0.024 (0.326)
Own a Motorcycle	-0.018 (0.445)	0.019 (0.672)	-0.044 (0.224)	0.029 (0.662)	-0.053 (0.360)
Use Public Transit	0.000 (0.927)	0.018* (0.014)	0.001 (0.939)	-0.003 (0.792)	-0.021 (0.100)
Own a Bicycle	-0.014** (0.007)	-0.014 (0.133)	0.000 (0.980)	-0.008 (0.505)	-0.045** (0.001)
Own Farmland	-0.039** (0.000)	-0.035** (0.004)	-0.029** (0.001)	-0.048** (0.000)	-0.046** (0.000)
Own a Stove (Gas or Electric)	0.004 (0.530)	-0.010 (0.300)	0.022* (0.013)	0.008 (0.641)	0.020 (0.264)
Own a Hotplate	-0.002 (0.664)	-0.009 (0.262)	-0.007 (0.435)	-0.007 (0.690)	0.038* (0.024)
Own a Mobile Phone	0.013* (0.026)	0.002 (0.858)	0.011 (0.215)	0.029 (0.051)	0.025 (0.120)
Highest Completed Grade of Education in Household	0.001 (0.364)	0.002 (0.083)	-0.001 (0.450)	0.003 (0.255)	-0.001 (0.843)
Age of Household Head	-0.001** (0.000)	-0.001** (0.000)	-0.001* (0.047)	-0.002** (0.000)	-0.002** (0.000)
Female Head of Household	-0.018** (0.000)	-0.012 (0.160)	-0.018* (0.040)	-0.035* (0.015)	-0.021 (0.157)
Lean Season dummy variable	-0.006 (0.115)	-0.014* (0.027)	-0.012 (0.057)	-0.003 (0.803)	0.024* (0.042)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-6: Estimates of the marginal effects of the determinants of low processed food in total food consumption (OLS)

Variable	Share of Low Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	-0.004** (0.000)	-0.005** (0.000)	-0.002 (0.117)	0.001 (0.667)	-0.002 (0.568)
Household shops at a supermarket in Lusaka	0.014 (0.612)	0.090** (0.000)			
Household shops at a supermarket in Kitwe	-0.072 (0.102)		0.096** (0.006)		
Household shops at a supermarket in Kasama	0.024 (0.648)			0.170** (0.001)	
Household shops at a supermarket in Mansa	0.319** (0.000)				0.298** (0.000)
Lusaka dummy variable	0.038** (0.000)				
Kitwe dummy variable	0.056** (0.000)				
Kasama dummy variable	-0.004 (0.579)				
Percentage Working Age Females Employed in Nonfarm Labor	0.006 (0.177)	0.000 (0.953)	-0.002 (0.770)	0.015 (0.193)	0.029* (0.017)
Dependency Ratio	-0.007 (0.436)	0.006 (0.673)	0.009 (0.597)	-0.024 (0.310)	-0.057* (0.016)
Adult Equivalents	0.001 (0.217)	-0.004* (0.024)	0.000 (0.874)	0.004 (0.186)	0.007* (0.014)
Own a Car	0.024** (0.001)	0.030** (0.006)	0.019 (0.108)	0.012 (0.625)	0.046* (0.018)
Own a Motorcycle	-0.003 (0.886)	-0.001 (0.985)	0.019 (0.608)	0.038 (0.511)	-0.064 (0.162)
Use Public Transit	-0.001 (0.858)	0.009 (0.246)	0.002 (0.776)	-0.023* (0.024)	0.008 (0.427)
Own a Bicycle	0.006 (0.184)	0.009 (0.335)	0.003 (0.742)	-0.004 (0.661)	0.011 (0.285)
Own Farmland	-0.028** (0.000)	-0.007 (0.571)	-0.043** (0.000)	-0.017 (0.072)	-0.033** (0.001)
Own a Stove (Gas or Electric)	0.019** (0.000)	0.023* (0.017)	0.037** (0.000)	0.023 (0.118)	0.000 (0.972)
Own a Hotplate	0.003 (0.519)	0.005 (0.496)	0.011 (0.235)	0.007 (0.677)	0.010 (0.465)
Own a Mobile Phone	0.012* (0.022)	0.004 (0.672)	-0.001 (0.914)	0.022 (0.077)	0.034** (0.008)
Highest Completed Grade of Education in Household	0.001 (0.482)	-0.001 (0.511)	0.002 (0.275)	0.002 (0.331)	0.001 (0.796)
Age of Household Head	0.000 (0.136)	0.001 (0.064)	0.001** (0.000)	-0.001* (0.050)	0.000 (0.533)
Female Head of Household	0.001 (0.884)	0.012 (0.131)	-0.012 (0.162)	-0.028* (0.021)	0.021 (0.070)
Lean Season dummy variable	0.042** (0.000)	0.037** (0.000)	0.040** (0.000)	0.037** (0.000)	0.066** (0.000)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-7: Estimates of the marginal effects of the determinants of high processed food in total food consumption (OLS)

Variable	Share of High Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.015** (0.000)	0.015** (0.000)	0.015** (0.000)	0.023** (0.000)	0.026** (0.000)
Household shops at a supermarket in Lusaka	-0.143** (0.000)	-0.181** (0.000)			
Household shops at a supermarket in Kitwe	-0.143** (0.007)		-0.131** (0.002)		
Household shops at a supermarket in Kasama	0.034 (0.597)			-0.018 (0.723)	
Household shops at a supermarket in Mansa	-0.188** (0.000)				-0.239** (0.000)
Lusaka dummy variable	-0.030** (0.000)				
Kitwe dummy variable	-0.045** (0.000)				
Kasama dummy variable	-0.054** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.003 (0.621)	-0.006 (0.578)	0.010 (0.329)	-0.004 (0.755)	0.001 (0.936)
Dependency Ratio	-0.003 (0.780)	-0.011 (0.554)	0.009 (0.651)	0.000 (0.996)	0.012 (0.692)
Adult Equivalents	0.007** (0.000)	0.011** (0.000)	0.006** (0.002)	0.002 (0.435)	0.006 (0.081)
Own a Car	-0.033** (0.000)	-0.031* (0.023)	-0.027 (0.058)	-0.029 (0.268)	-0.070** (0.006)
Own a Motorcycle	-0.015 (0.568)	0.019 (0.719)	-0.063 (0.146)	-0.008 (0.886)	0.011 (0.855)
Use Public Transit	0.000 (0.946)	0.010 (0.286)	-0.002 (0.860)	0.020 (0.056)	-0.029* (0.028)
Own a Bicycle	-0.020** (0.000)	-0.024* (0.046)	-0.003 (0.765)	-0.003 (0.740)	-0.055** (0.000)
Own Farmland	-0.011 (0.062)	-0.028 (0.063)	0.013 (0.197)	-0.032 (0.001)	-0.013 (0.316)
Own a Stove (Gas or Electric)	-0.016* (0.015)	-0.033** (0.006)	-0.015 (0.168)	-0.015 (0.317)	0.019 (0.292)
Own a Hotplate	-0.006 (0.351)	-0.014 (0.144)	-0.019 (0.098)	-0.014 (0.387)	0.028 (0.103)
Own a Mobile Phone	0.001 (0.922)	-0.002 (0.844)	0.012 (0.257)	0.007 (0.614)	-0.009 (0.611)
Highest Completed Grade of Education in Household	0.000 (0.816)	0.003 (0.053)	-0.003 (0.120)	0.001 (0.727)	-0.001 (0.697)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)
Female Head of Household	-0.019** (0.001)	-0.024 (0.018)	-0.006 (0.582)	-0.006 (0.611)	-0.042** (0.006)
Lean Season dummy variable	-0.049** (0.000)	-0.051** (0.000)	-0.052** (0.000)	-0.039** (0.000)	-0.043** (0.000)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-8: Second stage of two stage least squares; estimating the share of processed food in total food consumption (2SLS)

Variable	Share of Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.010** (0.000)	0.009** (0.000)	0.012** (0.000)	0.003 (0.981)	0.020 (0.762)
Household shops at a supermarket in Lusaka	-0.032 (0.791)	0.022 (0.853)			
Household shops at a supermarket in Kitwe	-0.179 (0.312)		0.343 (0.139)		
Household shops at a supermarket in Kasama	0.275 (0.188)			2.833** (0.004)	
Household shops at a supermarket in Mansa	0.575** (0.001)				0.825* (0.036)
Lusaka dummy variable	0.042** (0.010)				
Kitwe dummy variable	0.047** (0.003)				
Kasama dummy variable	-0.047** (0.001)				
Percentage Working Age Females Employed in Nonfarm Labor	0.007 (0.301)	-0.001 (0.898)	0.022* (0.046)	0.017 (0.289)	0.047* (0.016)
Dependency Ratio	-0.008 (0.633)	-0.005 (0.813)	0.014 (0.486)	-0.048 (0.221)	-0.041 (0.279)
Adult Equivalents	0.006** (0.003)	0.006* (0.025)	0.005* (0.025)	0.015* (0.014)	0.007 (0.072)
Own a Car	-0.015 (0.119)	-0.013 (0.296)	-0.005 (0.708)	-0.222** (0.008)	-0.018 (0.352)
Own a Motorcycle	0.007 (0.766)	0.031 (0.393)	-0.064* (0.021)	-0.101 (0.062)	-0.032 (0.300)
Use Public Transit	0.008 (0.274)	0.014 (0.164)	0.003 (0.765)	-0.013 (0.396)	-0.006 (0.698)
Own a Bicycle	-0.005 (0.490)	-0.010 (0.366)	0.013 (0.232)	0.014 (0.380)	-0.046** (0.002)
Own Farmland	-0.037** (0.000)	-0.040** (0.006)	-0.035** (0.003)	-0.023 (0.139)	-0.034 (0.052)
Own a Stove (Gas or Electric)	-0.004 (0.631)	-0.014 (0.284)	0.018 (0.150)	-0.118* (0.023)	-0.027 (0.444)
Own a Hotplate	-0.013** (0.090)	-0.018 (0.060)	-0.002 (0.859)	-0.051 (0.064)	0.012 (0.627)
Own a Mobile Phone	0.015 (0.096)	0.012 (0.310)	0.018 (0.182)	0.039 (0.052)	0.013 (0.590)
Highest Completed Grade of Education in Household	0.001 (0.709)	0.002 (0.455)	-0.002 (0.450)	-0.020* (0.018)	-0.005 (0.140)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.001 (0.075)	-0.002** (0.000)	-0.002* (0.017)
Female Head of Household	-0.012 (0.161)	-0.007 (0.543)	-0.022 (0.083)	-0.010 (0.690)	-0.031 (0.111)
Lean Season dummy variable	-0.006 (0.304)	-0.005 (0.496)	-0.011 (0.145)	0.000 (0.992)	0.036** (0.009)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-9: Second stage of two stage least squares; estimating the share of low processed food in total food consumption (2SLS)

Variable	Share of Low Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	-0.004** (0.000)	-0.005** (0.000)	-0.002 (0.186)	-0.022 (0.847)	-0.006 (0.902)
Household shops at a supermarket in Lusaka	-0.171 (0.163)	-0.060 (0.616)			
Household shops at a supermarket in Kitwe	-0.351* (0.045)		0.124 (0.647)		
Household shops at a supermarket in Kasama	0.049 (0.796)			2.996** (0.001)	
Household shops at a supermarket in Mansa	0.729** (0.000)				0.885** (0.006)
Lusaka dummy variable	0.059** (0.000)				
Kitwe dummy variable	0.074** (0.000)				
Kasama dummy variable	0.005 (0.662)				
Percentage Working Age Females Employed in Nonfarm Labor	0.001 (0.929)	-0.002 (0.831)	-0.002 (0.814)	0.020 (0.128)	0.043* (0.009)
Dependency Ratio	0.014 (0.337)	0.024 (0.230)	0.009 (0.652)	-0.049 (0.151)	-0.048 (0.122)
Adult Equivalents	-0.003 (0.055)	-0.004* (0.047)	-0.001 (0.506)	0.014** (0.005)	0.005 (0.234)
Own a Car	0.035** (0.001)	0.043** (0.001)	0.022 (0.165)	-0.217** (0.004)	0.039* (0.037)
Own a Motorcycle	-0.014 (0.538)	-0.017 (0.543)	-0.021 (0.659)	-0.082 (0.250)	-0.054 (0.300)
Use Public Transit	0.011 (0.096)	0.020* (0.030)	0.004 (0.721)	-0.032** (0.004)	0.018 (0.130)
Own a Bicycle	0.013 (0.077)	0.015 (0.182)	0.013 (0.176)	0.031* (0.044)	0.001 (0.933)
Own Farmland	-0.025** (0.000)	-0.012 (0.302)	-0.046** (0.000)	0.003 (0.830)	-0.026 (0.060)
Own a Stove (Gas or Electric)	0.022* (0.013)	0.017 (0.185)	0.040** (0.001)	-0.097* (0.025)	-0.045 (0.110)
Own a Hotplate	0.001 (0.847)	0.000 (0.976)	0.011 (0.422)	-0.042* (0.039)	-0.023 (0.225)
Own a Mobile Phone	0.009 (0.278)	0.009 (0.460)	-0.001 (0.966)	0.030 (0.081)	0.026 (0.158)
Highest Completed Grade of Education in Household	0.001 (0.343)	0.001 (0.517)	0.004 (0.064)	-0.021** (0.008)	-0.001 (0.852)
Age of Household Head	0.000 (0.151)	0.000 (0.462)	0.001** (0.000)	-0.001 (0.070)	0.000 (0.505)
Female Head of Household	-0.003 (0.700)	-0.001 (0.950)	-0.009 (0.493)	0.010 (0.635)	0.006 (0.661)
Lean Season dummy variable	0.037** (0.000)	0.037** (0.000)	0.033** (0.000)	0.037** (0.000)	0.072** (0.000)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-10: Second stage of two stage least squares; estimating the share of high processed food in total food consumption (2SLS)

Variable	Share of High Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.014** (0.000)	0.013** (0.000)	0.014** (0.000)	0.025 (0.490)	0.026 (0.264)
Household shops at a supermarket in Lusaka	0.139 (0.376)	0.082 (0.605)			
Household shops at a supermarket in Kitwe	0.172 (0.422)		0.220 (0.456)		
Household shops at a supermarket in Kasama	0.227 (0.308)			-0.163 (0.842)	
Household shops at a supermarket in Mansa	-0.154 (0.421)				-0.061 (0.903)
Lusaka dummy variable	-0.017 (0.295)				
Kitwe dummy variable	-0.027 (0.098)				
Kasama dummy variable	-0.052** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.007 (0.403)	0.001 (0.940)	0.024* (0.030)	-0.003 (0.814)	0.004 (0.867)
Dependency Ratio	-0.022 (0.191)	-0.029 (0.194)	0.005 (0.817)	0.001 (0.965)	0.007 (0.860)
Adult Equivalents	0.008** (0.000)	0.010** (0.000)	0.006** (0.005)	0.001 (0.775)	0.002 (0.586)
Own a Car	-0.050** (0.000)	-0.056** (0.001)	-0.026 (0.162)	-0.006 (0.933)	-0.058** (0.009)
Own a Motorcycle	0.021 (0.573)	0.048 (0.397)	-0.043 (0.303)	-0.019 (0.778)	0.022 (0.655)
Use Public Transit	-0.003 (0.659)	-0.006 (0.540)	0.000 (0.972)	0.019 (0.087)	-0.024 (0.136)
Own a Bicycle	-0.018* (0.021)	-0.025* (0.029)	0.000 (0.989)	-0.017 (0.220)	-0.048** (0.005)
Own Farmland	-0.012 (0.169)	-0.028 (0.095)	0.012 (0.374)	-0.026* (0.037)	-0.008 (0.668)
Own a Stove (Gas or Electric)	-0.026** (0.007)	-0.030* (0.028)	-0.022 (0.132)	-0.021 (0.603)	0.019 (0.669)
Own a Hotplate	-0.014 (0.088)	-0.019 (0.082)	-0.013 (0.349)	-0.009 (0.680)	0.035 (0.241)
Own a Mobile Phone	0.006 (0.536)	0.004 (0.769)	0.019 (0.151)	0.009 (0.499)	-0.014 (0.544)
Highest Completed Grade of Education in Household	-0.001 (0.615)	0.000 (0.862)	-0.005* (0.034)	0.001 (0.843)	-0.004 (0.204)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.001** (0.005)	-0.002** (0.004)
Female Head of Household	-0.009 (0.335)	-0.006 (0.622)	-0.013 (0.351)	-0.020 (0.172)	-0.038* (0.044)
Lean Season dummy variable	-0.043** (0.000)	-0.042** (0.000)	-0.044** (0.000)	-0.037** (0.000)	-0.036** (0.007)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-11: Fractional probit estimates of the determinants of household share of supermarket purchases in total food consumption (CRE)

Variable	Share Supermarket in Total Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.005** (0.000)	0.006** (0.000)	0.003** (0.000)	0.010** (0.000)	0.015** (0.000)
Lusaka dummy variable	0.004 (0.230)				
Kitwe dummy variable	-0.079** (0.000)				
Kasama dummy variable	-0.051** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	-0.003 (0.367)	0.011 (0.449)	0.004 (0.658)	0.013 (0.433)	-0.006 (0.675)
Dependency Ratio	0.000 (0.210)	0.036* (0.032)	-0.003 (0.748)	0.014 (0.413)	0.011 (0.684)
Adult Equivalents	0.015 (0.056)	0.004 (0.151)	-0.001 (0.546)	-0.002 (0.373)	0.005 (0.081)
Relative Distance to Supermarket in Lusaka (km) (log)	-0.008** (0.000)	-0.010** (0.000)			
Relative Distance to Supermarket in Kitwe (km) (log)	-0.004** (0.000)		-0.005** (0.000)		
Relative Distance to Supermarket in Kasama (km) (log)	-0.005** (0.009)			-0.002 (0.338)	
Relative Distance to Supermarket in Mansa (km) (log)	-0.007 (0.059)				-0.006* (0.027)
Own a Car	0.002 (0.264)	0.006 (0.693)	-0.005 (0.485)	0.048 (0.087)	-0.042* (0.023)
Own a Motorcycle	0.003 (0.821)	0.035 (0.218)	-0.003 (0.901)	0.001 (0.945)	-0.048 (0.054)
Use Public Transit	0.017 (0.295)	0.007 (0.456)	-0.007 (0.324)	-0.002 (0.814)	-0.008 (0.456)
Own a Bicycle	0.001 (0.860)	0.002 (0.892)	0.009 (0.317)	0.001 (0.960)	-0.019 (0.307)
Own Farmland	0.001 (0.922)	-0.014 (0.104)	0.004 (0.578)	0.008 (0.426)	-0.014 (0.384)
Own a Stove (Gas or Electric)	0.002 (0.810)	0.013 (0.496)	0.002 (0.842)	0.018 (0.243)	0.012 (0.537)
Own a Hotplate	0.011 (0.386)	0.010 (0.399)	0.004 (0.624)	-0.011 (0.364)	0.028 (0.131)
Own a Mobile Phone	0.009 (0.289)	0.000 (0.995)	0.012 (0.176)	-0.015 (0.393)	0.006 (0.722)
Highest Completed Grade of Education in Household	0.010 (0.333)	0.006** (0.003)	0.002 (0.468)	0.000 (0.898)	-0.009* (0.033)
Age of Household Head	0.004** (0.007)	0.000 (0.206)	0.000 (0.816)	0.000 (0.350)	-0.001 (0.053)
Female Head of Household	0.024 (0.052)	0.016 (0.117)	0.016* (0.046)	-0.006 (0.554)	0.002 (0.919)
Lean Season dummy variable	-0.017** (0.000)	-0.001 (0.837)	-0.006 (0.053)	-0.001 (0.819)	-0.001 (0.890)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-12: First stage of control function regression; OLS estimates of the determinants of the share of supermarket purchases in total food consumption (CRE)

Variable	Share Supermarket in Total Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.005** (0.000)	0.006** (0.000)	0.003** (0.000)	0.010** (0.000)	0.015** (0.000)
Lusaka dummy variable	-0.050** (0.000)				
Kitwe dummy variable	-0.050** (0.000)				
Kasama dummy variable	-0.020** (0.003)				
Percentage Working Age Females Employed in Nonfarm Labor	0.008 (0.359)	0.009 (0.492)	0.003 (0.696)	0.008 (0.494)	-0.006 (0.628)
Dependency Ratio	0.027* (0.033)	0.040* (0.025)	-0.002 (0.820)	0.016 (0.390)	0.022 (0.412)
Adult Equivalents	0.001 (0.651)	0.002 (0.549)	-0.001 (0.492)	-0.004 (0.091)	0.002 (0.421)
Relative Distance to Supermarket in Lusaka (km) (log)	-0.019** (0.000)	-0.021** (0.000)			
Relative Distance to Supermarket in Kitwe (km) (log)	-0.011* (0.017)		-0.008** (0.000)		
Relative Distance to Supermarket in Kasama (km) (log)	-0.014 (0.054)			-0.004 (0.441)	
Relative Distance to Supermarket in Mansa (km) (log)	-0.014** (0.003)				-0.007* (0.050)
Own a Car	0.005 (0.826)	0.012 (0.675)	-0.016 (0.349)	0.049 (0.276)	-0.077 (0.078)
Own a Motorcycle	0.033 (0.131)	0.055 (0.097)	0.011 (0.778)	-0.010 (0.865)	-0.074 (0.276)
Use Public Transit	0.000 (0.966)	0.007 (0.547)	-0.007 (0.185)	0.002 (0.841)	-0.007 (0.513)
Own a Bicycle	-0.001 (0.964)	-0.004 (0.859)	0.010 (0.194)	0.006 (0.736)	-0.021 (0.265)
Own Farmland	-0.016 (0.078)	-0.030 (0.061)	0.001 (0.882)	0.002 (0.797)	-0.006 (0.682)
Own a Stove (Gas or Electric)	0.016 (0.295)	0.018 (0.430)	0.005 (0.568)	0.025 (0.317)	0.019 (0.516)
Own a Hotplate	0.012 (0.326)	0.013 (0.413)	0.007 (0.468)	-0.021 (0.318)	0.039 (0.148)
Own a Mobile Phone	0.000 (0.926)	-0.001 (0.819)	0.006 (0.350)	-0.012 (0.274)	0.007 (0.520)
Highest Completed Grade of Education in Household	0.004* (0.017)	0.005* (0.013)	0.001 (0.648)	-0.001 (0.785)	-0.008 (0.058)
Age of Household Head	0.000 (0.296)	0.000 (0.248)	0.000 (0.697)	0.000 (0.964)	-0.001* (0.039)
Female Head of Household	0.014 (0.073)	0.014 (0.193)	0.014 (0.110)	-0.009 (0.465)	0.000 (1.000)
Lean Season dummy variable	-0.004 (0.291)	-0.002 (0.683)	-0.005 (0.124)	0.000 (0.932)	-0.003 (0.702)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-13: Second stage of control function regression; estimating the share of processed food in total food consumption (CRE)

Variable	Share of Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.010** (0.000)	0.009** (0.000)	0.011** (0.000)	0.006 (0.996)	0.021 (0.646)
Household shops at a supermarket in Lusaka	0.015 (0.890)	0.028 (0.817)			
Household shops at a supermarket in Kitwe	-0.085 (0.470)		0.388 (0.115)		
Household shops at a supermarket in Kasama	0.164 (0.216)			2.892** (0.008)	
Household shops at a supermarket in Mansa	0.409** (0.001)				0.796 (0.068)
Lusaka dummy variable	0.025 (0.080)				
Kitwe dummy variable	0.030* (0.021)				
Kasama dummy variable	-0.048** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	0.013* (0.013)	0.002 (0.890)	0.043* (0.024)	-0.028 (0.238)	0.052 (0.187)
Dependency Ratio	-0.009 (0.574)	-0.010 (0.660)	0.010 (0.634)	-0.048 (0.233)	-0.036 (0.331)
Adult Equivalents	0.006** (0.001)	0.006* (0.014)	0.005* (0.041)	0.018** (0.004)	0.007 (0.060)
Own a Car	-0.026 (0.122)	-0.037 (0.074)	0.014 (0.497)	-0.019 (0.861)	-0.006 (0.931)
Own a Motorcycle	-0.019 (0.600)	0.002 (0.971)	-0.071 (0.277)	-0.003 (0.939)	-0.075 (0.444)
Use Public Transit	-0.006 (0.559)	-0.003 (0.858)	-0.001 (0.956)	-0.031* (0.044)	-0.013 (0.531)
Own a Bicycle	-0.011 (0.496)	-0.028 (0.238)	0.004 (0.827)	0.032 (0.173)	0.021 (0.504)
Own Farmland	0.013 (0.398)	0.019 (0.475)	0.006 (0.722)	-0.003 (0.884)	-0.034 (0.224)
Own a Stove (Gas or Electric)	0.006 (0.730)	-0.002 (0.916)	0.024 (0.217)	-0.041 (0.379)	0.026 (0.571)
Own a Hotplate	-0.014 (0.336)	-0.019 (0.316)	0.006 (0.783)	0.020 (0.652)	-0.029 (0.233)
Own a Mobile Phone	0.018 (0.231)	0.028 (0.187)	0.004 (0.818)	0.012 (0.689)	-0.013 (0.694)
Highest Completed Grade of Education in Household	0.005 (0.103)	0.006 (0.149)	0.005 (0.234)	0.001 (0.825)	0.003 (0.698)
Age of Household Head	-0.001** (0.000)	-0.002** (0.000)	-0.001 (0.115)	-0.002** (0.000)	-0.002* (0.023)
Female Head of Household	-0.011 (0.195)	-0.006 (0.587)	-0.022 (0.079)	-0.007 (0.773)	-0.038 (0.047)
Lean Season dummy variable	-0.005 (0.380)	-0.004 (0.574)	-0.012 (0.093)	0.003 (0.742)	0.041** (0.002)
Endogenous Error from 1st Stage Regression	-0.186 (0.094)	-0.134 (0.280)	-0.450 (0.072)	-2.695* (0.014)	-0.679 (0.114)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-14: Second stage of control function regression; estimating the share of low processed food in total food consumption (CRE)

Variable	Share of Low Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	-0.004** (0.000)	-0.004** (0.002)	-0.003 (0.214)	-0.015 (0.983)	-0.004 (0.919)
Household shops at a supermarket in Lusaka	-0.099 (0.398)	-0.074 (0.555)			
Household shops at a supermarket in Kitwe	-0.205 (0.118)		0.096 (0.743)		
Household shops at a supermarket in Kasama	-0.069 (0.593)			2.589** (0.007)	
Household shops at a supermarket in Mansa	0.271 (0.018)				0.749* (0.026)
Lusaka dummy variable	0.037** (0.005)				
Kitwe dummy variable	0.053** (0.000)				
Kasama dummy variable	-0.008 (0.427)				
Percentage Working Age Females Employed in Nonfarm Labor	0.018* (0.018)	0.013 (0.483)	0.033 (0.078)	-0.014 (0.496)	0.052* (0.029)
Dependency Ratio	0.017 (0.248)	0.025 (0.213)	0.008 (0.684)	-0.046 (0.202)	-0.045 (0.128)
Adult Equivalents	-0.002 (0.109)	-0.004 (0.068)	-0.002 (0.309)	0.014** (0.009)	0.004 (0.281)
Own a Car	0.050** (0.003)	0.051* (0.017)	0.042 (0.147)	0.041 (0.565)	0.028 (0.675)
Own a Motorcycle	-0.020 (0.580)	-0.023 (0.582)	0.012 (0.882)	0.079 (0.089)	-0.050 (0.588)
Use Public Transit	0.010 (0.350)	0.027 (0.075)	-0.012 (0.486)	-0.031* (0.024)	-0.004 (0.812)
Own a Bicycle	0.014 (0.281)	0.015 (0.421)	0.006 (0.749)	0.006 (0.778)	-0.017 (0.442)
Own Farmland	-0.014 (0.325)	-0.005 (0.840)	-0.044* (0.029)	0.003 (0.843)	-0.027 (0.157)
Own a Stove (Gas or Electric)	-0.003 (0.836)	-0.002 (0.936)	-0.005 (0.828)	-0.083* (0.026)	-0.002 (0.958)
Own a Hotplate	-0.004 (0.792)	0.007 (0.702)	-0.042 (0.070)	0.044 (0.311)	-0.042 (0.066)
Own a Mobile Phone	-0.001 (0.968)	0.008 (0.689)	-0.019 (0.372)	0.054* (0.032)	-0.044* (0.024)
Highest Completed Grade of Education in Household	0.005 (0.097)	0.006 (0.129)	0.007 (0.089)	-0.003 (0.605)	0.005 (0.296)
Age of Household Head	0.000 (0.124)	0.000 (0.410)	0.001** (0.000)	-0.001* (0.035)	0.000 (0.452)
Female Head of Household	-0.001 (0.945)	0.000 (0.986)	-0.007 (0.575)	0.011 (0.619)	0.002 (0.900)
Lean Season dummy variable	0.041** (0.000)	0.042** (0.000)	0.034** (0.000)	0.040** (0.000)	0.075** (0.000)
Endogenous Error from 1st Stage Regression	0.111 (0.349)	0.173 (0.186)	-0.024 (0.935)	-2.416* (0.012)	-0.487 (0.143)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

Table 3-A-15: Second stage of control function regression; estimating the share of high processed food in total food consumption (CRE)

Variable	Share of High Processed Food				
	Pooled	Lusaka	Kitwe	Kasama	Mansa
Total Daily Household Expenditure per Adult Equivalent	0.014** (0.000)	0.013** (0.000)	0.014** (0.000)	0.022 (0.944)	0.028 (0.177)
Household shops at a supermarket in Lusaka	0.115 (0.383)	0.097 (0.515)			
Household shops at a supermarket in Kitwe	0.114 (0.428)		0.291 (0.345)		
Household shops at a supermarket in Kasama	0.271* (0.047)			0.055 (0.935)	
Household shops at a supermarket in Mansa	0.105 (0.473)				-0.096 (0.863)
Lusaka dummy variable	-0.011 (0.493)				
Kitwe dummy variable	-0.017 (0.271)				
Kasama dummy variable	-0.049** (0.000)				
Percentage Working Age Females Employed in Nonfarm Labor	-0.003** (0.003)	-0.010 (0.619)	0.011 (0.600)	-0.010 (0.615)	0.003 (0.926)
Dependency Ratio	-0.026 (0.111)	-0.034 (0.126)	-0.002 (0.940)	0.000 (0.997)	0.003 (0.947)
Adult Equivalents	0.009** (0.000)	0.009** (0.000)	0.007** (0.005)	0.004 (0.386)	0.005 (0.312)
Own a Car	-0.064** (0.000)	-0.076** (0.000)	-0.022 (0.469)	-0.049 (0.364)	-0.039 (0.464)
Own a Motorcycle	0.000 (0.993)	0.024 (0.740)	-0.069 (0.208)	-0.057 (0.289)	-0.033 (0.620)
Use Public Transit	-0.016 (0.169)	-0.030 (0.095)	0.011 (0.404)	-0.001 (0.963)	-0.010 (0.648)
Own a Bicycle	-0.024 (0.101)	-0.041* (0.049)	-0.004 (0.808)	0.025 (0.273)	0.031 (0.324)
Own Farmland	0.028 (0.080)	0.025 (0.364)	0.052** (0.004)	-0.003 (0.864)	-0.007 (0.811)
Own a Stove (Gas or Electric)	0.009 (0.652)	-0.001 (0.982)	0.029 (0.253)	0.047 (0.166)	0.033 (0.404)
Own a Hotplate	-0.009 (0.529)	-0.023 (0.213)	0.049* (0.049)	-0.025 (0.363)	0.017 (0.586)
Own a Mobile Phone	0.021 (0.189)	0.022 (0.308)	0.024 (0.223)	-0.044 (0.052)	0.033 (0.328)
Highest Completed Grade of Education in Household	0.001 (0.841)	0.001 (0.861)	-0.001 (0.794)	0.004 (0.489)	-0.003 (0.750)
Age of Household Head	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.001* (0.015)	-0.002** (0.004)
Female Head of Household	-0.010 (0.298)	-0.005 (0.667)	-0.014 (0.344)	-0.021 (0.131)	-0.039* (0.037)
Lean Season dummy variable	-0.044** (0.000)	-0.045** (0.000)	-0.046** (0.000)	-0.037 (0.000)	-0.036** (0.004)
Endogenous Error from 1st Stage Regression	-0.274* (0.046)	-0.281 (0.076)	-0.397 (0.203)	-0.053 (0.938)	-0.064 (0.907)

Source: Zambia Urban Consumption Survey 2007/2008

Note: Primary font indicates the estimated coefficient. Parentheses indicate the associated p-value. ** p < 0.01, * p < 0.05

REFERENCES

REFERENCES

- Asfaw, A. (2008). Does supermarket purchase affect the dietary practices of households? Some empirical evidence from Guatemala. *Development Policy Review*, 26(2), 227-243.
- Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *Review of Economics and statistics*, 79(4), 527-539.
- Battersby, J., & Peyton, S. (2014, June). The geography of supermarkets in Cape Town: Supermarket expansion and food access. In *Urban Forum* (Vol. 25, No. 2, pp. 153-164). Springer Netherlands.
- Berdegú, J. A., Carriazo, F., Jara, B., Modrego, F., & Soloaga, I. (2015). Cities, Territories, and Inclusive Growth: Unraveling Urban–Rural Linkages in Chile, Colombia, and Mexico. *World Development*.
- Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17), 7301-7306.
- Christiaensen, L., Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: the role of rural diversification and secondary towns¹. *Agricultural Economics*, 44(4-5), 435-447.
- D’Haese, M., & Van Huylenbroeck, G. (2005). The rise of supermarkets and changing expenditure patterns of poor rural households case study in the Transkei area, South Africa. *Food Policy*, 30(1), 97-113.
- Ecker, O., & Qaim, M. (2011). Analyzing nutritional impacts of policies: an empirical study for Malawi. *World Development*, 39(3), 412-428.
- Ferré, C., Ferreira, F. H., & Lanjouw, P. (2012). Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, lhs007.
- Goldman, A., Ramaswami, S., & Krider, R. E. (2002). Barriers to the advancement of modern food retail formats: theory and measurement. *Journal of Retailing*, 78(4), 281-295.
- Goldman, A., & Hino, H. (2005). Supermarkets vs. traditional retail stores: diagnosing the barriers to supermarkets’ market share growth in an ethnic minority community. *Journal of Retailing and Consumer Services*, 12(4), 273-284.
- Gómez, M. I., Barrett, C. B., Raney, T., Pinstруп-Andersen, P., Meerman, J., Croppenstedt, A., Carishma, B., & Thompson, B. (2013). Post-green revolution food systems and the triple burden of malnutrition. *Food Policy*, 42, 129-138.
- Gómez, M. I., & Ricketts, K. D. (2013). Food value chain transformations in developing countries: Selected hypotheses on nutritional implications. *Food Policy*, 42, 139-150.
- Harrison, K., Henley, D., Riley, H., & Shaffer, J. (1974). *Improving food marketing systems in developing countries: experiences from Latin America* (No. INVES-ET E14p H321). Michigan State Univ., East Lansing, Mich. (EUA). Latin American Studies Center.

- Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate behavioral research*, 45(4), 627-660.
- Hichaambwa, M., M. Beaver, M. Weber, and A. Chapoto. 2009. *Patterns of Urban Consumption and Expenditure in Zambia: A General Report Based on the CSO/MACO/FSRP Urban Consumption Survey in Urban Areas of Lusaka, Kitwe, Mansa, and Kasama, 2007-2008*. FSRP Draft Working Paper. Lusaka, Zambia: Food Security Research Project.
- Hino, H. (2014). Shopping at different food retail formats: Understanding cross-shopping behavior through retail format selective use patterns. *European Journal of Marketing*, 48(3/4), 674-698.
- 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(1), 45-56.
- Minten, B., Reardon, T., & Sutradhar, R. (2010). Food prices and modern retail: The case of Delhi. *World Development*, 38(12), 1775-1787.
- Moyer, R., & Hollander, S. C. (Eds.). (1968). *Markets and marketing in developing economies*. Homewood, IL: Richard D. Irwin, Inc.
- Neven, D., & Reardon, T. (2004). The rise of Kenyan supermarkets and the evolution of their horticulture product procurement systems. *Development policy review*, 22(6), 669-699.
- Neven, D., Reardon, T., Chege, J., & Wang, H. (2006). Supermarkets and consumers in Africa: the case of Nairobi, Kenya. *Journal of International Food & Agribusiness Marketing*, 18(1-2), 103-123.
- 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*, volume 11, 619-632.
- Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1), 121-133.
- Pingali, P. (2007). Westernization of Asian diets and the transformation of food systems: implications for research and policy. *Food policy*, 32(3), 281-298.
- Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition reviews*, 70(1), 3-21.
- Reardon, T., Timmer, C. P., Barrett, C. B., & Berdegue, J. (2003). The rise of supermarkets in Africa, Asia, and Latin America. *American journal of agricultural economics*, 85(5), 1140-1146.
- Reardon, T., Timmer, P., & Berdegue, J. (2004). The rapid rise of supermarkets in developing countries: induced organizational, institutional, and technological change in agrifood systems. *Electronic Journal of Agricultural and Development Economics*, 1(2), 168-183.
- Reardon, T., & Timmer, C. P. (2007). Transformation of markets for agricultural output in developing countries since 1950: How has thinking changed?. *Handbook of agricultural economics*, 3, 2807-2855.

- Reardon, T., Barrett, C. B., Berdegue, J. A., & Swinnen, J. F. (2009). Agrifood industry transformation and small farmers in developing countries. *World development*, 37(11), 1717-1727.
- Rischke, R., Kimenju, S. C., 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(4), 920-927.
- Smith, L. C. (Ed.). (2003). *The importance of women's status for child nutrition in developing countries* (Vol. 131). Intl Food Policy Res Inst.
- Smith, R., and R. Blundell (1986). An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. *Econometrica* 54, 679–685.
- Staiger, D. O., & Stock, J. H. (1994). Instrumental variables regression with weak instruments.
- Strydom, J. W. (2011). Retailing in disadvantaged communities: the outshopping phenomenon revisited. *Journal of Contemporary Management*, 8, 150-172.
- Thiele, S., & Weiss, C. (2003). Consumer demand for food diversity: evidence for Germany. *Food Policy*, 28(2), 99-115.
- Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015). The rise of a middle class in East and Southern Africa: Implications for food system transformation. *Journal of International Development*, 27(5), 628-646.
- Turrell, G., & Kavanagh, A. M. (2006). Socio-economic pathways to diet: modelling the association between socio-economic position and food purchasing behaviour. *Public health nutrition*, 9(03), 375-383.
- Veeck, A., & Veeck, G. (2000). Consumer segmentation and changing food purchase patterns in Nanjing, PRC. *World Development*, 28(3), 457-471.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- World Bank, (2009). World Development Report 2009: Reshaping Economic Geography, *World Bank*, Washington, DC.

CONCLUSIONS

These essays highlight multiple consumption patterns to show that the continuation of rising incomes, rapid urbanization and food retail modernization will induce growth and change within the agrifood system.

Essay one supports the hypotheses that with rising incomes households purchase greater shares of the food that they consume and follow the product transition as stated by Bennett's Law. I find that the surprisingly high levels of purchased shares of food in total household consumed food value across ESA (greater than 40% at any expenditure level in all four countries) are a result of households transitioning food consumption towards purchased food at subnumerary levels of income. This finding highlights the fact that transitions in food consumption patterns are not solely a middle-class story, as conventionally assumed, but poor households are accessing markets to purchase greater shares of their food and are consuming greater shares BSF foods as their incomes rise. Spatial considerations of increasing city size and reduced distance to cities have significant positive effects on purchased share, affirming the expectation that households will purchase more food (compared to consuming own production) when they have increased access to markets.

Essay two further highlights the heterogeneity in the effects of urbanization on food consumption patterns across income levels, city size and distances to cities, but with a focus on the consumption of processed foods. Processed foods have penetrated the diets of rural (36% of total food expenditure including consumed own production) and urban (63%) households at all levels across the income distribution of ESA. I find that income-induced diet change that begins among the poor contributes to these high levels of processed food consumption. The income-

induced transition towards processed food is driven by the consumption of highly processed food that begins to rise rapidly at subnumerary levels of income. Given the large portion of the population in ESA currently below the international poverty line, coupled with the pattern of increased processed share with income growth, these findings signal a strong future demand for increased food market infrastructure.

As anticipated, larger cities and lesser distance to cities result in greater consumption shares of processed food, but of note is the high consumption of processed food even by households who reside at considerable distances to urban areas.

Essay three incorporates food retail modernization into the food consumption pattern analysis, leading to the following three findings. First, supermarket use positively affects the consumption of processed food, which is expected given that on average eighty-five percent of the food value purchased from supermarkets is processed food. Second, and more surprising, conditional on the presence of a supermarket, households in smaller cities consume greater shares of food, by value, from supermarkets than households in larger cities. My conceptual model provides for this possibility by recognizing that congestion is greater in larger cities and that differences in investment cost of supermarkets and non-supermarkets could increase the relative density of non-supermarkets in larger cities, both of which would incentivize households to consume greater shares of their food from outlets other than supermarkets. Third, although supermarkets positively affect the consumption share of processed food, households that consume some food from supermarkets continue to purchase eighty percent of their processed food from non-supermarket retailers. The findings of this essay indicate that while food retail modernization within the agrifood system affects food consumption patterns, such as rising

consumption of processed food, these patterns are not solely dependent on food retail modernization.

Together, these three essays provide detail to how the macro trends of income growth, rapid urbanization and the beginning of food retail modernization significantly affect the consumption patterns of ESA. The fact that the diet transformation is happening among poor households, not just the minority above the poverty line, lends urgency to the need for improved public marketing infrastructure and regulatory approaches to accommodate and shape it.