32nd International Conference of Agricultural Economists
2-7 August 2024 I New Delhi I India

Healthiness of household food expenditure in urban and peri-urban Kenya: How much is explained by a spatial measure of the food environment

Ian Fisher[1], Mywish K. Maredia[2], David Tschirley[3]

1: Michigan State University Department of Agricultural, Food, and Resource Economics

Corresponding author email: fisheri6@msu.edu

Abstract

This paper addresses the pressing policy issue of food access and availability in low-income urban settings, particularly in the context of the nutrition transition, urbanization, and evolving food systems. By regressing food expenditure data against proximity-to-outlet measures for various outlet types, the study focuses on estimating distance elasticities—quantifying the responsiveness of household food shopping expenditure to variations in distances to different food outlets. The key finding underscores the significance of household location characteristics over average distance to outlets in predicting the healthiness of food purchases. The research further identifies variations in distance elasticities based on factors such as the main shopper's age, household poverty probability, and location. This study introduces a novel application of distance elasticity, paving the way for future investigations into food environment metrics within urban and peri-urban settings of low- and middle-income countries (LMICs). The insights gained aim to enhance the understanding of factors influencing food shopping behavior and guide strategies for promoting healthier food options through increased expenditures.

JEL Codes: Q180, Q000, C210

1: INTRODUCTION[1]

Access to healthy, nutritious food is crucial for economic and agricultural development. The Kenyan government has prioritized the availability and access to adequate, diverse, and healthy diets as part of its National Food and Nutrition Security Policy strategy document (Kenya Ministry of Agriculture 2011). Ongoing trends such as the nutrition transition (Popkin and Reardon 2018, Monteiro et al. 2013 & Tschirley et al. 2015), urbanization (Popkin 1999, Ruel et al. 2017, Hawkes, Harris, Gillespie 2017, & Crush, Battersby 2017), and food systems transformation (Reardon et al. 2017, Popkin and Reardon 2018, & Fraser et al. 2010) significantly impact food environments.

Access involves geographic relationships between households and food sources while availability is the adequacy of the supply of healthy food (Clapp et al. 2022 & Caspi et al. 2012). Food environments, defined as "the collective physical, economic, policy, and sociocultural surroundings, opportunities, and conditions that influence people's food and beverage choices and nutritional status" (Turner et al. 2018, Swinburn et al. 2013, Babey et al. 2018, & Wrigley 2002), have direct impacts on both the supply and demand dynamics of the food system (Global Panel 2017). Distance to food outlets, a key measurement of accessibility, is an important metric of determining healthiness of food environments (Athens, Duncan, Elbel 2016, Barnes et al. 2015, & Clapp et al. 2022).

Research in the areas of nutrition, economics, and health has highlighted the documented change in dietary behavior and food consumption leading to a rise in noncommunicable disease, particularly in low-income countries (Popkin 1999 and Popkin 2004). Environmental and spatial factors have been identified as key influencers on diet and nutrition, emphasizing the importance of studying the food environment. Distance to food outlets, a metric of accessibility, is critical in assessing the healthiness of food environments. Social, political, economic, and environmental forces, as well as a demographic and epidemiologic factors have been identified as key influencers on diet and nutrition (St-Onge, Keller, & Heymsfield, S.B. 2003, Popkin 1993,

[1] Research for this study is supported by the "Support for Applied Research and Analysis in Kenya and East Africa" (SARA-KEA) project with funding from USDA-FAS under grant No. FX21TA-10960R001 and FX22TA-10960R002. Data used for this research was collected with partial support from two grants- 1) the SARA-KEA project, and 2) the Feed the Future Innovation Lab for Food Security Policy, Research, Capacity, and Influence (PRCI) with funding from the United States Agency for International Development under grant No. 7200AA19LE00001.

Senauer et al. 1991). There is also a growing body of work on the role environmental/spatial factors have on the economics of diet and nutrition Glanz et al. 2005). Generally, the food environment has influence over what consumers purchase, emphasizing the importance of studying and measuring food environment (Herforth and Ahmed 2015 & Black, Moon, and Baird 2004, Rose et al. 2009, Farley et al. 2009, Tschirley et al. 2021) and showing evidence of effects of density of food outlets (Kruger et al. 2013 & Spence et al. 2009).

Existing literature presents a mixed perspective on the effectiveness of proximity or distance to food outlets as a determinant of shopping patterns and dietary behavior. Some studies suggest that supermarket proximity enhances access to a diverse diet, including fruits and vegetables but may also promote the consumption of energy-dense, highly processed foods (Hawkes 2008, Laraia et al. 2004, Pearson et al. 2005). Studies on grocery stores show associations between household proximity and nutrition, with greater availability of fresh vegetables near healthy food outlets linked to higher vegetable intake. Importantly, proximity to food outlets exhibits significant heterogeneity based on demographics and socioeconomic status (Zenk et al. 2008, Zenk et al. 2005, Black, Moon, Baird 2014, Sharkey et al. 2011, & Wang et al. 2007).

This research addresses gaps in existing knowledge by focusing on specific outlet types within the urban/peri-urban context of Kenya. It introduces distance elasticity, or the responsiveness of a measure to a change in distance, a term commonly used in trade literature (Schwartz 1973, Berthelon and Freund 2008, Bergstrand, Larch, and Yotoy 2015, Wang 2000, & Ebeke and Etoundi 2016) to food environment studies. In one of the few recent papers on the topic applied to FE distance metrics, it was found that for the developing country of Vietnam, a 10% decline in supermarket distance to a household increased consumer demand for supermarkets by about 7% (Mergenthaler, Weinberger, and Qaim 2009).

This research fills several knowledge gaps. First, it focuses on a more descriptive measure of linear outlet distance, disaggregating distances by outlet type. While some studies have looked at distance or proximity measures of the food environment (Laska et al. 2010; Rose and Richards 2004; Bodor et al., 2008), researchers have not studied how distances to specific outlet types, such as street vendors vs. supermarkets vs. vegetable sellers, etc., might affect the healthiness of food expenditures.

Second, this research applies the concept of distance elasticity to food environment research. Literature that has applied the distance elasticity concept has been in areas of measuring substitution patterns between stores (Chenarides and Jaenicke 2016) and analyzing demand systems from modern supply chains (Mergenthaler, Weinberger, and Qaim 2009). Few previous studies have looked at the margin of how additional distance to outlet affects healthiness of expenditure in the food environment context.

The main research questions of this research are:

1. How does proximity of household to the nearest outlet of each type affect food expenditure and significantly predict healthiness (or unhealthiness) of household food expenditure?
2. What, if any, socioeconomic, geographic, demographic group heterogeneity exist within the sample?

Based on the previous literature, this paper hypothesizes that healthiness of food offerings varies meaningfully across types of outlets and that distance to various outlet types meaningfully influences the healthiness of household food expenditure.

The remainder of this article is organized as follows. Section 2 presents the theoretical framework. Section 3 outlines context of the study. Section 4 outlines the data types and sampling method. Section 5 presents the descriptive statistics of the data. Section 6 details the methods used in this study and the approach of regressions used. Section 7 reports the regression results. Section 8 provides concluding remarks.

2: THEORETICAL FRAMEWORK

The theoretical foundation of this work rests on the prior works of microeconomic theory on consumer choice and preferences (Mas-Colell, Whinston, Green 1995 & Nicholson and Snyder 2008). The utility function of the household level aggregate (the main shopper of the household) is comprised of each food category type x_1, x_2, etc. and vector z, which is comprised of all other factors that influence utility (held constant). For simplicity, goods $x_1, x_2, \ldots x_n$. are a bundle of both healthy and unhealthy foods that each contribute towards raising utility levels through consumption. Each combination of these goods differs in terms of cost, nutrition, and utility. This utility function is represented as equation 1 below:

$$U_{HH} = f(x_1, x_2, x_3, \ldots, z) \quad (1)$$

Utility refers to the satisfaction gained from buying and using goods and services (Nicholson and Snyder, 2008), considering preferences for various food categories based on nutrition requirements and a budget constraint. Prices of each food and the cost of time influence the consumer choice problem, where main shoppers seek to maximize utility, subject to a budget constraint, nutritional needs, and time constraints. Simultaneously, an expenditure minimization problem arises, reframing the goal to allocate income, time, and nutritional requirements to achieve a minimum utility level at minimal expenditure. The expenditure function includes food category prices, personal utility and household utility, and outlet distance as an indicator of the opportunity cost of time. The model incorporates both healthy and unhealthy food prices, recognizing that households derive utility from both (Fung et al. 2018 and Bromage et al. 2021).

Equation 2 represents the dual expenditure minimization problem linked with utility maximization (Nicholson and Snyder 2008). Total expenditures per capita *Expendpercap* of *i* healthy and unhealthy foods are represented in equation 2 where p_i is the *i*th good price and x_i is the *i*th good. The goal is to minimize expenditure per capita subject to a minimum level of utility U_{HH} that is above some threshold y:

$$\text{minimize } Expendpercap = f(p_1 * x_1 + p_2 * x_2 + \cdots + p_i * x_i) \quad (2)$$
$$s.t. U_{HH} = f(x_1, x_2, x_3, \ldots, z) > y$$

Next, outlet distances are added as an additional input into the expenditure function. The function, where px_i represents the *i*th unhealthy or healthy food price, U_{HH} represents the desired utility level of the household is represented as equation 3. The left-hand-side variable is expenditure per capita. Vector *v* represents all other factors that have influence on expenditure (such as nutrition levels, etc.):

$$TotalFoodExpenditure = f(px_1, px_2, \ldots, px_i, outletdistance_j, v) \quad (3)$$

For each of the above equations, the key variable of interest is the effect of $outletdistance_j$, or the distance to the jth outlet type, on the dependent variable listed. In general, we assume that the above expenditure function, in accordance with microeconomic theory, follows the properties of homogeneity, non-decreasing in prices, and concavity in prices (Nicholson and Snyder, 2008).

3: METHOD AND DATA

3.1: Sampling

The study focuses on urban/peri-urban areas in Nairobi and Kisumu, Kenya. Using multi-stage sampling, areas were categorized into quartiles based on a combination of population level wealth and socio-economic status—e.g., household asset index, dwelling score, communication index, education index, and employment index. A two-stage cluster random sampling method identified 31 locations across the four quartiles, each with two enumeration areas (EAs). This resulted in 16 EAs in urban Nairobi, 16 in peri-urban Nairobi, 16 in urban Kisumu, and 14 in peri-urban Kisumu. Aiming for about 375 households per area, 23-28 households were targeted per EA. However, due to implementation challenges and non-responses our sample includes 18-24 households per EA.

3.2: Data

Data for this paper come from two sources—the household survey and the food environment census survey. These surveys were conducted between April and June 2022. The household food expenditure data come from interviews conducted with the main shopper in each household. Survey also collected data on demographic, economic and geographic characteristics of the main shopper and individual members of the household.

Food environment data come from a census of all food outlets operating in the home food environment of households surveyed. The home food environment is defined as a radius around the mean center, or centroid, of each enumeration area (EA). For Nairobi urban and peri-urban, as well as Kisumu urban, the radius was 0.4 kilometers; for Kisumu peri-urban, the radius was 0.6 kilometers. Each outlet was asked about what food groups (as defined by Bromage, 2021) are sold. We also recorded their GPS location.

Table 1 presents 14 different types of outlets we found across 61 home environments. Data suggests that the home food environments in our study area are dominated by traditional retail food businesses like duka, mama mboga, street vendors, and informal prepared food outlets. Modern retail outlets like supermarkets are less common.

3.3: Healthiness of food groups and outlet types

Food groups are classified as healthy and unhealthy as shown in Table 2. In total the list includes 25 food groups. According to Bromage et al. (2021), there are two types of unhealthy food categories: unhealthy at any consumption level and unhealthy at an excessive consumption level. For the sake of simplicity, this analysis combines 'unhealthy in excess' food groups with 'healthy' food groups, because in a developing country like Kenya, high fat dairy and red meat consumption is limited and fall mostly in the healthy quantity range.

Table 1: Food outlet type information and number of outlets enumerated across the study areas

Outlet Type	Description	# of outlets in survey
Small supermarket	Any self-service food outlet with 1-4 cash registers	30
Large supermarket	Any self-service food outlet with more than 4 cash registers	146
Duka (e.g. small traditional shop)	Traditional (not self-service) food outlets with permanent, constructed quarters	1712
Kiosk	Small, free-standing, "semi-movable" with rudimentary or transient structure	463
Mama mboga	Vegetable seller/vendor	1288
Street vendor	A seller located outside a market on the streets and selling from a mobile structure or from ground	926
Hawker	Seller that sells items on with or without a cart	89
Depot/wholesale	An outlet that primarily sells goods in bulk	59
Milk bar/milk atm	Outlets that primarily sell dairy products	72
Hotel/restaurant	An outlet selling prepared food for consumption on the premises, featuring permanent construction	469
Informal prepared food	Same as street vendor or kiosk—but specialized in selling prepared foods	709
Cereal shops and posho mills	Specialize in selling dry grains (cereals and pulses), and flours from these grains.	249
Bakery	An outlet that primarily sells baked goods	6
Butchery	An outlet that primarily sells red or white meat	315
	Total:	6533

Table 2: Food group healthiness classification per Bromage et al. (2021)

#	Food group type	Healthiness classification	#	Food group type	Healthiness classification
1	Citrus fruits	Healthy	17	High fat dairy	Unhealthy in excessive amounts
2	Deep orange fruits	Healthy	18	Red meat	Unhealthy in excessive amounts
3	Other fruits	Healthy	19	Processed meat	Unhealthy
4	Dark green leafy vegetables	Healthy	20	Refined grains and baked goods	Unhealthy
5	Cruciferous vegetables	Healthy	21	Sweets and ice cream	Unhealthy
6	Deep orange vegetables	Healthy	22	SSBs (sugar sweetened beverages)	Unhealthy

#	Food group type	Healthiness classification
7	Other vegetables	Healthy
8	Legumes	Healthy
9	Deep orange tubers	Healthy
10	Nuts and seeds	Healthy
11	Whole grains	Healthy
12	Liquid oils	Healthy
13	Fish and shellfish	Healthy
14	Poultry & game meat	Healthy
15	Low fat dairy	Healthy
16	Eggs	Healthy
23	Juice	Unhealthy
24	White roots and tubers	Unhealthy
25	Purchased deep fried foods	Unhealthy

Note: adapted from Bromage et al. (2021)

Using the census data, outlet healthiness characteristics are calculated in terms of the average number of healthy and unhealthy food groups on offer per outlet type (Table 3). Definitions of food group healthiness are based on Bromage et al. (2021) (see Table 2). Food outlets with the highest average total number of healthy food groups are mama mbogas, small supermarkets, and dukas, Admittingly, we do find some 'unhealthy'-labeled outlets, such as street vendors and hawkers, to have a higher average healthy-to-unhealthy food group ratio than some 'healthy'-labeled outlets, such as supermarkets and dukas.

Overall, we find that there is heterogeneity in food offerings across outlet types, with healthy food group averages ranging from 0.32-5.80, while unhealthy food group averages range from 0.10-4.10. Small supermarkets, as expected, have both the healthiest and the unhealthiest food groups on offer across all home food environments.

Table 1: Outlet healthiness characteristics

	Average across outlets				
Outlet type	Number of total food groups (25 total)	Number of healthy food groups (16 total)	Number of unhealthy in excess food groups (2 total):	Number of unhealthy food groups (7 total):	Ratio of number of healthy to unhealthy food groups
Small supermarket	8.9	4.47	0.33	4.1	1.09
Duka (e.g. small traditional shop)	7.33	3.29	0.41	3.63	0.91
Kiosk	5.41	3.17	0.16	2.11	1.50

Mama mboga	6.3	5.8	0.009	0.49	11.84
Street vendor	2.1	1.56	0.02	0.53	2.94
Hawker	1.34	0.82	0	0.34	2.41
Depot/wholesale	4.36	2.03	0.25	2.07	0.98
Milk bar/milk atm	1.92	0.51	0.97	0.43	1.19
Hotel/restaurant	1.12	0.32	0.06	0.75	0.43
Informal prepared food vendor	0.84	0.33	0.02	0.49	0.67
Cereal shop/poshomill	2.99	2.32	0.02	0.65	3.57
Butchery	1.32	0.39	0.84	0.1	3.90
Bakery	2.67	0.67	0.17	1.83	0.37

Scale: green (most) to red (least)
Source: Kenya food environment survey (2022)
Note: data unavailable for large supermarkets
Note: ratio column 'healthy' includes 'unhealthy in excess'

3.4: GIS data

The household survey collected GPS coordinates of household locations and the census collected GPS information for all the food outlets in the 0.4 (or 0.6) km radius (the "home food environment"). After plotting the household and outlet data, GIS tools were used to estimate distances between household points and outlet locations by outlet type. The ArcGISOnline "Proximity-Find Nearest" tool was used to find straight-line distances. Any missing distance values per observation are replaced with the sample mean of outlet distance (11 occurrences).

3.5: Survey weights and extreme values

Population weights are used in all analyses, such that the sample is representative of the population of urban and peri-urban Nairobi and Kisumu. In regression analysis, we cluster standard errors at the EA level, since multistage sampling technique was used, and EA was the unit of household sample selection. To minimize the bias that comes from extreme values, the expenditure values beyond the 99th percentile of the distribution are dropped from the analysis

4: DESCRIPTIVE STATISTICS OF THE DATA

Tables 4, 5, 6, and 7 present key descriptive statistics of the data. Table 4 highlights the average household size across the sample is 3, with variation from 3-4 on average across regions. The region with the highest average income is Nairobi urban, with 14,131 KES (Kenyan Shillings) while Kisumu peri-urban has the lowest reported average income 8,434 KES. Following Bromage et. al. (2021), aggregate average food expenditure amounts are calculated for healthy and unhealthy groups in Table 5. The overall mean monthly food expenditure per capita is 2,867 KES while average healthy food expenditure per capita is 1225 KES and average unhealthy food expenditure per capita is 1215 KES. In Table 6, main shopper outlet expenditure statistics are presented. The self-reported average distance (in kilometers) to each outlet varies from 0.24 km (milk bar/milk atm) to 9.55 km (depot). The top three closest outlets are: milk bar/milk atm (0.24 km), Kiosk (0.30 km), and mama mboga (0.30 km). The top three farthest outlets are: depot (9.55 km), wholesale (4.32 km), and hotel/restaurant (3.71 km). In Table 7, average distances to each outlet type are displayed as collected in the food environment census.

Table 2: Household and main shopper characteristics

Variable	Overall (mean and SD)	Nairobi u. (mean and SD)	Nairobi p.u. (mean and SD)	Kisumu u. (mean and SD)	Kisumu p.u. (mean and SD)
Household size (mean)	3 (1.85)	3 (1.87)	3 (1.66)	4 (3.31)	4 (2.42)
Household income per month (KES)	12,524 (14332.17)	14,131 (15773.39)	10,806 (11975.41)	10,021 (12374.1)	8,434 (12182.05)
Poverty score	61 (12.41)	61 (12.74)	61 (11.64)	58 (13.21)	57 (12.76)
Number of children (age < 18) in household	1 (1.43)	1 (1.47)	1 (1.27)	2 (1.68)	2 (1.81)
Household owns a bicycle	15%	17%	12%	13%	17%
Household owns a car	7%	9%	6%	3%	0.5%
Household owns a truck	2%	3%	1%	0.2%	0.2%
Household has 2 main shoppers	38%	35%	43%	38%	48%
Main shopper gender female	68%	66%	71%	75%	70%

Main shopper age (years)	36 (12.03)	35 (10.56)	37 (13.60)	34 (12.58)	36 (15.23)
Main shopper education (years)	13 (6.10)	14 (6.44)	13 (5.44)	12 (5.51)	14 (7.58)
% of main shoppers who value food nutrition as most important	46%	41%	54%	43%	43%
% of main shoppers who value food taste as most important	33%	32%	35%	30%	33%
% of main shoppers who value food convenience as most important	26%	28%	22%	26%	29%
% of main shoppers who value food price as most important	80%	80%	78%	82%	89%
% of main shoppers who value food nutrition as most important	34%	34%	31%	39%	46%
% of main shoppers who value food perishability as most important	33%	32%	35%	31%	33%
Sample size n	1496	368	354	382	392

Note: data weighted with survey population weights
Note: Poverty score is from Schreiner (2018). Responses are collected on 10 indicators and then used to estimate consumption-based poverty rates. Values signify the likelihood that the individual experiences poverty. Scores range from 0-100, with lower scores indicating higher poverty likelihood.
Note: under occupation, ‘other’ includes too young, retired, unpaid worker/volunteer

Table 5: Food expenditure descriptive statistics monthly per capita (KSH)

Variable	Overall (mean and SD)	Nairobi u. (mean and SD)	Nairobi p.u. (mean and SD)	Kisumu u. (mean and SD)	Kisumu p.u. (mean and SD)
Total	2,379 (1,624.76)	2,350 (1,541.16)	2,485 (1,758.97)	2,101 (1,571.57)	2,099 (1,417.62)
Healthy	1,322 (1,002.04)	1,329 (980.13)	1,333 (1,044.14)	1,167 (945.62)	1,307 (978.58)
Unhealthy	973 (697.23)	983 (710.46)	997 (679.53)	830 (700.85)	781 (626.52)
Sample size n	1425	352	335	362	376

Note: population weights applied
Note: for all variables, it is assumed 'unhealthy in excess' is grouped with the 'healthy' category
Note: Observations with extreme values greater than 99^{th} percentile are dropped

Table 6: Main shopper's (MS) self-reported outlet distances and expenditure statistics

Outlet type	MS self-reported distance (km) on average (mean and SD)	Amount spent at outlet per capita per month in KES (mean and SD)	Top 5 items purchased at outlet (% HHs) on average
Small supermarket	2.02 (5.82)	292 (1,538.48)	1. Sugar (white, granulated or lump) 2. Vegetable oil 3. Maize meal/flour, sifted 4. Rice (white, milled, polished grain) 5. Wheat flour (refined, fortified, sifted)
Large supermarket	5.2 (15.33)	941 (2,739.26)	1. Sugar (white, granulated or lump) 2. Maize meal/flour, sifted 3. Vegetable oil 4. Rice (white, milled, polished grain) 5. Wheat flour (refined, fortified, sifted)
Duka (e.g. small traditional shop)	0.50 (3.06)	2,932 (4,281.77)	1. Sugar (white, granulated, or lump) 2. Vegetable oil 3. Maize meal/flour (sifted) 4. Milk (cow, whole, fresh) 5. Bread (white)
Kiosk	0.30 (2.26)	259 (946.05)	1. Tomato (red, ripe) 2. Kale (Sukuma wiki) 3. Vegetable oil

			4. Onion (mature, red skinned, peeled) 5. Sugar (white, granulated, or lump)
Mama mboga	0.30 (0.90)	1004 (1,862.63)	1. Tomato (red, ripe) 2. Kale (sukuma wiki) 3. Onion (mature, red skinned, peeled) 4. Cabbage (leaf head, white) 5. Onion (spring, raw)
Street vendor	1.49 (3.08)	98 (660.25)	1. Sardines 2. Basic Mandazi 3. Milk (cow, whole, fresh) 4. Tomato (red, ripe) 5. White Chapati
Hawker	0.58 (4.69)	15 (185.51)	1. Kunde (cowpeas, leaves, picked) 2. Mrenda (jute mallow, picked leaves) 3. Sardines 4. Watermelon 5. Managu (leafy green vegetable)
Depot	9.55 (28.76)	31 (309.32)	1. Rice (white, milled, polished grain) 2. Yellow beans dry raw 3. Vegetable oil 4. Gram (green, dry) 5. Milk (cow, whole, fresh)
Wholesale	4.32 (10.62)	199 (1,083.887)	1. Vegetable oil 2. Maize meal/flour (sifted) 3. Rice (white, milled, polished grain) 4. Sugar (white, granulated, or lump) 5. Wheat flour (refined/fortified/sifted)
Milk bar/milk atm	0.24 (0.42)	62 (285.30)	1. Milk (cow, whole, fresh) 2. Milk (cow, whole, fermented) 3. Milk (cow, condensed, skimmed, sweetened) 4. Milk (cow, powder, skimmed) 5. Milk (cow, powder, whole)
Hotel/restaurant	3.71 (11.88)	46 (545.66)	1. White Chapti 2. Ugali from refined maize flour 3. Red beans stew 4. Bean stew 5. Chai ya Maziwa (mixed tea), with sugar

Informal prepared food	2.32 (23.04)	187 (557.75)	1. Basic Mandazi 2. White Chapati 3. Boiled beans 4. Githeri (fresh maize and dry beans) 5. Githeri (dry maize and dry beans)
Posho mill or cereal shop	0.78 (2.12)	108 (477.0)	1. Whole maize meal flour 2. Maize grain (white variety, whole, dry) 3. Millet (finder, flour) 4. Cornflour (from maize starch) 5. Rice flour
Butchery	2.45 (27.91)	376 (1,072.18)	1. Beef (with bones) 2. Beef (medium fat, without bones) 3. Matumbo (tripes) 4. Beef (high fat, without bones) 5. Beef (lean)

Note: population weights applied
Note: data source is main shopper census survey

Table 7: Linear distance (km) and counts within the home food environment

Outlet type	Count in home FE	Entire sample (km)	Nairobi u. (km)	Nairobi p.u. (km)	Kisumu u. (km)	Kisumu p.u. (km)
All outlets	6533	0.08 (0.25)	0.08 (0.19)	0.08 (0.33)	0.07 (0.09)	0.15 (0.15)
Informal prepared food outlet	709	0.24 (0.57)	0.13 (0.21)	0.36 (0.70	0.15 (0.12)	0.84 (1.68)
Milk bar/milk atm	72	1.47 (2.57)	0.43 (0.44)	2.42 (2.92)	1.01 (0.84)	8.84 (3.95)
Depot/wholesale	59	1.87 (2.76)	0.43 (0.58)	3.54 (3.19)	2.33 (2.01)	6.63 (4.05)
Hawker	89	1.10 (1.64)	0.91 (1.19)	1.10 (1.93)	2.25 (1.49)	2.54 (2.90)
Street vendor	926	0.40 (1.08)	0.24 (0.42)	0.62 (1.59)	0.25 (0.35)	0.81 (1.82)
Mama mboga	1288	0.18 (0.51)	0.11 (0.21)	0.23 (0.56)	0.12 (0.12)	0.93 (1.68)
Kiosk	463	0.26 (0.67)	0.13 (0.21)	0.34 (0.70)	0.35 (0.44)	1.41 (2.28)
Duka	1712	0.12 (0.34)	0.09 (0.20)	0.12 (0.33)	0.12 (0.11)	0.50 (1.16)
Small supermarket	30	1.90 (2.75)	0.86 (0.81)	2.71 (3.38)	2.61 (1.58)	9.00 (2.94)

Large supermarket[a]	146	1.51	1.16	2.18	--	--
		(1.27)	(0.68)	(1.75)		
Hotel/restaurant	469	0.62	0.24	1.17	0.42	1.33
		(1.57)	(0.40)	(2.33)	(0.46)	(2.33)
Poshomill or cereal shop	249	1.87	0.50	2.99	0.84	3.14
		(2.61)	(0.66)	(3.60)	(0.82)	(2.78)
Butchery	315	0.43	0.26	0.37	0.58	3.69
		(1.03)	(0.41)	(0.70)	(0.50)	(3.34)

a: large supermarket statistics for Kisumu urban and peri-urban are currently under development.
Note: values in parentheses are standard deviation

5: METHOD AND APPROACH

5.1: Regression analysis

A multivariate regression using log-log functional form and OLS estimation is employed. After accounting for missing data, our analytical sample is 1495 households.

Our key dependent variable is the log of household main shopper's healthy monthly expenditure, while the key independent variables are the log of each type of outlet distance as calculated using the food environment survey (not self-reported distance). A 1 is added to each observation to avoid undefined values when taking the log.

Table 8: Details of vectors for model

Vector Name	Controls included
$HH\_characteristics_i$	HHsize, poverty score, # of children, if own bicycle, if own car, if own truck
$MS\_characteristics_i$	gender, age, education, occupation, food values (nutrition, taste, convenience, price, availability, perishability)
$Location\_controls_i$	Region, EA (enumeration area), census wealth index location strata
$Log\_outlet\_distance_{ij}$	Log of linear distance from main shopper household I to nearest shopped outlet of type j: informal prepared outlet, milk bar/milk atm, depot/wholesaler, hawker, street vendor, mama mboga, kiosk, duka, small supermarket, hotel/restaurant, poshomill/cereal shop, butchery

Food prices are not explicitly controlled for in the model for two reasons. First, it is assumed that food prices are homogenous within a given EA. Thus, once location controls are included in the model, all factors that are constant within the EA, including food prices, weather infrastructure, and wages, are also controlled for. Second, prices are self-reported by the main shoppers and thus are endogenous to the model and not suitable for use as a control.

Due to the presence of heteroskedasticity, standard errors are clustered at the EA level. As indicated earlier, survey weights are applied to the data.

Our general empirical model that will be estimated for the *i*th household is as follows:

$$\log\_health\_expend\_percap_i = \alpha + \beta_1 \log\_outlet\_distance_i + \delta_1 HH\_characteristics_i + \delta_2 MS\_characteristics_i + \delta_3 Location\_controls_i + \varepsilon_i \quad (4)$$

The sequence of regressions is first a parsimonious model, followed by the addition of household controls, main shopper controls, and location controls. Our main dependent variables that will be utilized is the log of healthy food expenditure per capita.

5.2: Group heterogeneity tests

For each set of groups outlined in Table 9, a fully saturated model is estimated. We use log of monthly healthy expenditure per capita as the dependent variable for this specification. Then, a Wald test is run to indicate if the coefficients of the two groups are equal. Any significance of the interacted outlet distance variables is identified. The associated hypotheses are:

H_0: groups' coefficient is the same:

$$\beta_i = \beta_j \text{ or } \beta_i - \beta_j = 0$$

H_1: groups' coefficient is different:

$$\beta_i \neq \beta_j \text{ or } \beta_i - \beta_j \neq 0$$

Table 9: Identification of dummy variables for heterogeneity test

Characteristic	Variable	Description (using medians)
Age	$young_i$	• 1 if main shopper's age is less than 34 • 0 if main shopper's age is greater than or equal to 34
Gender	$male_i$	• 1 if main shopper's gender is male • 0 if main shopper's gender is female
Location (urban vs. peri-urban)	$urban_i$	• 1 if household location is urban • 0 if household location is peri-urban
Poverty probability	$lowpov_i$	• 1 if household's chance of poverty is greater than 59 • 0 if household's chance of poverty is less than or equal to 59
Strata location (based on census wealth index)	$poor_i$	• 1 if household belongs to a location in quartiles 1 or 2 (more poor) • 0 if household belongs to a location in quartiles 3 or 4 (more rich)
Transportation (car)	car_i	• 1 if main shopper's household owns a car • 0 if main shopper's household does not own a car

5.3: Robustness check

Two classes of robustness checks are employed: 1.) While the main analysis groups 'unhealthy in excess' food items with the 'healthy' category for the dependent variable, 2 reclassifications are tested: a.) Adding 'unhealthy in excess' to 'unhealthy' and b.) excluding 'unhealthy in excess' food items entirely and 2.) With partially complete data on large supermarkets only in Nairobi, the model is rerun to check for consistency.

6: RESULTS

6.1: Regression results and interpretation

Columns A-D of Table 10 represent models with different combination of controls added to the model. The key variables of interest are the logged outlet distances and their corresponding distance elasticities.

Results show as expected the same coefficient but of greater magnitude: for a 1% increase in distance to nearest cereal shop/poshomill, healthy expenditure increases by 0.34% on average for the fully saturated model. As the results move from only the parsimonious model in column A to the fully saturated model in column D, we find several significant variables. But, once location controls are added in column D, much of this significance disappears.

Regression results are robust for select outlet types in columns A-C of Table 10 (prior to location controls). We find that a 1% increase in the distance to the nearest milk bar/milk atm is correlated with a 0.10% increase in monthly healthy food expenditure per capita on average. We also find a consistent negative and significant coefficient on hawkers (0.07). Mama mboga outlet distance and hotels and restaurants distance also have a similar inverse relationship with respect to healthy expenditure.

Our main hypothesis, that as distance increases from the main shopper's household location to each of the relatively healthy outlet types (mama mbogas, small supermarkets, large supermarkets, dukas, kiosks, cereal shops/poshomills, and depots/wholesalers), healthy expenditure share will decrease, is partially confirmed. If we focus on column C in Table 10, which controls for household and main shopper characteristics, we find negative, significant coefficients for mama mboga outlets only.

Table 10: Regression results with the log of healthy monthly expenditure per capita as dependent variable

	A	B	C	D
Log of informal prepared distance	0.10	0.04	0.05	-0.02
	(0.06)	(0.06)	(0.06)	(0.10)
log of milk bar/milk atm distance	0.11**	0.082**	0.10***	-0.17
	(0.04)	(0.04)	(0.03)	(0.11)
log of depot/wholesaler distance	-0.07	-0.02	-0.02	-0.14
	(0.05)	(0.04)	(0.03)	(0.18)
log of hawker distance	-0.08**	-0.05*	-0.07***	0.03
	(0.03)	(0.03)	(0.02)	(0.06)
log of street vendor distance	-0.05	-0.05	-0.05	-0.06
	(0.05)	(0.05)	(0.05)	(0.10)
log of mama mboga distance	-0.07*	-0.11***	-0.09**	0.01
	(0.04)	(0.04)	(0.04)	(0.07)
log of kiosk distance	-0.02	0.07	0.0922**	0.09
	(0.05)	(0.05)	(0.04)	(0.06)
log of duka distance	0.02	0.01	-0.02	-0.04
	(0.05)	(0.05)	(0.05)	(0.07)
log of small supermarket distance	0.00	-0.01	0.03	-0.07
	(0.05)	(0.05)	(0.04)	(0.14)
log of hotel/restaurant distance	-0.10*	-0.12**	-0.13***	-0.09
	(0.05)	(0.05)	(0.04)	(0.09)
log of butchery distance	-0.07	-0.02	-0.03	0.00
	(0.06)	(0.05)	(0.05)	(0.09)
log of cereal/poshomill distance	0.12***	0.11**	0.09**	0.34*
	(0.05)	(0.04)	(0.04)	(0.18)
Household controls	No	Yes	Yes	Yes
Main shopper controls	No	No	Yes	Yes
Location controls	No	No	No	Yes
Observations	1,429	1,429	1,429	1,429
R-squared	0.05	0.13	0.17	0.22

Household sample weights applied

Robust standard errors in parentheses clustered at EA level

*** p<0.01, ** p<0.05, * p<0.1

6.2: Heterogeneity test results and interpretation

Across all results, there is select detected heterogeneity among groups. When comparing young and old main shoppers, distance elasticities differ only for mama mboga outlets (age).

Results also show that location of main shopper's household matters for informal prepared outlets and small supermarkets. Chance of poverty (street vendor and small supermarkets), wealth (dukas), and access to car transportation (large supermarkets) also have heterogeneity for select outlet types. No difference between male and female main shopper groups is found for any outlet type.

Table 11: Wald test results for difference in coefficient value

Variable	Age less than or greater than/equal to median	household poverty probability less than vs. greater than/equal to median	urban vs. peri-urban location	male vs. female shopper	location strata (quartiles 1-2 vs. 3-4)	HH owns a car (Yes vs. No)
			P-value			
Log of informal prepared distance	0.68	0.26	0.08***	0.97	0.30	0.65
Log of milk bar/milk atm distance	0.38	0.67	0.20	0.58	0.16	0.002***
Log of depot/wholesaler distance	0.87	0.23	0.91	0.20	0.84	0.74
Log of hawker distance	0.94	0.77	0.76	0.10	0.63	0.42
Log of street vendor distance	0.94	0.003***	0.12	0.24	0.16	0.17
Log of mama mboga distance	0.06***	0.35	0.82	0.54	0.36	0.97
Log of kiosk distance	0.20	0.18	0.33	0.67	0.89	0.07***
Log of duka distance	0.35	0.93	0.75	0.71	0.005***	0.64
Log of small supermarket distance	0.08***	0.12	0.008***	0.14	0.15	0.35
Log of hotel restaurant distance	0.16	0.88	0.26	0.52	0.15	0.20
Log of butchery distance	0.59	0.71	0.85	0.33	0.72	0.59
Log of cereal shop/poshomill distance	0.65	0.90	0.13	0.40	0.88	0.56

*** if p-value is < 0.1

6.3: Robustness check results

Appendix A highlights regression results for the previously outlined robustness checks. Again, the purpose of the first check was to test the robustness of the results under different classifications of the dependent variable, while the second check addresses if large supermarkets are added to the Nairobi.

The regressions in Appendix A show that results are generally robust and show similar patterns. Signs and magnitudes of coefficients are comparable; significance also follows a similar pattern. We even continue to see the same phenomenon with location controls; much of the coefficient significance is dropped once accounting for household location characteristics.

The second part of our robustness checks shows how the results change if large supermarkets are included. Table 12 outlines regression results and it shows that results are robust across all 4 specifications A-D. We find similar significance, magnitude, and sign of coefficients as well as ascending R-squared values. The log of large supermarket distance is not found to be significant for any of the specified models A-D.

Table 12: Robustness check regression results including large supermarkets for Nairobi urban and peri-urban sub-sample

	A	B	C	D
Log of informal prepared distance	0.12*	0.05	0.05	-0.01
	(0.07)	(0.06)	(0.07)	(0.12)
log of millk bar/milk atm distance	0.14**	0.09**	0.10**	-0.15
	(0.05)	(0.04)	(0.04)	(0.11)
log of depot/wholesaler distance	-0.05	-0.02	-0.05	-0.28
	(0.06)	(0.05)	(0.04)	(0.22)
log of hawker distance	-0.08**	-0.05*	-0.07***	0.03
	(0.03)	(0.03)	(0.02)	(0.08)
log of street vendor distance	-0.05	-0.06	-0.05	-0.09
	(0.05)	(0.05)	(0.05)	(0.12)
log of mama mboga distance	-0.08*	-0.11**	-0.10*	0.03
	(0.04)	(0.05)	(0.05)	(0.08)
log of kiosk distance	-0.01	0.11**	0.14**	0.09
	(0.05)	(0.05)	(0.05)	(0.06)
log of duka distance	-0.02	-0.01	-0.04	-0.10
	(0.06)	(0.06)	(0.06)	(0.08)
log of small supermarket distance	-0.01	-0.03	-0.01	-0.04
	(0.05)	(0.04)	(0.04)	(0.15)
Log of large supermarket distance	-0.07	0.03	0.03	-0.40

	(0.09)	(0.07)	(0.07)	(0.26)
log of hotel/restaurant distance	-0.16***	-0.16***	-0.16***	-0.16
	(0.05)	(0.04)	(0.04)	(0.10)
log of butchery distance	-0.04	-0.01	-0.01	0.05
	(0.08)	(0.07)	(0.06)	(0.11)
log of cereal/poshomill distance	0.15***	0.14***	0.14***	0.41*
	(0.05)	(0.04)	(0.04)	(0.20)
Household controls	No	Yes	Yes	Yes
Main shopper controls	No	No	Yes	Yes
Location controls	No	No	No	Yes
Observations	695	695	695	695
R-squared	0.06	0.15	0.19	0.22

Household sample weights applied

Robust standard errors in parentheses clustered at EA level

*** p<0.01, ** p<0.05, * p<0.1

7: CONCLUSION

7.1: Summary

This paper analyzed the strength of the relationship between a simple food environment measure, linear distance to 14 different food outlets, and household main shoppers' healthy food expenditure in urban and peri-urban Kenya. The key takeaways, among all results, from this research are:

1. Once location is controlled for, much of the variation in healthiness of expenditure is attributed to where people live and their characteristics, not distance to specific outlets.
2. The main hypothesis, that as distance increases from the main shopper's household location to each of the healthy outlet types, healthy expenditure will decrease, is partially confirmed. It was found that only for mama mboga outlets, who are street vegetable sellers, does increasing distance result in a decrease in healthy expenditure. Surprisingly, increasing distance to large and small supermarkets results in increases in healthy food expenditure, ceteris paribus.
3. Heterogeneity by group exists in the results: various model coefficients are statistically significant in terms of age, gender, location, poverty probability, and wealth location socio-economic strata. More research is needed to determine the nuances of this striking

difference, especially since mama mboga outlets supply a plethora of healthy food groups.

4. From the descriptive statistics, we find that there is considerable heterogeneity in terms of what different types of foods outlets have on offer in terms of healthiness as well as what sampled main shoppers expend their resources on.
5. Robustness checks indicate that results are robust across variation definitions of the 'healthy' food category.

This research disaggregates distances to specific food outlets to create a more nuanced measure of the food environment in urban and peri-urban Kenya and apply the concept of distance elasticity. The results show that distance to some outlets relative to others can have an impact on how the main shopper responds in terms of food expenditure. It is also emphasized in the findings that how food groups are classified into 'healthy' and 'unhealthy' also can dictate results.

7.2: Study Limitations

This paper presented a simple measure of household-outlet proximity, one that only captures linear distance between two points. While this is a validated FE measurement technique, it does not fully reflect real-world navigation of city streets. The distance measured used in this study also does not account for mode of transportation. Lack of exogenous food price data also limits the study's ability to control for micro-level food prices. Due to data limitations, only small supermarkets were included in this analysis of this study. Future results from this study would greatly benefit from the inclusion of large supermarkets.

7.3: Concluding remarks and implications

The Kenyan government has made it clear that, among many policy-related issues, ensuring households are able to purchase foods that contribute towards an adequate, diverse, and healthy diet is a key policy goal (Kenya Ministry of Agriculture, 2011). With major systemic shifts in nutrition, urbanization, and food systems on the horizon, it is apparent that how low-income consumers in urban areas access and acquire food for themselves and their household is also rapidly changing. The dynamic nature of a household's food environment warrants a closer look at how changes in outlet proximity affect the nutritional quality of food that is purchased.

This study helps fill this gap in understanding through the analysis of a rich dataset capturing both household-level consumption, characteristics, and food purchases and food outlet information in an urban and peri-urban setting. With a focus on estimating distance elasticities that quantify the responsiveness of household food shopping expenditure to variations in distances to different food outlets, this research builds new knowledge in understanding the sensitivity of nutritional choices of low-income urban consumers in Kenya to alterations in their surrounding food environment. Further, heterogeneity in distance elasticities is investigated across economic, geographic, or gender differences of the main shopper of the household. The key finding of this work emphasizes that household location characteristics possess greater predictive power than average distance to outlets of healthiness of food purchases.

Future research on this topic can address questions such as how people substitute the healthiness of their purchases if outlet distances vary and what additional factors of outlets beyond distance might impact healthy food expenditure. Additional questions can be asked about how price changes signal consumers to switch between outlets, and how food expenditure decisions shift between healthy and unhealthy choices with price changes. Lastly, future research can investigate how the entrance of a new healthy or unhealthy outlet in a food environment changes the welfare of both existing outlets and consumers. Ongoing urbanization and food system transformation in Kenya has led to rapid changes in the socioeconomic distribution of people across urban areas and how and where food is acquired. A better understanding of the nature of these dynamics will lend itself to more responsive policy, improved economic welfare, and most importantly, a more food secure population.

BIBLIOGRAPHY

Athens, J.K., Duncan, D.T., & Elbel, B. (2016). Proximity to fast food outlets and supermarkets as predictors of fast food dining frequency. *Journal of the Academy of Nutrition and Dietetics*, 116(8): 1266-1275. DOI: 10.1016/j.jand.2015.12.022

Babey, S.H. et al. (2008). Designed for disease: the link between local food environments and obesity and diabetes. *UCLA Center for Health Policy Research*. https://escholarship.org/uc/item/7sf9t5wx

Barnes, T.L. et al. (2015). Geographic measures of retail food outlets and perceived availability of healthy foods in neighborhoods. *Journal of Public Health Nutrition*, 19(8): 1368-1374. DOI: 10.1017/S1368980015002864

Bergstrand, J.H., Larch, M., & Yotov, Y.V. (2015). Economic integration agreements, border effects, and distance elasticities in the gravity equation. *European Economic Review*, 78: 307-327. https://doi.org/10.1016/j.euroecorev.2015.06.003

Berthelon, M. & Freund, C. (2008). On the conversation of distance in international trade. *Journal of International Economics*, 75: 310-320. https://EconPapers.repec.org/RePEc:eee:inecon:v:75:y:2008:i:2:p:310-320

Black, C., Moon, G., & Baird, J. (2014). Dietary inequalities: What is the evidence for the effect of the neighborhood food environment? *Health & Place*, 27: 229-242. https://doi.org/10.1016/j.healthplace.2013.09.015

Bodor, J.N. et al. (2008). Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutrition*, 11:413-420. DOI: 10.1017/S1368980007000493

Bromage S. et al. (2021). Development and validation of a novel food-based global diet quality score (GDQS*)*. *Journal of Nutrition*, 151: 75-92. DOI: 10.1093/jn/nxab244

Caspi, C.E. et al. (2012). The local food environment and diet: A systematic review. *Health & Place*, 18: 1172-1187. https://doi.org/10.1016/j.healthplace.2012.05.006

Chenarides, L. & Jaenicke, E.C. (2016). Store choice and consumer behavior in food deserts: an empirical application of the distance metric method. Working Paper: 1-25. https://www.tsefr.eu/sites/default/files/TSE/documents/sem2016/food_eco/jaenicke_2.pdf

Clapp, J. et al. (2022). Viewpoint: The case for a six-dimensional food security framework. *Food Policy*, 106: 102164. https://doi.org/10.1016/j.foodpol.2021.102164

Crush, J. & Battersby, J. (2017). Rapid urbanization, urban food deserts, and food security in Africa. London.

Ebeke, C. & Etoundi, M.N. (2016). Does distance still matter for agricultural trade? Working paper. https://www.ifpri.org/publication/does-distance-still-matter-agricultural-trade

Farley, T.A. et al. (2009). Measuring the Food Environment: Shelf Space of Fruits, Vegetables, and Snack Foods in Stores. *Journal of Urban Health*, 86: 672-682. DOI: 10.1007/s11524-009-9390-3

Fraser, L.K. et al. (2010). The geography of fast food outlets: a review. *International Journal of Environmental Research and Public Health*, 7: 2290-2308. doi:10.3390/ijerph7052290

Fung, T.T., et al. (2018) International food group–based diet quality and risk of coronary heart disease in men and women. *Clinical Nutrition*, 107: 120-129. https://doi.org/10.1093/ajcn/nqx015

Glanz, K. et al. (2005). Healthy nutrition environments: Concepts and measures. *American Journal of Health Promotion*, 19: 330-343. 10.4278/0890-1171-19.5.330

Global Panel (2017). Improving nutrition through enhanced food environments. Policy Brief No. 7. London, UK: Global Panel on Agriculture and Food Systems for Nutrition. https://glopan.org/sites/default/files/FoodEnvironmentsBrief.pdf

Hawkes, C., Harris, J., & Gillespie, S. (2017). Changing diets: Urbanization and the nutrition transition. In 2017 Global Food Policy Report. Chapter 4. Pp 34-41. Washington, DC: International Food Policy Research Institute (IFPRI). https://doi.org/10.2499/9780896292529_04

Hawkes, C. (2008). Dietary implications of supermarket development: a global perspective. *Development Policy Review*, 26: 657-692. https://doi.org/10.1111/j.1467-7679.2008.00428.x

Herforth, A. & Ahmed, S. (2015). The food environment, its effects on dietary consumption, and potential for measurement within agriculture-nutrition interventions. *Food Security*, 7: 505-520. 10.1007/s12571-015-0455-8

Kenya Ministry of Agriculture (2011). National food and nutrition security policy: 2011. *The Kenya Institute for Public Policy Research and Analysis*. http://repository.kippra.or.ke/handle/123456789/1620

Kruger, D.J. et al. (2013). Local Concentration of fast-food outlets is associated with poor nutrition and obesity. *American Journal of Health Promotion*, 28:340-343. DOI: 10.4278/ajhp.111201-QUAN-437

Laraia, B.A. et al. (2004). Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventative Medicine*, 39: 869-875. DOI: 10.1016/j.ypmed.2004.03.018

Laska, M.N. et al. (2010). Neighbourhood food environments: are they associated with adolescent dietary intake, food purchases and weight status? *Public Health Nutrition*, 13: 1757-1763. DOI: 10.1017/S1368980010001564

Mas-Colell, A., Whinston, M.D. & Green, J.R. (1995). Microeconomic Theory. Oxford University Press.

Mergenthaler, M., Weinberger, K., & Qaim, M. (2009). The food system transformation in developing countries: a disaggregate demand analysis for fruits and vegetables in Vietnam. *Food Policy*, 34: 426-436. https://doi.org/10.1016/j.foodpol.2009.03.009

Monteiro, C.A. et al. (2013). Ultra-processed products are becoming dominant in the global food system. *Obesity Review*, 14: 21-28. DOI: 10.1111/obr.12107

Nicholson, W. & Snyder, C.M. (2008). Microeconomic theory: basic principles and extensions. Cengage Learning Press.

Pearson, T. et al. (2005). Do 'food deserts' influence fruit and vegetable consumption? A cross-sectional study. *Appetite*, 45: 195-197. DOI: 10.1016/j.appet.2005.04.003

Popkin, B.M. & Reardon, T. (2018). Obesity and the food system transformation in Latin America. *Obesity Reviews*, 19: 1028-1064. https://doi.org/10.1111/obr.12694

Popkin, B.M. (1993). Nutritional patterns and transitions. *Population and Development Review*, 19: 138-157. https://doi.org/10.2307/2938388

Popkin, B.M. (1999). Urbanization, lifestyle changes and the nutrition transition. *World Development*, 27: 1905-1916. https://doi.org/10.1016/S0305-750X(99)00094-7

Popkin, B.M (2004). The nutrition transition: an overview of world patterns of change. *Nutrition Review*, 62: 140-143. DOI: 10.1111/j.1753-4887.2004.tb00084.x

Reardon, T. et al. (2017). The rise of supermarkets in Africa, Asia, and Latin America. American *Journal of Agricultural Economics*, 85: 1140-1146. https://www.jstor.org/stable/1244885

Rose, D. & Richards, R. (2004). Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Healthy Nutrition*, 7: 1081-1088. DOI: 10.1079/PHN2004648

Rose, D. et al. (2009). Neighborhood Food Environments and Body Mass Index The Importance of In-Store Contents. *American Journal of Preventative Medicine*, 37: 214-9. DOI: 10.1016/j.amepre.2009.04.024

Ruel, M.T. et al. (2017). Urbanization, Food Security and Nutrition. In: de Pee, S., Taren, D., Bloem, M. (eds) Nutrition and Health in a Developing World . Nutrition and Health. Humana Press, Cham. https://doi.org/10.1007/978-3-319-43739-2_32

Schreiner, M. (2018). Simple poverty scorecard tool: Kenya. Scorocs, 1-69. https://www.simplepovertyscorecard.com/KEN_2015_ENG.pdf

Schwartz, A. (1973). Interpreting the effect of distance on migration. *Journal of Political Economy*, 81: 1153-1169. https://www.jstor.org/stable/1830643

Senauer, B. (1991). *Food trends and the changing consumer*. Eagan Press.

Sharkey, J.R. et al. (2011). Association between proximity to and coverage of traditional fast-food restaurants and non-traditional fast-food outlets and fast-food consumption among rural adults. *International Journal of Health Geographics*, 10:37. http://www.ij-healthgeographics.com/content/10/1/37

Spence, J.C. et al. (2009). Relation between local food environments and obesity among adults. *BMC Public Health*, 9:192. doi:10.1186/1471-2458-9-192

St-Onge, M.P., Keller, K.L., & Heymsfield, S.B. (2003). *The American Journal of Clinical Nutrition*. 78: 1068-1073. https://doi.org/10.1093/ajcn/78.6.1068

Swinburn, B. et al. (2013). INFORMAS (International Network for Food and Obesity/non-communicable diseases Research, Monitoring and Action Support): overview and key principles. *Obesity Review*, 14: 1-12. doi: 10.1111/obr.12087

Tschirley, D. et al. (2015). The rise of a middle class in East and Southern Africa: implications for food system transformation. *Journal of International Development*, 27: 628-646. https://doi.org/10.1002/jid.3107

Tschirley et al. (2021). Measuring the Food Environment and its Impact on Diets: Alternative Metrics Deliver Very Different Results in Urban Nairobi. Working Paper.

Turner, C. et al. (2018). Concepts and critical perspective for food environment research: a global framework with implications for action in low- and middle-income countries. *Global Food Security*, 18: 93-101. https://doi.org/10.1016/j.gfs.2018.08.003

Wang, Z. et al. (2000). The impact of distance on us ag exports: an econometric analysis. Working paper. file:///Users/ianfisher/Downloads/dot_38943_DS1.pdf

Wang, M.C. et al. (2007). Socioeconomic and food-related physical characteristics of the neighborhood environment are associated with body mass index. *Journal of Epidemiology and Community Health*, 61: 491-498. DOI: 10.1136/jech.2006.051680

Wrigley, N. (2002). "Food Deserts" in British Cities: Policy Context and Research Priorities. *Urban Studies, 39: 2029*-2040. https://doi.org/10.1080/0042098022000011344

Zenk, S.N. & Powell, L.M. (2008). US secondary schools and food outlets. *Health Place*, 14: 336-346. DOI: 10.1016/j.healthplace.2007.08.003

Zenk, S.N. et al. (2005). Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *American Journal of Public Health*, 95: 660-667. DOI: 10.2105/AJPH.2004.042150

APPENDIX A: ROBUSTNESS CHECK REGRESSIONS

Table 13: Regression results with the log of healthy monthly expenditure per capita as dependent variable with 'unhealthy in excess' added to 'unhealthy'

	A	B	C	D
Log of informal prepared distance	0.10	0.04	0.04	-0.01
	(0.07)	(0.06)	(0.06)	(0.11)
log of millk bar/milk atm distance	0.12***	0.10**	0.13***	-0.13
	(0.04)	(0.04)	(0.04)	(0.08)
log of depot/wholesaler distance	-0.09	-0.05	-0.05	-0.16
	(0.06)	(0.05)	(0.04)	(0.20)
log of hawker distance	-0.07**	-0.05	-0.06**	0.00
	(0.03)	(0.03)	(0.03)	(0.09)
log of street vendor distance	-0.06	-0.07	-0.06	-0.09
	(0.06)	(0.06)	(0.06)	(0.15)
log of mama mboga distance	-0.05	-0.08*	-0.07	0.06
	(0.04)	(0.04)	(0.05)	(0.08)
log of kiosk distance	0.05	0.12**	0.13**	0.16*
	(0.05)	(0.05)	(0.05)	(0.08)
log of duka distance	-0.05	-0.05	-0.07	-0.14*
	(0.06)	(0.06)	(0.06)	(0.08)
log of small supermarket distance	0.04	0.03	0.07	0.10
	(0.05)	(0.06)	(0.05)	(0.17)
log of hotel/restaurant distance	-0.10**	-0.12**	-0.12***	0.01
	(0.05)	(0.05)	(0.04)	(0.09)
log of butchery distance	-0.02	0.02	0.01	-0.05
	(0.06)	(0.05)	(0.05)	(0.09)
log of cereal/poshomill distance	0.07	0.06	0.04	0.20
	(0.05)	(0.05)	(0.04)	(0.16)
Household controls	No	Yes	Yes	Yes
Main shopper controls	No	No	Yes	Yes
Location controls	No	No	No	Yes
Observations	1,434	1,434	1,434	1,434
R-squared	0.03	0.09	0.13	0.19

Household sample weights applied

Robust standard errors in parentheses clustered at EA level

*** p<0.01, ** p<0.05, * p<0.1

Table 14: Regression results with the log of healthy monthly expenditure per capita as dependent variable with 'unhealthy in excess' excluded

	A	B	C	D
Log of informal prepared distance	0.10	0.04	0.04	-0.01
	(0.07)	(0.06)	(0.06)	(0.11)
log of millk bar/milk atm distance	0.12***	0.10**	0.13***	-0.13
	(0.04)	(0.04)	(0.04)	(0.08)
log of depot/wholesaler distance	-0.09	-0.05	-0.05	-0.16
	(0.06)	(0.05)	(0.04)	(0.20)
log of hawker distance	-0.07**	-0.05	-0.06**	0.00
	(0.03)	(0.03)	(0.03)	(0.09)
log of street vendor distance	-0.06	-0.07	-0.06	-0.09
	(0.06)	(0.06)	(0.06)	(0.15)
log of mama mboga distance	-0.05	-0.08*	-0.07	0.06
	(0.04)	(0.04)	(0.05)	(0.08)
log of kiosk distance	0.05	0.12**	0.13**	0.16*
	(0.05)	(0.05)	(0.05)	(0.08)
log of duka distance	-0.05	-0.05	-0.07	-0.14*
	(0.06)	(0.06)	(0.06)	(0.08)
log of small supermarket distance	0.04	0.03	0.07	0.10
	(0.05)	(0.06)	(0.05)	(0.17)
log of hotel/restaurant distance	-0.10**	-0.12**	-0.12***	0.01
	(0.05)	(0.05)	(0.04)	(0.09)
log of butchery distance	-0.02	0.02	0.01	-0.05
	(0.06)	(0.05)	(0.05)	(0.09)
log of cereal/poshomill distance	0.07	0.06	0.04	0.20
	(0.05)	(0.05)	(0.04)	(0.16)
Household controls	No	Yes	Yes	Yes
Main shopper controls	No	No	Yes	Yes
Location controls	No	No	No	Yes
Observations	1,434	1,434	1,434	1,434
R-squared	0.03	0.09	0.13	0.19

Household sample weights applied

Robust standard errors in parentheses clustered at EA level

*** p<0.01, ** p<0.05, * p<0.1