GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

GlobalFood
www.uni-goettingen.de/globalfood

RTG 1666 GlobalFood

Transformation of Global Agri-Food Systems:
Trends, Driving Forces, and Implications for Developing Countries

Georg-August-University of Göttingen

GlobalFood Discussion Papers

No. 29

Supermarkets and the Nutrition Transition in Kenya

Ramona Rischke
Simon Chege Kimenju
Matin Qaim
Stephan Klasen

January 2014

RTG 1666 GlobalFood · Heinrich Düker Weg 12 · 37073 Göttingen · Germany
www.uni-goettingen.de/globalfood

ISSN (2192-3248)

Supermarkets and the Nutrition Transition in Kenya

Ramona Rischke[a]*, Simon Chege Kimenju[b], Matin Qaim[b], Stephan Klasen[a]

[a] Chair for Development Economics,
[b] Department for Agricultural Economics and Rural Development

RTG 1666 "GlobalFood", Georg-August-University of Goettingen, www.uni-goettingen.de\globalfood

*Corresponding author, email: ramona.rischke@agr.uni-goettingen.de

Abstract

Many low income countries are experiencing a "nutrition transition" towards the consumption of more energy-dense, highly processed foods and beverages that are often high in caloric sweeteners, fat and salt. Changing lifestyles and urbanisation have coincided with a 'retail revolution', a rapid advance of supermarkets even in remote areas. Among the consequences of the nutrition transition have been expanding waistlines and surging rates of nutrition-related non-communicable diseases, including diabetes, heart diseases and certain cancers. Given the still prevailing rates of under-nutrition, affected countries face a double burden of malnutrition, and individuals that have overcome food poverty risk often remain health-poor. The effect of supermarkets on consumers' diets and the nutrition transition remains unclear: By offering stable and consistent access to a wide range of foods with different dietary qualities, supermarkets could either discourage or contribute to the consumption of a well-balanced diet. This paper investigates the effect of supermarkets on consumption patterns using cross-sectional household survey data collected in Kenya in 2012. In order to establish causality, our sample was designed to be quasi-experimental in nature, with study sites differing in terms of supermarket access. We employ instrumental variable techniques to account for potential endogeneity due to selection effects regarding supermarket purchases. Our findings suggest that supermarket purchases increase the consumption of processed foods at the expense of unprocessed foods. Supermarkets are associated with higher expenditure shares and calorie shares of processed foods, and with increased per capita calorie availability. The latter effect is supported by lower prices per calorie for processed food items. Supermarket purchases have a positive effect on dietary diversity, but implications for the nutrient adequacy of consumers remain unclear.

Keywords: Nutrition transition, supermarkets, food consumption, malnutrition, overweight, obesity.

JEL Classification: D12, I14, I15, I32

Acknowledgements: This research was financially supported by the German Research Foundation (DFG). We are grateful to the University of Nairobi, particularly Dr. Jonathan Nzuma and his students, for support in fieldwork coordination. The authors declare no conflict of interest.

1. Introduction

Many low and middle income countries have been experiencing a nutrition transition, which is understood as a rapid change of diets towards more energy-dense, often (highly) processed and convenience foods and beverages that tend to be rich in fat, caloric sweeteners and salt. In some countries, China and Brazil for example, the onset of these trends was in the mid-1990s already (Popkin, 1997). This "westernization" of diets (Pingali, 2007, p. 4) and a concurrent trend towards more sedentary lifestyles, another component of the nutrition transition, were soon being observed with concern, because they were found to contribute to surging rates of overweight and obesity, which are risk factors for nutrition related non-communicable diseases (NCDs) such as diabetes, cardiovascular diseases and certain types of cancer (e.g. Rosin, 2008; Sturm, 2002). Given still prevailing rates of under-nutrition and related deficiencies, many low income countries are now facing a double burden of malnutrition where under-nutrition and obesity coexist, sometimes even in the same households (Popkin *et al.*, 2012; Roemling and Qaim, 2013).

These nutritional transformations have been associated with changes on both the demand as well as the supply side: changing demand patterns, commonly linked to rising incomes, and urbanisation processes, coincided with a 'retail revolution', a rapid spread of supermarkets (SMs) and fast food outlets. While Mergenthaler et al. (2009) provide case study evidence to suggest demand side factors to predominate, both trends are often believed to be mutually reinforcing (Hawkes, 2008; Popkin *et al.*, 2012; Reardon *et al.*, 2004).

Whereas the concept of epidemiological and nutritional transitions is not new, the high speed at which it is currently observed in low income countries certainly is. Poor health systems exacerbate the situation and add to the vulnerability of low income countries (Popkin, 2004; Schmidhuber and Shetty, 2005). In settings where obesity replaces rather than adds to under-nutrition, found in Brazil for example (Monteiro *et al.*, 2002), there is a hazard for individuals to shift away from food poverty but to remain health poor (Schmidhuber and Shetty, 2005).

In terms of population groups affected, evidence for dietary and health changes, in one case or another, has been found among all socioeconomic groups in both low and middle income countries. At the same time, there is only a limited understanding of which context specific factors determine which socioeconomic groups are most affected (Monteiro *et al.*, 2004; Popkin *et al.*, 2012; Subramanian *et al.*, 2011). Reviewing the evidence, Monteiro *et al.* (2004, p. 943) conclude that "The burden of obesity in a particular developing country tends to shift towards the groups with lower SES [socioeconomic status] as the country's GNP increases."

The consumption of processed and highly processed foods and beverages is often singled out as an important factor contributing to unhealthy diets, as this category includes high calorie foods with only poor micronutrient content, such as sugary beverages, sweets, and all kinds of salted snacks (Monteiro *et al.*, 2010). The spread of supermarkets and fast food outlets, in turn, is argued to improve the availability of these products and to increase their desirability even among poor households in remote areas (Asfaw, 2008; Hawkes *et al.*, 2009). On the other hand, supermarkets could also provide more stable and affordable access to a greater variety of foods and drinks, which might improve the dietary diversity and overall dietary quality of consumers (Asfaw, 2008; Hawkes, 2008).

In any case, supermarkets have the potential for affecting dietary choices to the better or worse, and it is important to better understand the dynamics at play. For this reason, this paper addresses the following research questions:

1. How do supermarkets affect consumption patterns of households?
2. What factors determine where consumers source their food from?

For our empirical analysis, we rely on survey data collected from July to August 2012 in Kenya. Very rich and highly disaggregated food consumption data allow us to analyse consumption patterns with a particular focus on goods associated with the nutrition transition, and at different levels of processing.

Our contribution to the literature is threefold: first, we use detailed data on actual food purchases from different retail formats in addition to measures of physical access which the food environment literature is often restricted to (notable exceptions are Asfaw, 2008; Tessier *et al.*, 2008). Secondly, in contrast to most other studies, we account for potential endogeneity of supermarket purchases (selection effects) using instrumental variable techniques and further improve identification by having a quasi-experimental survey design. Lastly, given the very few studies on this issue in developing countries, we add the first case study of this issue in Sub-Saharan Africa.

For our quasi-experimental design, we chose survey locations among small towns such that they differ in terms of when, if at all, a local supermarket was established, while being comparable in other aspects. While most households in large Kenyan towns have fairly good access to supermarkets, this is not yet true for small towns. Small towns in Kenya (less than 50,000 inhabitants) are of particular relevance also because they accommodate 70% of the urban population, and the manifestation of lifestyle changes are less obvious and less well studied (KNBS, 2010a; KNBS, 2010b).

In a contribution to the non-empirical literature, we provide a detailed account of the current food environment and different retail formats in Kenya and shed some light on the rationale behind consumer decisions and attitudes. This is relevant as it creates a reference point in a highly dynamic market (Neven *et al.*, 2006; PlanetRetail, 2013). In order to understand the potential interactions between the food environment and consumption patterns, we refine a theoretical framework from the literature and apply it to the setting at hand.

Our findings suggest that supermarket purchases increase the consumption of processed foods at the expense of unprocessed foods. Supermarkets are associated with higher expenditure shares and higher calorie shares of processed foods. Contrary to our initial expectation, this is mainly driven by primary processed food. Overall consumption measured in per capita calorie availability at home increases with supermarket purchases. While the food budget share is not significantly affected, prices per calorie are reduced through supermarkets (an effect mainly supported by cheaper prices per calorie in primary processed foods). Supermarket purchases have a positive effect on dietary diversity but further research is needed to assess the nutritional consequences of these shifts. Concerning our second research question, the perception of lower prices and higher product variety motivates consumers to shop in supermarkets, while traditional kiosks (i.e. small corner stalls) still dominate the retail scene for the main reason of being physically much more accessible.

The paper is structured as follows. Chapter 2 comprises theoretical aspects and a literature review. The concept of food environments and different retail formats are introduced, and a theoretical framework is

developed that motivates our hypotheses. Chapter 3 gives a background on the study design and data collection and more thoroughly presents the study context. Chapter 4 discusses our empirical methodology, before we present and discuss our empirical results in chapter 5. Chapter 6 concludes.

2. Theoretical framework and literature review

The term food environment refers to the "[food related] physical and infrastructural features of the area" (Giskes *et al.*, 2011, p. e96) such as access to and the density of different types of retail outlets. There are several potential pathways through which the food environment and supermarkets in particular, can influence consumption patterns, other than by merely making goods available. Before describing these in detail, it useful to elaborate a bit more on the supermarket revolution, and different retail formats. Afterwards, we develop a theoretical framework of impact channels and derive our hypotheses. We will conclude this section by summarising existing case studies.

2.1. Supermarket revolution and outlet types

The retail or supermarket revolution as Reardon *et al.* (2004) describe it, started in the early 1990s in Latin American countries and reached South and East African countries by the late 1990s. The spread of supermarkets, often accompanied or closely followed by spreading fast food chains, and the related introduction of new products, were enabled by favourable political conditions as well as transformations of the food retail system. Upon entering a new market, supermarkets usually target a high income niche, but quickly try to gain market share and expand their product range and customer base to include middle- and lower income households and covering a wider catchment area (Reardon *et al.*, 2004). At the same time, rising levels of incomes, accelerating urbanisation and more modern media coverage created lifestyle changes that ensured a rising demand for all kinds of goods, including higher value, processed and convenience foods and beverages. In sum, unsaturated markets emerged that made it profitable for foreign and increasingly domestic investors to establish or to extend existing supermarkets chains (Hawkes *et al.*, 2009; Popkin *et al.*, 2012; Reardon *et al.*, 2009; Reardon *et al.*, 2004; Schmidhuber and Shetty, 2005).

In Kenya and typical to a low-income country, common alternatives to supermarkets are smaller self-service stores[1] and, more traditionally, kiosks. In terms of overlapping product ranges, at early stages of the supermarket revolution and certainly in small towns in Kenya, supermarkets are not yet competing with other traditional retail formats, most notably open air markets for fresh fruits and vegetables. Table 1 provides an overview of supermarket characteristics and relevant competitors. Several features stand out in particular: supermarkets are self-service stores, while kiosks are strictly over-the-counter shops. Supermarkets, as opposed to kiosks, stock large varieties of different kinds of food and non-food products[2]. This is in terms of product ranges and in terms of brands and features of the same product type, i.e. different flavours, functionalities (e.g. nutrients added to foods) and levels of processing. High value non-food items, such as electronics, furniture and clothes are uniquely offered by supermarkets. Also, only supermarkets stock frozen food and have large (and working) fridges. Generally speaking, the characteristics of small self-service stores lie right in between those of supermarkets and kiosk.

[1] In other studies also called small supermarkets, mini-supermarkets or neighbourhood stores (Neven et al., 2006).
[2] For simplicity, we implicitly include beverages unless stated otherwise.

Table 1: Defining features of different retail outlets – the case of Kenya

	Supermarket	Small self-service store	Kiosk (traditional retail)
Size indicators	**> 150 m²** (Neven and Reardon, 2004) ▪ Typically >1 floor ▪ Typically >2 modern cash counters	< 150 m², though size in small towns typically **10-30 m²** ▪ Typically 1 floor ▪ Typically 0-2 modern cash counters	▪ **1-10 m²** ▪ No modern cash counter
Service features	▪ **Self-service** ▪ **One-stop shopping** ▪ More sophisticated shopping atmosphere: - Spacious isles - Full shelves - Clean & bright ▪ **No credit**	▪ **Self-service** ▪ Narrow aisles, often little light ▪ **No credit**	▪ **Over-the-counter service** ▪ Direct contact to shop owner ▪ **Gives credit**
Product features	▪ Large variety of different food and non-food products ▪ Large variety of brands and features within product categories ▪ Frozen and refrigerated foods ▪ Small to very large packaging sizes ▪ High value non-food items, e.g. electronics, furniture, clothes	▪ Large variety of different food products ▪ Limited variety of non-food products, brands and product features ▪ Neither frozen, nor cooled foods ▪ Small to fairly large packaging sizes ▪ No high value non-food items	▪ Limited but often fair variety of different food products ▪ Only fast-moving non-food products, limited brands and product features ▪ Neither frozen, nor cooled foods ▪ Very small to small packaging sizes ▪ No high value non-food items

Source: Own observation unless stated otherwise.

2.2. Impact channels

Figure 1 illustrates potential relationships between food environments and consumption choices as developed from the reviewed literature.

The basic reasoning behind anticipating an effect of supermarkets on diets is that:

1. The food environment has an impact on where people do their shopping, which in turn influences their dietary practice (Asfaw, 2008) and that
2. Introducing supermarkets significantly alters the food environment.

Note that the term dietary practice does not only refer to what people eat and drink but also to quantities consumed and to eating habits (e.g. meal patterns, snacking) (see Figure 1, column 3). How do we expect supermarkets to influence dietary practices in a low income country setting?

Supermarkets and food access (Figure 1, column 1)

Supermarkets improve access to and increase the availability of goods. By adding stocks to what is in the market already, they could contribute to stabilizing the food supply (in terms of quantities and prices) in settings that are vulnerable to food supply/ food price shocks. This would boost or at least help in smoothening food consumption, especially among poor households. Even though this line of argument brings small scale producers to mind that might suffer from the competition of supermarkets, this is unlikely to play a role in Kenya: in small towns, supermarkets are not yet offering fresh fruits and

vegetables. In rural areas, supermarkets are too far away to distort local prices for perishable goods, and in large towns, for the vast majority, farming is not the main if any source of income.

Note that most goods can be thought of as being accessible even before a local supermarket is established, but some at prohibitively high prices if transport is required to travel to a large town, for example.

Supermarkets and food availability (Figure 1, column 1)

While the concept of food access, strictly speaking, merely refers to physical access, the term 'food availability' more broadly describes what kind of goods are available in a particular outlet.

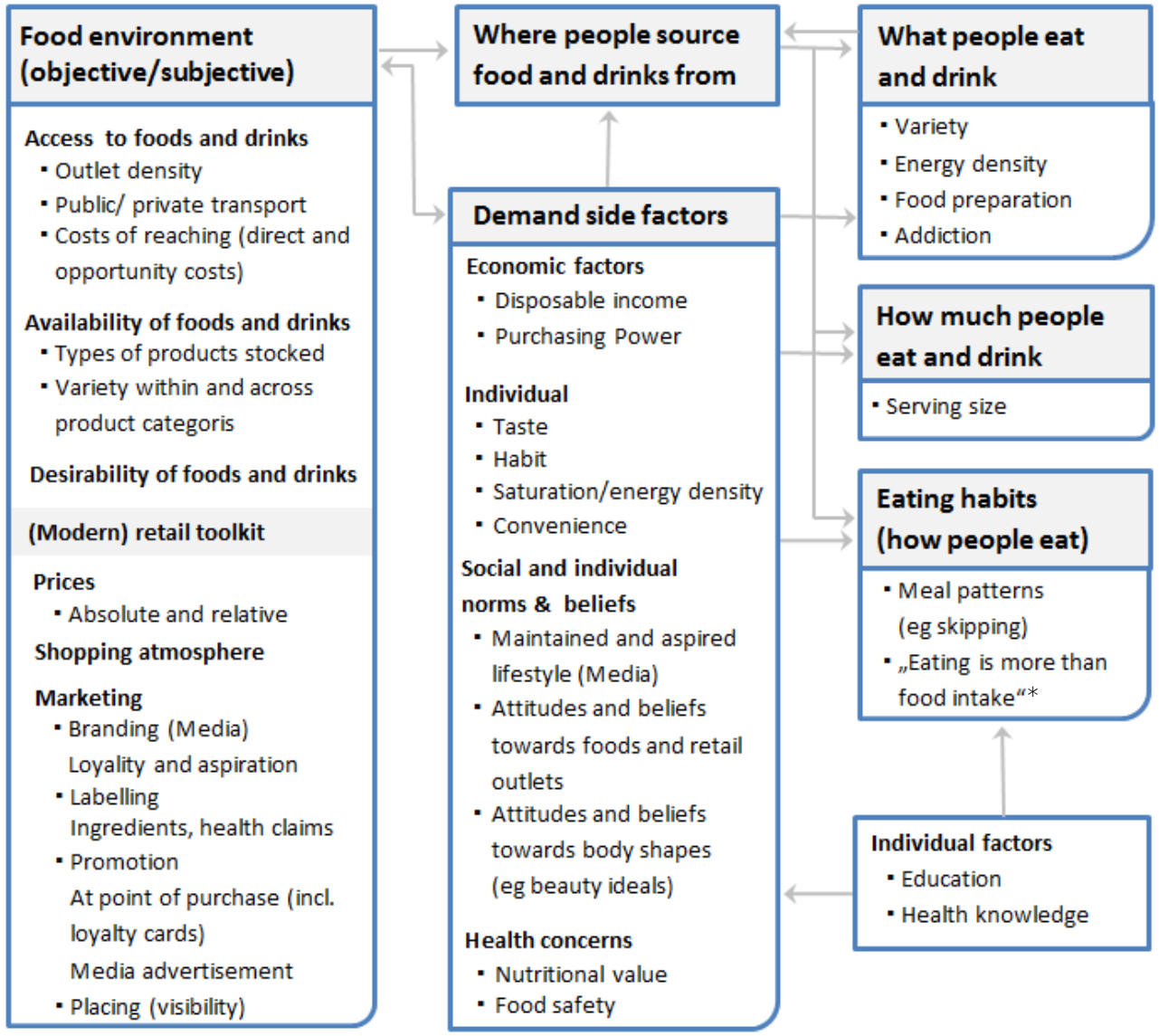
Thus, qualitatively, by offering more types of goods, more brands, flavours, functional foods and levels of processing (e.g. cholesterol free and low fat versions, ready-to-eat, ready-to-heat food) supermarkets offer a larger variety of all types: healthy, 'health neutral' and unhealthy products, regardless of the consumer's dietary needs. This is expected to increase the dietary diversity of consumers. At the same time, changing quantities and substitution within and across food categories could be enhancing as well as deteriorating for dietary quality (Asfaw, 2008; Hawkes, 2008). Thus, the expected magnitude of these effects has to be further elaborated on and will be closely linked to expected effects on relative prices.

Supermarkets, processed foods and relative prices (Figure 1, column 1)

Reardon *et al.* (2004) argue that regarding logistics and owing to economies of scale in procurement and stocking capacities, supermarkets in low income countries have a comparative advantage in offering industrially processed, dry and packed goods with long shelf-lives at relatively lower prices as compared to smaller outlets that need to focus on fast-moving consumption goods. Thinking about processed food in this context, we are neither referring to "minimally processed" foods that are processed "with the purpose of preserving them and making them more available and accessible […]", such as drying, packing or pasteurizing them nor to "processed culinary or food industry ingredients" to use Monteiro *et al.*'s definition (2004, p. 7). The latter includes processing with the aim of extracting or producing ingredients that can still be called raw, such as rice, maize, wheat flour or vegetable oil. Instead, we refer to highly processed food products, which cannot be considered raw but are produced, for example, by "salting, sugaring, baking, frying, deep frying […] canning, and also frequently the use of preservatives and cosmetic additives, the addition of synthetic vitamins and of minerals […]." (Ibid, p. 7). It is the latter category for which supermarkets are expected to have the strongest advantage over other retail formats, as they are relatively slow-moving, especially in low income country markets. In addition, they are more likely to be imported, so that logistical advantages can be used to their fullest. Even though this classification puts products such as flour enriched with vitamins and potato chips in the same processing category, it is food in this category that tends to be high in salt, sugar and saturated fats and is often considered unhealthy. This is because they tend to be energy-dense, while they are often low in micronutrients and dietary fibre (Asfaw, 2011; Monteiro *et al.*, 2004).

There is also scientific evidence that high levels of processing (e.g. ready-to-eat/ ready-to-heat) have negative health effects, in that they contribute to unfavourable eating habits (e.g. snacking) that result in overweight, obesity and nutrition related diseases (Monteiro *et al.*, 2010). Asfaw (2011) further points to medical literature suggesting that energy from processed foods high in concentrated sugar such as corn syrup and fructose is easily absorbed by the body and that these foods can contribute to addictions by altering the body's hormonal balance.

Figure 1: Conceptual framework - food environment, consumption and influencing factors



Source: Own illustration based on literature review. *Chandon & Wansink (2012, p.583)

Because prices are arguably one of the most important drivers of consumption decisions, especially against the background of past food price hikes, it is crucial to think about implications for relative prices. Chandon and Wansink (2012, p. 572) make the point that highly processed food products are highly differentiated, branded products which are less bound to commodity prices because: "With these branded products, marketers can establish their own price depending on which consumer segment they wish to target." As an example to the contrary, Popkin *et al.* (2012) point out that advances in the production of edible oil led to price reductions that had already by the mid 1990's enabled even poor households in low income countries to increase their energy intake. Reviewing existing evidence on pricing strategies and supermarkets in low income countries, Hawkes (2008) finds that supermarkets tend to be more expensive upon market entry but become more price-competitive later, and first among processed foods as discussed above.

Independent of relative prices and qualities, by offering large packaging sizes in addition to small ones, supermarkets facilitate bulk shopping and likely offer quantity discounts. Especially poor consumers, however, cannot utilise these discounts. In fact, one advantage of kiosks is that they often offer credit as well as smallest amounts of products, even though economic theory suggests that the latter results in higher unit prices.

What should we take away from this? First, processed and highly processed products come in a wide range of nutritional qualities and prices. For analysing the impact of supermarkets and retail environments, it is imperative to carefully differentiate not only between different food categories but processing levels also, which we intend to do in this paper. Second, there is no conclusive or better uniform evidence on supermarkets and pricing strategies. Price premiums have been detected in some cases (Schipmann and Qaim, 2011) and consistently smaller prices in others (Hawkes, 2008). Because we assume that supermarkets in our study context, small towns with a limited catchment area, need to target a broad customer base, we hypothesize that they offer a wide range of product qualities and prices, including the smallest and highest prices for key items consumed. Third, because supermarkets in general do seem to have a price advantage with highly energy-dense products, we conclude that they are likely to increase the calorie availability and hypothesize that the price paid per calorie is decreasing with a higher share of processed items consumed.

Supermarkets and marketing strategies (Figure 1, column 1)

How could supermarkets affect consumption other than by influencing relative prices? There are a number of studies and review articles stressing the role of marketing strategies (by retailers as well as food manufacturers) on consumption decisions, to a fair extent even without consumers consciously noticing. Monteiro *et al.* (2010, p. 12) go as far as commenting that "[the] idea that eating and drinking behaviours are simply a matter of conscious choice that can be educated is fundamentally wrong". Hawkes (2008, p. 682) is more illustrative by pointing to the concept of "food desirability". She argues that supermarkets and the food industry are making food products desirable by strategically using marketing strategies such as media advertisement or point-of-sale promotions (e.g. discounts, freebies of certain brands). Chandon and Wansink (2012) reviewed studies on how marketing strategies influence what and how much people eat and provide a comprehensive overview, even though the study settings are not always clear: some studies under review highlight the importance of factors such as where products are placed (e.g. eye height) and how they are displayed, which certainly falls in the category of subconscious factors. Others, more surprisingly, show that temporary price discounts can increase the consumption of respective goods

(rather than merely shifting consumption across brands or time). Offering larger quantities of a good (e.g. multipacks) can equally increase consumption, even if unit prices are kept constant. This is particularly relevant since, as we have shown before, supermarkets are stocking larger packaging sizes than their counterparts. Following this line of argument, supermarkets are hypothesized to increase overall consumption of all food groups (Hawkes, 2008).

Other demand side factors (Figure 1, column 2)

There are a number of demand side factors that can have a direct influence on both, dietary practices as well as where people do their shopping. These include: economic factors (e.g. disposable income), individual and household preferences (e.g. for taste or habits), social and individual norms and beliefs (e.g. attitudes towards modern or traditional foods and outlets, the maintained and aspired lifestyle and beauty ideals) and personal health concerns. We will incorporate proxies for them as control variables in the empirical analysis.

2.3. Case study evidence and hypotheses

Most empirical studies on supermarkets and the food environment have been carried out in high income countries. In this context, most authors assume that supermarkets improve access to healthy foods, such as fresh fruits and vegetables (FFV), improve the dietary quality of consumers and help in reducing overweight and obesity. Empirical results to this are mixed (e.g. Cummins et al., 2005; Laraia et al., 2004; Moore et al., 2008; Morland et al., 2006; Pearce et al., 2008; Powell et al., 2007; Wrigley et al., 2003). We have already seen that it is other types of products, namely processed and highly processed foods and beverages, which are of concern in low income countries, thus the expectations are certainly different. At the same time, it is worthwhile to bear in mind that the healthiness of diets is a relative concept that depends on the context and the health risk being looked at (e.g. thinking of sugar-free drinks high in artificial sweeteners or food safety concerns).

Two studies were conducted in a developing country context, which further contribute to the literature by actually considering supermarket purchases rather than supermarket access. Asfaw (2008) finds that supermarket purchases in Guatemala increase the share of partially and highly processed foods at the expense of staple foods and that supermarket purchases are positively associated with BMI. Tessier et al. (2008) in a similarly titled paper conclude that regular users of supermarkets in Tunis have a slightly improved dietary quality.

Thus, existing studies confirm that the impact of supermarkets on diets is context specific in nature and that important research gaps remain with respect to mediating factors. Following the discussion of the previous chapter, in the case of Kenya , we hypothesize that:

- $H1_1$: Supermarket customers eat differently: supermarket purchases increase per capita consumption shares of processed and highly-processed foods.
- $H1_2$: Supermarket customers eat more: supermarket purchases increase total per capita consumption.
- $H1_3$: Supermarket customers eat more types of food: supermarket purchases increase the dietary diversity of consumers.

3. Study Design and Descriptive Statistics

3.1. The case of Kenya

Supermarkets have been spreading rapidly throughout Kenya and the pattern has been similar to the retail revolution described in other low income countries (Neven *et al.*, 2006; Reardon *et al.*, 2004). In the early 2000s, Kenya's retail sector was already classified as one of the most dynamic in Sub-Sahara Africa (Neven *et al.*, 2006). Today, despite being highly fragmented, it is among the most developed retail sectors in Sub-Sahara Africa (PlanetRetail, 2013). This fragmentation explains why the top three retailers in 2013 only had a market share of around 5% while in 2003 already, supermarkets more generally had a 20% market share of the urban food retail market (Neven and Reardon, 2004; PlanetRetail, 2013). Interestingly from a domestic policy perspective and somewhat different from what early supermarket revolution countries experienced (Reardon *et al.*, 2004), none of today's top five supermarket chains[3] in Kenya are owned by international corporations or foreign firms, but by Kenyan enterprises. It should also be noted that quite a number of supermarkets do not belong to chains at all or have only a few outlets, however they do not qualitatively differ from chain supermarkets.

A Demographic and Health Survey conducted in 2008/09 revealed rates of overweight and obesity among women between the ages of 15-49 to be 25%, while 12% were underweight, attesting the relevance of investigating this double burden of malnutrition. The same survey found 30% of children below the age of 5 to be stunted and 6% wasted (KNBS and ICFMacro, 2010). Reliable nation-wide prevalence data for women and men of all ages, however, are not available.

Semi-structured (not recorded) interviews with representatives from the supermarket chains, the Ministry of Health, nutritionists and consumer groups confirmed the presence and relevance of consumption and lifestyle changes that are characteristic to the nutrition transition. Rising rates of overweight, obesity and non-communicable diseases, especially diabetes, were identified as pressing health concerns. In the case of diabetes, awareness campaigns had already been rolled-out with the help of international donors.

Just as lifestyle and consumption patterns differ between large and small towns so do target groups and product ranges of supermarkets and competing retail outlets, and it helpful to be aware of the differences: international and other fast food chains are only present in large towns. Also in large towns only do supermarkets offer fresh fruits and vegetables, have built-in butcheries, restaurants and large bakeries. Western style convenience processing (pre-cut vegetables, prepared salads, frozen or tinned ready-to-heat food) are only available here. Visiting large town supermarkets or hypermarkets which are ten times larger in size (Neven *et al.*, 2006), it becomes evident that lifestyle and status play a significant role and that shopping atmosphere is not an abstract concept but a strong force one can hardly escape from. However, Neven *et al.* (2006), who analyse patterns of the retail revolution in Kenya and consumer attitudes in Nairobi, already put forward that the introduction of supermarkets in small towns, from a consumer perspective, is likely to be as impressive and as powerful in influencing consumer choices, as the introduction of hypermarkets in large towns or mini-supermarkets in rural areas. Product ranges of supermarkets, small self-service stores and kiosks in small towns are surprisingly similar (see Table A 1 in

[3]Nakumatt, Tuskys, Uchumi, Naivas, and Ukwalla

Appendix for a detailed account). The main differences are qualitative in nature and as outlined in the previous chapter. It is mainly frozen and some cooled products that supermarkets add to the table.

3.2. Data collection and identification strategy

This study uses data from a household consumption survey that was conducted in three small towns in Central Province, Kenya. Fieldwork took place from May to September 2012, with interviews being conducted between July and August. A total of 453 households were interviewed.

Our identification strategy to test for a causal relationship between supermarkets and consumption patterns relies on a selection of our study sites that aimed at being quasi-experimental in nature: we selected three towns that differ in terms of their access to supermarkets while being comparable in other aspects:

1) One with a well-established supermarket (Ol Kalou: one supermarket since 2002),
2) One with a supermarket opened fairly recently but with a sufficient time lag to allow inhabitants to get used to it (Mwea: one supermarket since August 2011) and
3) One town with no supermarket up to that point in time.

We used a systematic random sampling strategy based on a sampling frame produced for this survey. Our sampling area included the town centres and close peripheries (about 2.5 km radius), which covered the core and peri-urban areas, and in each case the most densely populated parts of the town outskirts.

Due to interview non-participation, we were forced to replace 22% of households initially selected. This was mostly for the reasons of interview partners being busy/ not found at home or having a lack of interest. A number of cases also resulted from households that had just moved to our survey sites (less than 6 month ago), which we decided to replace reasoning that consumption patterns could still reflect the food environment of the former rather than the current place of residence. We tried to avoid introducing selection bias to the best of our abilities. Using a dummy for replacement households in robustness checks never turned out significant.

In sum, while we cannot rule out limitations of our sampling strategy and potential biases resulting from non-response, we are confident that our sample is representative of the surveyed towns and towns of equal size more generally.

4. Methodology

4.1. Empirical strategy

In general terms, our model can be specified as proposed by Asfaw (2008):

$$\boldsymbol{D}_i = \propto \boldsymbol{X}_i + \beta S_i + \varepsilon_i \quad (1)$$

$$S_i = \gamma \boldsymbol{X}_i + \delta \boldsymbol{Z}_i + \omega_i \quad (2)$$

where $\boldsymbol{D}_i$ refers to dietary indicators of household *i*, $\boldsymbol{X}_i$ to explanatory variables and S_i to the measure of supermarket purchases, our main variable of interest. Because supermarket purchases are likely to be

endogenous, we use an instrumental variable approach and thus add equation (2) to the model, where $\boldsymbol{Z_i}$ refers to the excluded instruments used. ε_i and ω_i are error terms.

Supermarket purchases are conceptualized using the share of supermarket purchases from the overall food basket. Note that this share can be positive for non-supermarket locations because out-of-town shopping is possible.

The potential sources of endogeneity are related to self-selection on non-observables, i.e. systematic differences between frequent supermarket customers and those who do not or only rarely shop in supermarkets. As far as instruments are concerned, we use distance to the nearest supermarket. This reflects our initial hypothesis that supermarket access will induce people to shop there. At the same time, we claim distance to supermarkets, or supermarket locations to be exogenous: while investors' decisions regarding supermarket locations is driven by market potential, we argue that this potential boils down to demand side factors, which we control for, and to road infrastructure so as to facilitate logistics. We reason differences in road infrastructure to be exogenous to consumption patterns: while the main roads within our survey sites are paved and in fairly good shape, the roads connecting our survey locations and neighbouring towns differ in quality. In fact, the supermarket in Mwea, the largest town in our sample, happened to be introduced shortly after the inter-town road infrastructure was greatly improved. On the other end, Njabini, the non-supermarket location, has the worst inter town connections. Within our survey sites, supermarket managers themselves pointed out that the specific location was mainly driven by the availability of a large plot for construction and thus is exogenous for our analysis.

Distance is measured as physical linear distance between the household and nearest supermarket based on GPS readings. Note that there is only one supermarket per location (town), and consumers mostly go there by foot. For the town without a supermarket, the closest supermarkets can only be reached using public or private transport. Linear distances approximate walking distances in our supermarket sites well.

Our explanatory variables mirror the demand side and individual factors from our conceptual framework presented earlier (see Figure 1). Individual level factors, such as education or age, refer to either the household head or to the person responsible for food purchases and preparation.

4.2. Variables of interest – consumption patterns

Food consumption was captured with a 30 day recall period because we expect decisions regarding where to shop to vary during a wage cycle (e.g. households shopping in bulk in supermarkets after getting paid while increasingly shopping for small portion sizes at kiosks towards the end of the month).[4] In very disaggregated form (e.g. differentiating between fortified and unfortified flour and different types of cooking oil), we asked how much quantity was consumed by the household during the last month. This was for consumption inside the house, since food eaten outside the home is more specific to the individual and usually not sourced from supermarkets, but from street hawkers, restaurants and sometimes kiosks. For each good consumed inside the house, we asked for the quantity that was consumed from purchases, own production, and other sources (e.g. gifts). For purchases, the respondents

[4] While being aware that what we call food consumption is merely food availability and different from actual food intake (e.g. due to wastage), we use the terms interchangeably.

were additionally reporting how much they spent. Further, they indicated what quantity of these purchases they bought in supermarkets, smaller self-service stores and traditional retail outlets (i.e. all other outlets, e.g. kiosk, butchery, open air market). Because outlets falling in the latter category only have few overlapping products, we can still and most notably identify the consumption that came from kiosks. Monetary values for own production and other sources are imputed so as to include it in the food expenditure aggregate. For this, we use median unit values reported for the same good by neighbouring households. The expenditure share of a particular retail outlet is from the total food expenditure of that household.

For capturing general demand patterns, we differentiate products by levels of industrial processing. For the majority of products, we follow the classifications used by Asfaw (2011) and Monteiro et al. (2010) into unprocessed foods, such as fresh fruits and vegetables, primary processed foods including rice, sugar and cooking oils, and highly processes foods such as breakfast cereals, bread and sweets. All foods and drinks are classified into these mutually exclusive categories, with the exception of alcoholic beverages, which are excluded here.

We conceptualise consumption patterns by expenditure and calorie shares on different types of foods, i.e. different levels of processing. Overall consumption is considered in terms of total calories consumed and we will consider households' food budget shares also.

5. Empirical results

5.1. Descriptive Statistics

Town characteristics

The survey locations differ quite substantially in terms of their size: Njabini is the smallest and least urbanised town with an estimate of 1870 households (estimate based on our sampling frame). Mwea on the other end is the largest town with an estimate of four times the number of households (7650 households). Still, in terms is physical and social infrastructure (e.g. main roads being tarmac roads, having access to banks, a hospital, several health centres and other services, having similar administrative structures), all survey locations are comparable. In terms of ethnicity and religion, Kikuyu and Christian are by far the most prevalent in all survey towns, with rates exceeding 80% and 90%, respectively.

Household characteristics

Table 2 summarizes household characteristics by our survey locations. The sample size across the survey locations is fairly similar, ranging from 134 to 161 households[5]. The average number of household members in Njabini exceeds that of the other locations by one additional member. Three quarters of all households in the sample are male headed. The age of the head is around 38 years on average, with significant differences for Ol Kalou (younger heads) and Njabini (older ones). Despite having older heads, Njabini seems to be lagging behind in regards to the highest level of education of the household head.

[5] Five observations were excluded from the initial sample for the reason of unrealistic consumption figures.

Table 2: Household characteristics of sample

	All	Njabini (no SM)		Mwea (SM since 2011)		Ol Kalou (SM since 2002)	
	mean	mean	diff to others	mean	diff to others	mean	diff to others
Household size	3.63	4.28	1.01***	3.14	-0.70***	3.38	-0.38**
	(1.93)	(2.38)	(0.18)	(1.44)	(0.20)	(1.57)	(0.19)
Male head (%)	0.74	0.77	0.05	0.69	-0.06	0.74	0.00
Monthly p.c. exp. (food + non-food) in KSh	9425.15	8105.58	-2059.81***	10415.12	1412.44*	9946.68	792.02
	(7995.69)	(8788.48)	(782.13)	(6840.21)	(823.26)	(7923.59)	(796.61)
Age of head	37.51	40.61	4.84***	36.87	-0.91	34.80	-4.11***
	(13.01)	(14.21)	(1.26)	(12.37)	(1.34)	(11.56)	(1.28)
Education of head completed							
No formal educ.	0.03	0.06	0.04**	0.01	-0.02	0.02	-0.02
Primary	0.38	0.48	0.16***	0.32	-0.09*	0.33	-0.08
Secondary	0.38	0.30	-0.11**	0.44	0.09*	0.39	0.03
Tertiary	0.21	0.16	-0.09**	0.22	0.02	0.25	0.07*
Observations	448	161	161	134	134	153	153

Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.
Source: Own calculation.

Average monthly per capita expenditure amount to 9,425 KSh, while being significantly and quite substantially smaller in Njabini. We are not aware of an up to data poverty estimate, but based on the latest poverty line available to us (year 2005) and subsequent consumer price statistics that are publicly available, we extrapolate today's poverty line to be around 7,500 KSh per capita per month. This would yield a poverty headcount of 47% in our sample. The latest poverty estimate according to World Bank statistics is from 2005 also and around 46%.

Shopping behaviour and relative prices

Table 3 provides an overview of access to different retail outlets and shopping behaviour: in our supermarket locations, the distance to the local supermarket is below 1km on average, while the nearest supermarket is 40km away from Njabini. For each household in our sample, traditional retail outlets (i.e. kiosks, open air markets) are very close, with reported travelling times of maximum 12 minutes. In fact, for most households, the closest kiosk can be reached much faster, within 1-3 minutes by foot.

The food expenditure shares, i.e. the value of last month's food consumption from different retail outlets paint the expected picture: across towns, Ol Kalou has the highest food expenditure share from supermarkets, followed by Mwea and Njabini. In Ol Kalou, the average supermarket share (of food expenditure) is 17%, while in Mwea already 11% worth of food expenditure is spent in supermarkets. Even in Njabini, the mean supermarket share is positive despite being only 2%, and 14% of households reported positive food purchases from supermarkets during the last month. In Ol Kalou, 84% of households frequented the supermarket, 80% in Mwea. Interestingly, in all towns, the frequency of shopping in kiosks is very high, it does not vary much from the overall mean of 25 times last month and

traditional retail is by far the most important source for food with expenditure shares ranging from 66% to 75% across towns.

Table 3: Access to retail outlets and shopping behaviour

	All mean/sd	**Njabini (no SM)** mean/sd	**Mwea (SM since 2011)** mean/sd	**Ol Kalou (SM since 2002)** mean/sd
Number of times shopping in [...] last month				
Supermarket	3.05	0.36	2.70	5.77
	(5.36)	(0.98)	(3.27)	(7.46)
Small self-service store	2.50	4.08	0.53	2.71
	(5.73)	(8.44)	(1.91)	(3.66)
Kiosk	25.62	23.84	29.33	24.18
	(16.82)	(17.69)	(15.78)	(16.38)
Distance to SM in km	14.55	39.29	0.67	0.68
	(20.44)	(14.35)	(0.49)	(0.41)
Travelling time to [...] (min. one way)				
Supermarket	47.64	103.68	16.54	15.90
	(47.29)	(33.73)	(9.08)	(10.59)
Kiosk	5.33	8.30	2.95	4.31
	(5.82)	(7.58)	(2.73)	(4.15)
Traditional retail (av. to kiosk & market)	9.21	12.00	7.46	7.80
	(7.74)	(9.69)	(5.84)	(5.88)
Share of HHs buying in supermarket	0.58	0.14	0.80	0.84
Expenditure shares in [...]				
Supermarket	0.10	0.02	0.11	0.17
	(0.12)	(0.06)	(0.10)	(0.13)
Small self-service store	0.05	0.08	0.02	0.05
	(0.11)	(0.13)	(0.10)	(0.08)
Traditional retail	0.70	0.71	0.75	0.66
	(0.19)	(0.20)	(0.17)	(0.17)
Own production	0.11	0.16	0.08	0.09
	(0.15)	(0.17)	(0.13)	(0.13)
Observations	448	161	134	153

Expenditure shares don't add up to 100% because left out category gift and other sources.

Source: Own calculation.

Asked for the most important reasons to shop in different retail outlets, more than half of the respondents in supermarket locations indicated that perceived lower prices are their main motivation to shop in supermarkets (see Table 4). Improved availability, e.g. more variety of food and non-food products or packaging sizes, were reported by 16% of respondents in Ol Kalou and 8% in Mwea. The possibility for one-stop-shopping, along with other factors that we group under convenience, such as self-service and the possibility to read labels were identified as the main motivation to shop in SMs by 11% of respondents in SM locations. For kiosks on the other hand, physical access was by far the most important reason, ranging from 52% in Njabini to 69% in the other towns. Looking at the second and third reasons reported (not shown) this picture does not change dramatically. Note that the importance of perceived lower prices in supermarkets and physical access in the case of kiosks is consistent to what Neven et al. (2006) found in the case of consumers in Nairobi.

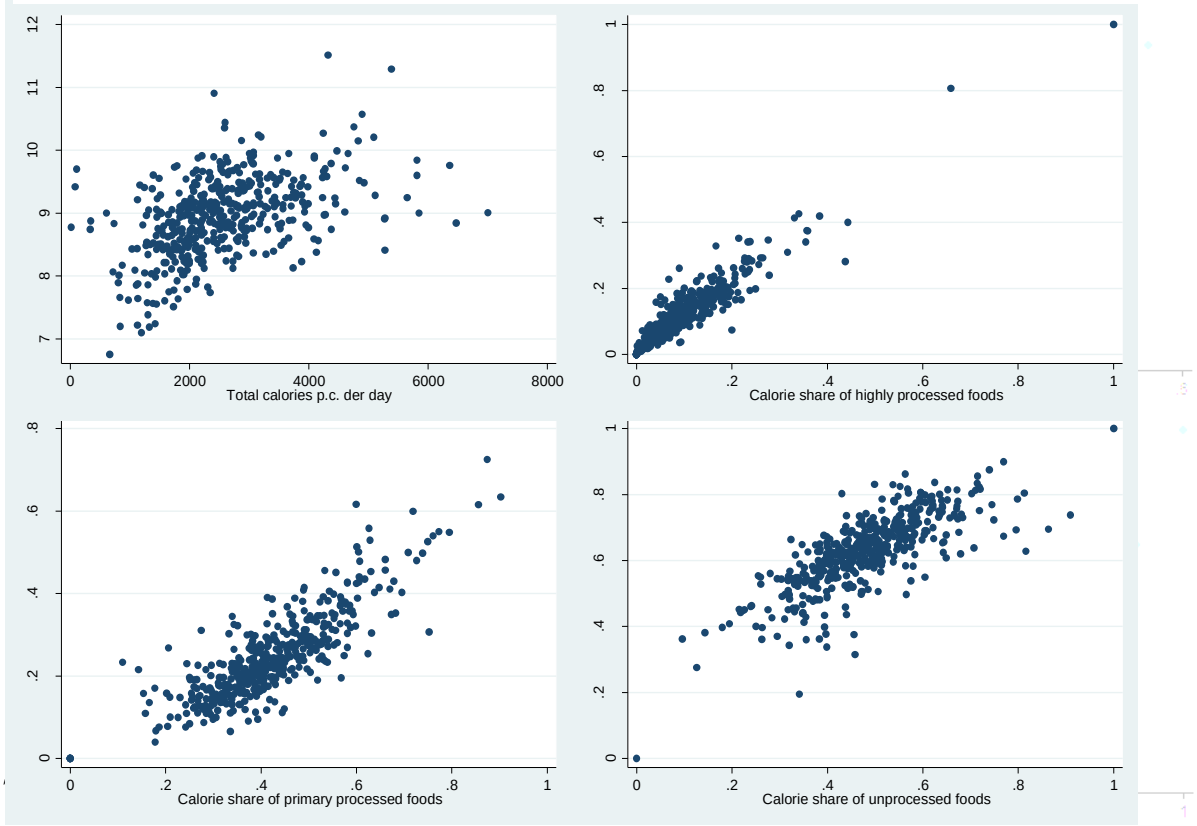
Table 4: Reasons for shopping in LSM/ kiosk stores were uses during last month

	All	Njabini (no SM)	Ol Kalou (SM 2002)	Mwea (SM 2011)
	mean	mean	Mean	mean
Most important reason for shopping in SM				
Not applicable (doesn't shop there)	0.42	0.85	0.22	0.15
Economic (e.g. lower prices)	0.37	0.07	0.54	0.53
Improved availability (e.g. more variety)	0.09	0.02	0.08	0.16
Convenience (e.g. one-stop-shopping)	0.07	0.00	0.11	0.11
Physical access (e.g. close to home/ work)	0.02	0.04	0.01	0.01
Social (e.g. meet people, talk to staff)	0.01	0.00	0.01	0.03
Most important reason for shopping in kiosk				
Not applicable (doesn't shop there)	0.04	0.04	0.06	0.03
Economic (e.g. lower prices)	0.12	0.22	0.06	0.06
Improved availability (e.g. more variety)	0.06	0.04	0.13	0.04
Convenience (e.g. one-stop-shopping)	0.06	0.08	0.03	0.07
Physical access (e.g. close to home/ work)	0.62	0.52	0.68	0.69
Social (e.g. meet people, talk to staff)	0.05	0.07	0.01	0.06
Observations	448	161	134	153

Source: Own calculation.

Comparing price ranges across outlets (not shown) cannot easily support the perception of generally lower prices (per kg) offered by supermarkets. Irrespective of quality differences, most price ranges do not seem to differ much across stores. We will return to this issue below, but at this point the question remains, if food expenditure serves as an interesting indicator of food consumption given that they might reflect price differences. Figure 2 plots several expenditure indicators against calorie indicators, both variables using adult equivalence scales and including not only purchased items but also own production and other sources (for home consumption). The first plot could suggest a non-linear relationship between p.c. expenditure and calorie availability, which is particularly pronounced once calorie availability becomes very large and likely to exceed actual p.c. calorie intake (e.g. due to food wastage or hosting guests). Plotting food expenditure shares against calorie shares for different levels of processing (remaining plots) reveal strong positive and rather linear relationships so that both indicators seem to capture the same aspects of food consumption and are thus interesting for further analysis. Only in the case of highly processed foods, however, are expenditure and calorie shares so close to each other in absolute terms. Note that 5% of households do not report any consumption of highly processed foods (median 12%). In the case of primary processed foods, median expenditure shares are 23%, calorie shares 40%, suggesting that indeed, prices per calorie are lowest in this food category. Unprocessed foods contribute around 47% of calories for the median consumer, while 63% of food expenditure is spent on these items.

Figure 2: Expenditure and calorie indicators



Source: Own calculation.

Empirical results

Food expenditure shares by levels of processing

Table 5 displays main empirical results with respect to expenditure shares on products at different levels of processing using OLS and IV specifications. Summary statistics of all variables used, first stage results and robustness checks are found in the Appendix. Robust standard errors are used in all specifications. We tested each model for cluster effects at the neighbourhood level, our primary sampling unit, and use cluster robust standard errors whenever required. Note that all IV specifications reported in this paper have first stage test statistics with respect to exclusion and weak instrument criteria meeting or well exceeding conventional thresholds.

The OLS results in Table 5 confirm our initial expectations: supermarket purchases are positively associated with expenditure shares of highly as well as primary processed foods, while the share of unprocessed foods is declining. Once we account for endogeneity in supermarket purchases using two-stage least squares IV models, supermarket purchases lose their significance in the case of highly processed foods, while remaining significant in all other cases. At the same time, correcting for endogeneity changes the effect size of supermarkets in some cases, which we take as an indication that endogeneity is a relevant issue here that is controlled through the IV approach. What effects are remaining? While the effect on highly processed foods disappears, supermarkets have an even stronger effect in the case of primary processed foods, which is carried over to the effect on all processed foods where the point estimate increases from 0.21 to 0.38.

Table 5: OLS and IV regression results – Food expenditure shares by levels of industrial processing

	(1) OLS Expenditure share *highly processed* food	**(2) IV** Expenditure share *highly processed* food	**(3) OLS** Expenditure share *primary processed* food	**(4) IV** Expenditure share *primary processed* food	**(5) OLS** Expenditure share *all processed* food	**(6) IV** Expenditure share *all processed* food	**(7) OLS** Expenditure share for *unprocessed* foods	**(8) IV** Expenditure share for *unprocessed* foods
SM expenditure share	0.0766*	0.0712	0.1336***	0.2109**	0.2134***	0.3781***	-0.2127***	-0.3220***
	(0.041)	(0.091)	(0.039)	(0.086)	(0.041)	(0.101)	(0.046)	(0.077)
Ln p.c. expenditure	0.0225***	0.0227**	-0.0829***	-0.0863***	-0.0595***	-0.0668***	0.0313**	0.0361***
	(0.008)	(0.010)	(0.009)	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)
Household size	-0.0009	-0.0009	0.0062	0.0062	0.0045	0.0044	-0.0141***	-0.0141***
	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
=1 if head is married	-0.0228**	-0.0228**	-0.0089	-0.0089	-0.0313***	-0.0314***	0.0412***	0.0413***
	(0.009)	(0.009)	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Education of head in	0.0041***	0.0041***	-0.0009	-0.0014	0.0032**	0.0021	-0.0016	-0.0009
years	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook	-0.0061***	-0.0061***	0.0002	0.0003	-0.0055***	-0.0053**	0.0055***	0.0054***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook squared	0.0001***	0.0001***	0.0000	0.0000	0.0001**	0.0001**	-0.0001**	-0.0001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does	-0.0346***	-0.0347***	-0.0243**	-0.0224**	-0.0609***	-0.0569***	0.0702***	0.0675***
farming	(0.008)	(0.008)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
Other controls[6]								
	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.0462	0.0445	0.9562***	0.9810***	0.9955***	1.0487***	0.2164**	0.1812*
	(0.079)	(0.090)	(0.077)	(0.084)	(0.090)	(0.094)	(0.101)	(0.099)
Observations	448	448	448	448	448	448	448	448
R^2	0.256	0.256	0.316	0.310	0.233	0.208	0.240	0.229

Standard errors in parentheses. Robust (1),(2),(5),(6) and cluster robust (3),(4),(7),(8) standard errors used. * p < 0.10, ** p < 0.05, *** p < 0.01.

Source: Own calculation

[6] See Table A 3 - Table A 6 in Appendix for full models.

How are these coefficients to be interpreted? If the supermarket expenditure share increased by 1 percentage point (the average share is 9%), the expenditure share on processed foods would increase by 0.38 percentage points. However, considering that the average share in our supermarket locations is 14% against 1% where no SM is present, looking at a 10 percentage point increase in purchases does not seem implausible, and would be associated with a 3.8 percentage point increase in expenditure shared on processed food (an increase from 34 to around 38% for the average consumer in the non-SM location).

The coefficients of total per capita expenditure and the education of the household head, which we take as measures of disposable income and status respectively, suggest positive income effects regarding highly and unprocessed foods, and negative income effects with respect to primary processed food. Note that these effects include quality effects of unknown magnitude. The variable for household heads being married, and especially the age and age squared of the person responsible for food purchases and cooking can be thought of as capturing information on traditional values and they have the expected sign.

Robustness checks include testing different sets of control variables, and restricting the sample to the supermarket locations only. The most insightful results are shown in the Appendix (see Table A 3 - Table A 6). Generally we find the direction of main effects and their statistical significance to be robust, but effect sizes are sensitive to model specifications. For the case of processed food expenditure shares, for example, point estimates of SM purchases considering the whole sample range from 0.18 in OLS to 0.39 in IV specifications. Interestingly, for all expenditure shares, the effects remain stable when excluding our non-supermarket location from the sample. Another interesting finding regards interaction effects that we find between supermarket shares and a dummy indicating if the households' kiosk consumption exceeds the town average. The reasoning behind expecting an effect is that depending on their shopping intensity in traditional retail outlets, households might frequent supermarkets for different reasons and with different outcomes. Indeed, in the case of primary and all processed foods, controlling for frequent kiosk consumption increases the effect of supermarket purchases, but less among frequent kiosk consumer. It is the other way around for unprocessed foods. Note, however that the interaction effects should be interpreted with care because first, frequent consumers tend to have lower supermarket expenditure shares to begin with and kiosk purchases might be subject to same kind of selection effects that we expect in the case of supermarkets. Other interaction effects with measures of income or education, for example, were not found to be significant.

Calorie consumption

Turning to the models on calorie shares, supermarkets have less pronounced effects than before (see Table A 7 in Appendix). A significantly positive relationship between supermarket purchases and calorie shares remains significant over both OLS and IV specifications in the case of all processed foods only. The direction of all other effects is as expected but mostly insignificant. In accordance with our previous findings, negative income effects are found for primary foods and positive for both highly and unprocessed foods. Considering the robustness checks in Table A 8 (see Appendix), the effect size of supermarket purchases on calorie shares from processed food ranges from 1.1 percentage point in OLS to 2.2 percentage points in IV specifications given a 10 percentage point increase in supermarket purchases. The average household in our non-supermarket location consumed 49% of their calories from primary processed food. Again, the effect size rises when we include a dummy for frequent kiosk consumers and interactions term with supermarket purchases but in this case, for high frequency consumer, the effect of supermarket almost cancelled out.

What do we take away up to this point? Supermarkets indeed influence consumption patterns in that they are associated with higher consumption shares of processed foods (incl. beverages). This is in terms of expenditure as well as calorie shares of these goods and at the expense of unprocessed foods. These results partly confirm our hypothesis 1. The contradicting part concerns highly processed foods, where we expected stronger and significant effects of supermarkets.

In order to address our second hypothesis that supermarket purchases would be associated with higher overall consumption, we analyse per capita calorie availability. For the reason of a high standard deviation in daily per capita calorie availability (see Table A 2 in Appendix), we use the log of p.c. calories in our regressions. This produces more robust results as compared to using absolute values, while the results remain sensitive to model specifications. Table 6 presents our main results. We find supermarkets to be positively and significantly associated with higher per capita calorie availability so that we cautiously confirm our hypothesis. In the OLS specification, the semi-elasticities indicate that per capita calories increase by 0.37% upon a 1 percentage point increase of supermarket purchases. In case of our example used before, a 10% increase in supermarket purchases would increase calories by 3.7% or 90 calories per capita per day in the case of an average consumer in the non-supermarket location. Models (4) and (6) again show a significant interaction between frequent kiosk consumers and supermarket purchases. Above median kiosk purchases are associated with higher calorie availability while supermarket purchases among frequent kiosk consumers have a negative effect on calories. Effect sizes of supermarket purchases are higher in the IV as compared to the OLS specification. This might reflect measurement errors in calories consumed which would bias OLS results towards zero if they are random. IV techniques account for such random measurement errors.

The finding that supermarket purchases are associated with higher calorie availability is interesting in itself. However, it is worthwhile to investigate further demand effects: since overall calorie consumption is significantly increased at constant overall expenditure levels, we expect households either to spend a higher proportion of their expenditure on food, or to source calories at cheaper prices. Note that this concerns prices per calories and not prices per physical unit (kg) as discussed earlier. In fact, we cannot find significant effects of supermarket purchases on the food budget share (controlling for income, see Table A 10 in Appendix). Prices per calories however, are indeed significantly negatively affected by supermarket purchases in the IV specifications, which are much more reliable in this case because of reversed causality between prices and expenditure shares by construction.

It is not straightforward to assess the consequences of these findings on overall dietary quality in terms of nutrient adequacy. Both, the effect on changing consumption shares as well as changing calorie availability are driven by primary processed rather than highly processed foods and primary processed foods in turn are characterised by relatively low prices per calorie.

Table 6: OLS and IV regression results – Calorie availability at home

	(1) OLS log of *per capita calories* per day	(2) IV log of *per capita calories* per day	(3) 1st stage SM expenditure share	(4) OLS log of *per capita calories* per day	(5) IV log of *per capita calories* per day
SM expenditure share	0.3706**	0.8485*		0.9140***	1.2479*
	(0.186)	(0.504)		(0.291)	(0.672)
Ln p.c. expenditure	0.3599***	0.3397***	0.0348***	0.3943***	0.3854***
	(0.056)	(0.068)	(0.009)	(0.057)	(0.067)
HH size using adult	-0.0055	-0.0067	0.0067**	-0.0091	-0.0104
equivalent scales	(0.024)	(0.024)	(0.003)	(0.022)	(0.023)
= 1for male head	-0.2220***	-0.2155***	-0.0071	-0.2151***	-0.2105***
	(0.060)	(0.060)	(0.011)	(0.058)	(0.059)
Education of head in	0.0025	0.0001	0.0033**	0.0031	0.0018
years	(0.008)	(0.008)	(0.001)	(0.008)	(0.008)
Age of cook	-0.0060	-0.0051	-0.0029	-0.0062	-0.0058
	(0.008)	(0.008)	(0.002)	(0.008)	(0.008)
Age of cook squared	0.0001	0.0001	0.0000	0.0001	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does farming	0.1996***	0.2066***	-0.0090	0.2220***	0.2279***
	(0.053)	(0.053)	(0.011)	(0.055)	(0.054)
Livelihood: public	-0.1599**	-0.1963***	0.0616***	-0.2059***	-0.2317***
sector employment	(0.070)	(0.075)	(0.019)	(0.074)	(0.084)
Livelihood: private	0.0202	-0.0063	0.0324**	-0.0482	-0.0680
sector employment	(0.066)	(0.074)	(0.013)	(0.065)	(0.078)
Livelihood: self-	-0.0862	-0.0972	0.0008	-0.1443**	-0.1550**
employment	(0.063)	(0.060)	(0.011)	(0.066)	(0.061)
Livelihood: casual	0.0864	0.0760	0.0067	0.0056	-0.0114
labour	(0.084)	(0.088)	(0.014)	(0.085)	(0.101)
Ln distance to SM			-0.0250***		
			(0.002)		
=1 for >median				0.2941***	0.3317***
KIOSK consumpt.				(0.082)	(0.100)
Interaction				-1.2260***	-1.5326**
i.KIOSK*SMshare				(0.387)	(0.684)
Constant	4.6185***	4.7705***	-0.1928**	4.1970***	4.2512***
	(0.493)	(0.579)	(0.096)	(0.535)	(0.588)
Observations	448	448	448	448	448
R^2	0.238	0.229	0.379	0.277	0.274

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Own calculation

One crude proxy of dietary quality used in the literature is dietary diversity, usually measured by the number of distinct food products or general food categories (e.g. cereals, roots and tubers, dairy) consumed (Ruel, 2002). What we find is that supermarket purchases do indeed increase the dietary diversity of households (see Table A 11 in Appendix). Increasing supermarket purchases by 10 percentage points increases the number of distinct major food categories by around 0.17 and 3.2 new food products are added to the diet. Given that all food groups and the majority of food items available in supermarkets are available in other retail outlets also, this is a notable result. However, there are several weaknesses to this measure that need to be considered when interpreting these results. Firstly, measures of dietary diversity usually use shorter recall periods. Also, even if we take positive relationship between dietary diversity and nutrient adequacy as a given, determining the threshold value of a high quality vs. a low quality diet is a sensitive and context specific matter that requires further research (Ibid).

Discussion

One weakness of our empirical setup regards the treatment of town dummies, which we do not include in our main specifications. Inclusion would be appealing in order to capture systematic town differences, such as general price or consumption differences. However, including town dummies in the IV specification renders our instrument to work poorly: because we only have three towns in our sample town dummies are highly correlated with the distance to supermarkets and cause distance to become insignificant in our first stage results. However, once livelihood sources are controlled for, towns remain significant only in few OLS cases and furthermore, coefficients of our supermarket purchases remain fairly robust (see robustness checks in Appendix). Furthermore, using expenditure shares rather than absolute expenditures as measure of consumption should reduce the impact on general price differences across towns.

Note again that the food consumption we are analysing here is limited to the food that is consumed or better available for consumption at home. Because most food that is bought in supermarkets and competing outlets is consumed at home, it is this consumption that we expect to be most severely affected by the local food environment. However, substitution effects with consumption outside home are possible but not explicitly addressed. For the robustness checks, we do additionally control for the food expenditure share away from home, which does not alter our main results. On average, the share of monthly total household expenditure spent on food away from home is around 10% in all towns, with median shares between 5-9% per town (the budget share dedicated to food consumed inside the home varies from 42-49% across towns).

Finally for the reason of our fairly small sample size, we cannot rule out that some effects remain undetected for reasons of statistical power.

6. Conclusion

This paper was motivated by continuing reports of the nutrition transition and negative health consequences in low income countries. Alongside other lifestyle changes, dietary changes have been linked in the literature to rising rates of nutrition-related non-communicable diseases and were argued to be demand as well as supply side driven.

Along this line, the rapid spread of supermarkets in low income countries is suspected to advance the nutrition transition by increasing the availability, affordability and by purposeful marketing associated foods and beverages to consumers. For this reason, we have analysed the effect of supermarkets on consumption patterns using very detailed household survey data collected for this purpose in a quasi-experimental setting in Kenya in 2012.

With respect to the affordability of food products, we established that lower (perceived) prices are by far the most important reason for consumer to shop at supermarkets. Our quantitative price data cannot clearly confirm nor reject this perception. The strongest incentive to shop at kiosks, the main traditional competitor to supermarkets, is physical access. In sum, drivers of retail outlet choices in small urban towns are similar to the ones that have been reported for large towns (Neven *et al.*, 2006), which suggests that our findings are relevant beyond the important group of small towns that were our focus.

In terms of consumption patterns, we find that supermarket purchases are positively associated with the consumption of processed at the cost of unprocessed foods. This holds in terms of expenditure shares as well as calorie shares and is mainly driven by an increased consumption of primary processed goods. While we had expected a stronger effect on consumption shares of highly processed foods (hypothesis $H1_1$), this does nevertheless suggest that the nutrition transition is advancing with spreading supermarkets.

As consumption patterns are changing towards more processed food, we find a positive effect of supermarket purchases on p.c. calorie availability, which confirms our hypothesis that frequent supermarket consumers consume more (hypothesis $H1_2$). As we do not find a significant effect on the food budget share (controlling for total expenditure) this increase in total calories is realised by a negative effect of supermarket purchases on the price paid per calorie. Particularly primary processed foods come at lower prices per calorie and support this effect.

More research is needed on the effect of these changes for the nutrient adequacy of consumers. We have shown that supermarket purchases are positively associated with dietary diversity, confirming our hypothesis ($H1_3$). However, it is out of the scope of this paper to investigate what this implies for nutrient adequacy that we are ultimately concerned with. Given that supermarket purchases have not been found to increase the calorie share of highly processed food at this point, the effects on nutritional status are likely at least, to be less adverse than expected and might even have beneficial effects for some parts of the population if and to the extent that they can contribute to a well-balanced diet.

Methodologically, our results confirm the adequacy of addressing the issue of endogeneity, which former studies have often neglected. For the reason of using a quasi-experimental design and IV methods, we are confident in interpreting the effect of supermarket purchases on consumption patterns as being causal.

7. Appendix

Table A 1: Product range of different retail formats in small towns

Typical products categories:	Supermarket	Small self-service store	Traditional kiosk
Non-food items of daily use	Yes	Yes	Yes
Crisps & salted snacks	Yes	Yes	Yes
Milk and yoghurt	Yes, fresh & long life	Yes, long life	**No**
Meat and fish	Yes, cooled sausages, frozen chicken & fish	**No**	**No**
Cooking fat, incl. cholesterol free	Yes	Yes	Yes
Fortified products (e.g. added vitamins)	Yes	Yes	**No**
Tinned products	Yes, but very limited	**No**	**No**
Instant noodles, breakfast cereals	Yes	Yes	Yes
Soft drinks, juices with sugar added, drinking chocolate	Yes	Yes	Yes
Fruit juice without added sugar	Yes	**No**	**No**
Alcoholic Beverages	Yes, but limited	**No**	**No**
Built-in over the counter retail (e.g. bakery, butchery, fast food stall)	**No** (only few cases)	**No**	**No**
Fresh fruits & vegetables	**No** (if yes, only very limited)		

Source: Own observation.

Table A 2: Summary statistics of main dependent and explanatory variables

DEPENDENT VARIABLES

	All	**Njabini (no SM)**		**Mwea (SM since 2011)**		**Ol Kalou (SM since 2002)**	
	Mean	Mean	Diff to others	Mean	Diff to others	Mean	Diff to others
Food expenditure shares							
Unprocessed food	0.63	0.65	0.03***	0.62	-0.02	0.62	-0.02
	(0.11)	(0.12)	(0.01)	(0.12)	(0.01)	(0.10)	(0.01)
Primary processed food	0.25	0.24	-0.00	0.25	0.01	0.24	-0.00
	(0.11)	(0.12)	(0.01)	(0.10)	(0.01)	(0.09)	(0.01)
Highly processed food	0.12	0.10	-0.03***	0.13	0.01	0.13	0.02**
	(0.10)	(0.10)	(0.01)	(0.11)	(0.01)	(0.08)	(0.01)
All processed food	0.36	0.34	-0.04***	0.38	0.02*	0.38	0.02*
	(0.11)	(0.12)	(0.01)	(0.12)	(0.01)	(0.10)	(0.01)
Calorie shares							
Unprocessed food	0.48	0.50	0.03**	0.47	-0.02	0.47	-0.01
	(0.12)	(0.13)	(0.01)	(0.12)	(0.01)	(0.11)	(0.01)
Primary processed food	0.42	0.42	0.00	0.43	0.01	0.42	-0.01
	(0.13)	(0.14)	(0.01)	(0.12)	(0.01)	(0.12)	(0.01)
Highly processed food	0.10	0.08	-0.03***	0.11	0.01	0.11	0.02*
	(0.09)	(0.09)	(0.01)	(0.10)	(0.01)	(0.08)	(0.01)
All processed food	0.52	0.50	-0.03**	0.53	0.02	0.52	0.01
	(0.12)	(0.13)	(0.01)	(0.12)	(0.01)	(0.11)	(0.01)
Calories p.c. per day (adult equivalent)	2561.01	2311.84	-388.94***	2608.23	67.38	2781.84	335.36***
	(1049.87)	(958.24)	(101.84)	(1095.87)	(108.41)	(1052.26)	(103.50)

	All	Njabini (no SM)		Mwea (SM since 2011)		Ol Kalou (SM since 2002)	
Calorie density (Kcal/per kg)	1338.22	1260.51	-121.30***	1354.72	23.54	1405.54	102.23***
	(255.71)	(248.63)	(24.54)	(259.33)	(26.39)	(239.14)	(25.04)
Price per calorie	0.04	0.04	-0.00	0.05	0.00	0.04	-0.00
	(0.02)	(0.02)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Food budget share (inside home)	0.46	0.49	0.06***	0.42	-0.05***	0.45	-0.01
	(0.15)	(0.15)	(0.01)	(0.15)	(0.02)	(0.13)	(0.01)
Food diversity							
# products consumed (less alcohol)	39.72	35.29	-6.92***	44.12	6.28***	40.53	1.23
	(12.69)	(12.55)	(1.21)	(12.53)	(1.28)	(11.48)	(1.26)
# food groups consumed	10.86	10.53	-0.52***	11.04	0.25*	11.05	0.29**
	(1.35)	(1.36)	(0.13)	(1.61)	(0.14)	(0.97)	(0.13)
Observations	448	161	161	134	134	153	153

Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

EXPLANATORY VARIABLES

	All	Njabini (no SM)		Mwea (SM since 2011)		Ol Kalou (SM since 2002)	
	Mean	Mean	Diff to others	Mean	Diff to others	Mean	Diff to others
SM expenditure share	0.10	0.02	-0.13***	0.11	0.03**	0.17	0.11***
	(0.12)	(0.06)	(0.01)	(0.10)	(0.01)	(0.13)	(0.01)
Ln p.c. expenditure	8.94	8.74	-0.31***	9.09	0.22***	9.02	0.12*
	(0.64)	(0.70)	(0.06)	(0.55)	(0.06)	(0.58)	(0.06)
Household demographics							
# female adults	1.09	1.17	0.14**	1.07	-0.02	1.01	-0.12*
	(0.64)	(0.77)	(0.06)	(0.60)	(0.07)	(0.52)	(0.06)
# male adults	0.91	1.04	0.21***	0.80	-0.16**	0.87	-0.06
	(0.70)	(0.85)	(0.07)	(0.57)	(0.07)	(0.59)	(0.07)
# adolescents	0.64	0.93	0.46***	0.49	-0.21**	0.46	-0.27***
	(0.96)	(1.17)	(0.09)	(0.81)	(0.10)	(0.73)	(0.09)
# children	1.00	1.13	0.21**	0.78	-0.31***	1.05	0.08
	(1.00)	(1.12)	(0.10)	(0.82)	(0.10)	(0.98)	(0.10)
Household size	3.63	4.28	1.01***	3.14	-0.70***	3.38	-0.38**
	(1.93)	(2.38)	(0.18)	(1.44)	(0.20)	(1.57)	(0.19)
Characteristics of head/ person responsible for food							
=1 if head is married	0.65	0.69	0.06	0.60	-0.06	0.65	-0.00
Education of head in years	10.13	9.01	-1.75***	10.75	0.88**	10.77	0.97***
	(3.80)	(4.04)	(0.37)	(3.44)	(0.39)	(3.59)	(0.38)
Age of cook	33.58	35.26	2.62**	33.21	-0.53	32.14	-2.19*
	(11.82)	(13.39)	(1.16)	(10.61)	(1.22)	(10.88)	(1.17)
=1 if HH does farming	0.59	0.68	0.15***	0.50	-0.13**	0.57	-0.03
	(0.49)	(0.47)	(0.05)	(0.50)	(0.05)	(0.50)	(0.05)
Main source of livelihood							
=1 for public sector	0.12	0.09	-0.05	0.11	-0.01	0.16	0.07**
=1 for private sector	0.28	0.20	-0.11**	0.23	-0.06	0.39	0.18***
=1 for self-employment	0.30	0.25	-0.07	0.38	0.12**	0.27	-0.04
=1 for agricultural	0.15	0.25	0.16***	0.15	0.01	0.03	-0.17***
=1 for casual labour	0.16	0.20	0.07**	0.13	-0.05	0.14	-0.03
Observations	448	161	161	134	134	153	153

Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Source: Own calculation.

Table A 3: Expenditure share highly processed food, main model and robustness checks

MAIN MODEL

	(1) OLS Expenditure share *highly processed* food	**(2) IV** Expenditure share *highly processed* food	**(3) 1st stage** SM expenditure share
SM expenditure share	0.0766*	0.0712	
	(0.041)	(0.091)	
Ln p.c. expenditure	0.0225***	0.0227**	0.0353***
	(0.008)	(0.010)	(0.009)
HH size	-0.0009	-0.0009	0.0043
	(0.003)	(0.003)	(0.003)
=1 if head is married	-0.0228**	-0.0228**	0.0010
	(0.009)	(0.009)	(0.011)
Education of head in years	0.0041***	0.0041***	0.0051***
	(0.001)	(0.001)	(0.001)
Age of cook	-0.0061***	-0.0061***	-0.0025
	(0.002)	(0.002)	(0.002)
Age of cook squared	0.0001***	0.0001***	0.0000
	(0.000)	(0.000)	(0.000)
=1 if HH does farming	-0.0346***	-0.0347***	-0.0135
	(0.008)	(0.008)	(0.010)
Ln distance to SM			-0.0252***
			(0.002)
Constant	0.0462	0.0445	-0.2056**
	(0.079)	(0.090)	(0.093)
Observations	448	448	448
R^2	0.256	0.256	0.351

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

SELECTED ROBUSTNESS CHECKS – only main variable of interest shown.

OLS	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	0.0766*	0.0736	0.0942**	0.0759*	0.0900*
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
R^2	0.256	0.256	0.259	0.256	0.201
IV	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	0.0712	IV invalid	0.0441	0.0690	IV invalid
	(0.09)		(0.09)	(0.09)	
R^2	0.256		0.256	0.256	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as in model (1).

Source: Own calculation.

Table A 4: Expenditure share primary processed food, main model and robustness check

	(1) OLS main Expenditure share *primary processed* food	(2) IV main Expenditure share *primary processed* food	(3) 1st stage SM expenditure share	(4) OLS Expenditure share *primary processed* food	(5) IV Expenditure share *primary processed* food
SM expenditure	0.1336***	0.2109**		0.2191***	0.2679***
share	(0.039)	(0.086)		(0.045)	(0.099)
Ln p.c. expenditure	-0.0829***	-0.0863***	0.0358***	-0.0757***	-0.0769***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.009)
HH size	0.0062	0.0062	0.0034	0.0058	0.0057
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
=1 if head is	-0.0089	-0.0089	0.0019	-0.0084	-0.0082
married	(0.012)	(0.011)	(0.009)	(0.012)	(0.012)
Education of head	-0.0009	-0.0014	0.0050***	-0.0008	-0.0010
in years	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Age of cook	0.0002	0.0003	-0.0018	0.0000	0.0001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook	0.0000	0.0000	0.0000	0.0000	0.0000
squared	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does	-0.0243**	-0.0224**	-0.0157*	-0.0182*	-0.0168*
farming	(0.009)	(0.009)	(0.008)	(0.009)	(0.010)
Mwea (SM 2011)	0.0247**	0.0241**	-0.0532***	0.0228**	0.0224**
	(0.010)	(0.009)	(0.018)	(0.010)	(0.010)
Ln distance to SM			-0.0305***		
			(0.003)		
=1 for >median				0.0494***	0.0548***
KIOSK consumpt.				(0.014)	(0.015)
Interaction				-0.1916**	-0.2361**
i.KIOSK*SMshare				(0.092)	(0.104)
Constant	0.9562***	0.9810***	-0.1976*	0.8638***	0.8695***
	(0.077)	(0.084)	(0.100)	(0.081)	(0.082)
Observations	448	448	448	448	448
R^2	0.316	0.310	0.384	0.344	0.343

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SELECTED ROBUSTNESS CHECKS – only main variable of interest shown.

OLS	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure	0.1336***	0.1095**	0.1628***	0.1363***	0.1014*
share	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
R^2	0.316	0.317	0.348	0.316	0.290
IV	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure	0.2109**	IV invalid	0.1854**	0.2255***	IV invalid
share	(0.09)		(0.08)	(0.08)	
R^2	0.310		0.348	0.308	

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as model (1).

Source: Own calculation.

Table A 5: Expenditure share all processed food, main model and robustness checks

	(1) OLS main Expenditure share *all processed* food	**(2) IV main** Expenditure share *all processed* food	**(3) 1st stage** SM expenditure share	**(4) OLS** Expenditure share *all processed* food	**(5) IV** Expenditure share *all processed* food
SM expenditure share	0.2134***	0.3781***		0.2969***	0.4296***
	(0.041)	(0.101)		(0.054)	(0.128)
Ln p.c. expenditure	-0.0595***	-0.0668***	0.0353***	-0.0502***	-0.0534***
	(0.010)	(0.011)	(0.009)	(0.010)	(0.010)
HH size	0.0045	0.0044	0.0043	0.0041	0.0039
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
=1 if head is married	-0.0313***	-0.0314***	0.0010	-0.0314***	-0.0308***
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Education of head in years	0.0032**	0.0021	0.0051***	0.0034**	0.0028*
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Age of cook	-0.0055***	-0.0053**	-0.0025	-0.0057***	-0.0056***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook squared	0.0001**	0.0001**	0.0000	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does farming	-0.0609***	-0.0569***	-0.0135	-0.0533***	-0.0496***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Ln distance to SM			-0.0252***		
			(0.002)		
=1 for >median KIOSK consumpt.				0.0564***	0.0711***
				(0.015)	(0.016)
Interaction i.KIOSK*SMshare				-0.1513*	-0.2725**
				(0.087)	(0.107)
Constant	0.9955***	1.0487***	-0.2056**	0.8768***	0.8929***
	(0.090)	(0.094)	(0.093)	(0.095)	(0.092)
Observations	448	448	448	448	448
R^2	0.233	0.208	0.351	0.270	0.258

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SELECTED ROBUSTNESS CHECKS – only main variable of interest shown.

OLS	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure share	0.2134***	0.1831***	0.2897***	0.2166***	0.1698***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
R^2	0.233	0.242	0.264	0.234	0.256
IV	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure share	0.3781***	IV invalid	0.3625***	0.3942***	IV invalid
	(0.10)		(0.09)	(0.10)	
R^2	0.208		0.259	0.205	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as model (1).

Source: Own calculation.

Table A 6: Expenditure share for unprocessed food, main model and robustness checks

	(1) OLS main Expenditure share *unprocessed* foods	**(2) IV main** Expenditure share *unprocessed* foods	**(3) 1st stage** SM expenditure share	**(4) OLS** Expenditure share *unprocessed* foods	**(5) IV** Expenditure share *unprocessed* foods
SM expenditure share	-0.2127***	-0.3220***		-0.2968***	-0.3635***
	(0.046)	(0.077)		(0.055)	(0.129)
Ln p.c. expenditure	0.0313**	0.0361***	0.0354***	0.0231*	0.0248*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
# female adults	0.0371***	0.0376***	-0.0011	0.0352***	0.0351***
	(0.010)	(0.010)	(0.005)	(0.010)	(0.010)
HH size	-0.0141***	-0.0141***	0.0046	-0.0133***	-0.0132***
	(0.005)	(0.005)	(0.003)	(0.005)	(0.004)
=1 if head is married	0.0412***	0.0413***	0.0010	0.0409***	0.0406***
	(0.012)	(0.012)	(0.008)	(0.012)	(0.011)
Education of head in years	-0.0016	-0.0009	0.0051***	-0.0018	-0.0015
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Age of cook	0.0055***	0.0054***	-0.0025	0.0056***	0.0056***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook squared	-0.0001**	-0.0001**	0.0000	-0.0001***	-0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does farming	0.0702***	0.0675***	-0.0135	0.0631***	0.0612***
	(0.010)	(0.010)	(0.009)	(0.010)	(0.011)
Ln distance to SM			-0.0252***		
			(0.002)		
=1 for >median KIOSK consumpt.				-0.0533***	-0.0607***
				(0.013)	(0.017)
Interaction i.KIOSK*SMshare				0.1698**	0.2307*
				(0.076)	(0.128)
Constant	0.2164**	0.1812*	-0.2058*	0.3226***	0.3146***
	(0.101)	(0.099)	(0.116)	(0.108)	(0.112)
Observations	448	448	448	448	448
R^2	0.240	0.229	0.351	0.270	0.267

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SELECTED ROBUSTNESS CHECKS – only main variable of interest shown.

OLS	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	-0.2127***	-0.2048***	-0.2864***	-0.2121***	-0.2968***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
R^2	0.240	0.244	0.234	0.240	0.270
IV	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	-0.3220***	IV invalid	-0.3083***	-0.3249***	IV invalid
	(0.08)		(0.08)	(0.10)	
R^2	0.229		0.234	0.229	

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as model (1).

Source: Own calculation.

Table A 7: Share of calories from different food categories – OLS and IV estimates

	(1) OLS Calorie share *highly processed* foods	**(2) IV** Calorie share *highly processed* foods	**(3) OLS** Calorie share *primary processed* foods	**(4) IV** Calorie share *primary processed* foods	**(5) OLS** Calorie share *all processed* food	**(6) IV** Calorie share *all processed* food	**(7) OLS** Calorie share *unprocessed* foods	**(8) IV** Calorie share *unprocessed* foods
SM expen-	0.0261	0.0381	0.0949*	0.1475	0.1209***	0.1857*	-0.1167***	-0.1787*
diture share	(0.035)	(0.079)	(0.048)	(0.116)	(0.042)	(0.111)	(0.042)	(0.108)
Ln p.c.	0.0286***	0.0281***	-0.0712***	-0.0735***	-0.0426***	-0.0454***	0.0387***	0.0414***
expenditure	(0.007)	(0.008)	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)
HHsize (ad.	-0.0018	-0.0018	0.0016	0.0016	-0.0002	-0.0002	-0.0002	-0.0002
equiv.)	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
=1 if head is	-0.0127	-0.0127	-0.0011	-0.0012	-0.0138	-0.0139	0.0147	0.0148
married	(0.009)	(0.009)	(0.015)	(0.014)	(0.013)	(0.012)	(0.013)	(0.013)
Education of	0.0036***	0.0035***	-0.0002	-0.0005	0.0034*	0.0030*	-0.0031*	-0.0027
head in years	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook	-0.0054***	-0.0054***	0.0017	0.0018	-0.0037	-0.0036	0.0037	0.0036
	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age of cook	0.0000***	0.0000***	-0.0000	-0.0000	0.0000	0.0000	-0.0000	-0.0000
squared	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH	-0.0357***	-0.0354***	-0.0380***	-0.0367***	-0.0737***	-0.0721***	0.0749***	0.0734***
does farming	(0.007)	(0.008)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Constant	-0.0405	-0.0366	1.0224***	1.0393***	0.9819***	1.0027***	0.0495	0.0296
	(0.067)	(0.078)	(0.110)	(0.110)	(0.111)	(0.117)	(0.107)	(0.113)
Observations	448	448	448	448	448	448	448	448
R^2	0.264	0.264	0.141	0.139	0.148	0.145	0.147	0.144

Robust (1)-(4) and cluster robust (5)-(8) standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Source: Own calculation.

Table A 8: Calorie share of processed foods, main results and robustness checks

	(1) OLS main Calorie share *all processed* food	**(2) IV main** Calorie share *all processed* food	**(3) 1st stage** SM expenditure share	**(4) OLS** Calorie share *all processed* food	**(5) IV** Calorie share *all processed* food
SM expenditure share	0.1209***	0.1857*		0.2249***	0.2445*
	(0.042)	(0.109)		(0.054)	(0.140)
Ln p.c. expenditure	-0.0426***	-0.0454***	0.0357***	-0.0345***	-0.0349***
	(0.012)	(0.012)	(0.009)	(0.012)	(0.012)
HHsize using adult	-0.0002	-0.0002	0.0053*	-0.0004	-0.0005
equivalent scales	(0.006)	(0.005)	(0.003)	(0.005)	(0.005)
=1 if head is married	-0.0138	-0.0139	0.0010	-0.0135	-0.0134
	(0.013)	(0.014)	(0.011)	(0.014)	(0.013)
Education of head in yrs	0.0034*	0.0030	0.0050***	0.0036*	0.0035
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Age of cook	-0.0037	-0.0036	-0.0026	-0.0040*	-0.0040*
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Age of cook squared	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if hh does farming	-0.0737***	-0.0721***	-0.0136	-0.0667***	-0.0662***
	(0.012)	(0.013)	(0.010)	(0.012)	(0.013)
Ln distance to SM			-0.0252***		
			(0.002)		
=1 for >median KIOSK				0.0571***	0.0593***
consumpt.				(0.017)	(0.020)
Interaction				-0.2474***	-0.2653*
i.KIOSK*SMshare				(0.092)	(0.139)
Constant	0.9819***	1.0027***	-0.2073**	0.8777***	0.8800***
	(0.111)	(0.104)	(0.092)	(0.112)	(0.112)
Observations	448	448	448	448	448
R^2	0.148	0.145	0.352	0.176	0.176

Cluster robust standard errors in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SELECTED ROBUSTNESS CHECKS – only main variable of interest shown.

OLS	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure	0.1209***	0.1124**	0.1695***	0.1310***	0.0627
share	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
R^2	0.148	0.151	0.193	0.153	0.159
IV	Standard controls	Including all towns	Additional controls	HH replace-ment control	SM location sample only
SM expenditure	0.1857*	IV not valid	0.1386	0.2177**	IV not valid
share	(0.11)		(0.10)	(0.11)	
R^2	0.145		0.181	0.148	

Cluster robust standard errors in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as model (1).

Source: Own calculation.

Table A 9: Log daily calorie availability per capita, robustness checks

OLS	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	0.3706**	0.1297	0.4574**	0.3359*	0.3887
	(0.19)	(0.19)	(0.19)	(0.18)	(0.28)
R^2	0.238	0.249	0.355	0.242	0.345
IV	Standard controls	Including all towns	Additional controls	HH replacement control	SM location sample only
SM expenditure share	0.8485*	IV invalid	0.6258	0.7552	IV invalid
	(0.50)		(0.43)	(0.51)	
R^2	0.229		0.354	0.235	

Robust standard errors in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls as in model (1), Table 6.

Source: Own calculation.

Table A 10: Food budget shares and prices per calories, OLS and IV estimation

	(1) OLS Food budget share	**(2) IV** Food budget share	**(3) OLS** Price per calorie	**(4) IV** Price per calorie	**(5) OLS** Price per calorie	**(6) IV** Price per calorie
SM expenditure share	-0.0244	-0.1494	-0.0109*	-0.0534***	-0.0167***	-0.0472***
	(0.046)	(0.106)	(0.006)	(0.012)	(0.006)	(0.011)
Ln p.c. expenditure	-0.1280***	-0.1220***	0.0138***	0.0157***	0.0123***	0.0133***
	(0.012)	(0.014)	(0.002)	(0.002)	(0.002)	(0.002)
HHsize using adult equivalent scales	-0.0074	-0.0074	-0.0006	-0.0005	-0.0006	-0.0005
	(0.005)	(0.005)	(0.001)	(0.001)	(0.000)	(0.000)
=1 for male HH head	-0.0183	-0.0193*	0.0038***	0.0033**	0.0038***	0.0035***
	(0.011)	(0.011)	(0.001)	(0.001)	(0.001)	(0.001)
Education of head in years	-0.0044**	-0.0036	0.0000	0.0003	-0.0000	0.0002
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Age of cook	-0.0034	-0.0036	-0.0001	-0.0002	-0.0001	-0.0001
	(0.003)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
Age of cook squared	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
=1 if HH does farming	0.0150	0.0118	-0.0045***	-0.0054***	-0.0053***	-0.0062***
	(0.011)	(0.011)	(0.001)	(0.001)	(0.001)	(0.001)
Exp share on food away from home	-0.3593***	-0.3680***				
	(0.061)	(0.065)				
=1 for >median KIOSK consumpt.					-0.0063***	-0.0078***
					(0.001)	(0.001)
Constant	1.8027***	1.7598***	-0.0722***	-0.0859***	-0.0549***	-0.0601***
	(0.117)	(0.132)	(0.016)	(0.017)	(0.015)	(0.017)
Observations	448	448	448	448	448	448
R^2	0.492	0.484	0.437	0.348	0.472	0.428

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Own calculation.

Table A 11: Food diversity indicators, OLS and IV estimation

	(1) OLS # food groups consumed by HH (excl. alcohol)	**(2) IV** # food groups consumed by HH (excl. alcohol)	**(3) OLS** # products consumed by HH (excl. alcohol)	**(4) IV** # products consumed by HH (excl. alcohol)
SM expenditure share	1.6550***	2.8555***	11.1922***	31.7750***
	(0.534)	(1.076)	(3.866)	(6.308)
Ln p.c. expenditure	0.2472*	0.1940	8.0892***	7.1769***
	(0.125)	(0.146)	(1.277)	(1.148)
HHsize (adult	0.1180	0.1167*	1.3234***	1.3022***
equivalent scales)	(0.072)	(0.071)	(0.414)	(0.399)
=1 for male	-0.6251***	-0.6105***	-5.4646***	-5.2139***
household head	(0.159)	(0.162)	(1.222)	(1.263)
Education of head in	0.0837**	0.0752**	0.7676***	0.6230***
years	(0.032)	(0.034)	(0.163)	(0.180)
Age of cook	-0.0587**	-0.0571**	-0.1706	-0.1419
	(0.027)	(0.028)	(0.202)	(0.207)
Age of cook squared	0.0005	0.0004	-0.0000	-0.0003
	(0.000)	(0.000)	(0.002)	(0.002)
=1 if HH does	0.3997***	0.4264***	4.5453***	5.0019***
farming	(0.136)	(0.136)	(1.244)	(1.142)
Constant	8.8784***	9.2635***	-38.5969***	-31.9951***
	(1.088)	(1.179)	(10.774)	(9.666)
Observations	448	448	448	448
R^2	0.172	0.163	0.327	0.297

Standard errors in parentheses.* p < 0.10, ** p < 0.05, *** p < 0.01.

Source: Own calculation

8. Literature

Asfaw, A. 2008. Does Supermarket Purchase Affect the Dietary Practices of Households? Some Empirical Evidence from Guatemala. *Development Policy Review* 26 (2):227-243.

Asfaw, A. 2011. Does Consumption of Processed Foods Explain Disparities in the Body Weight of Individuals? The Case of Guatemala. *Health Economics* 20:184–195.

Chandon, P., and Wansink, B. 2012. Does Food Marketing Need to Make us Fat? A Review and Solutions. *Nutrition Reviews* 70 (10):571-593.

Cummins, S., Petticrew, M., Higgins, C., Findlay, A., and Sparks, L. 2005. Large Scale Food Retailing as an Intervention for Diet and Health: Quasi-experimental Evaluation of a Natural Experiment. *Journal of Epidemiology & Community Health* 59 (12):1035-1040.

Giskes, K., van Lenthe, F., Avendano-Pabon, M., and Brug, J. 2011. A Systematic Review of Environmental Factors and Obesogenic Dietary Intakes Among Adults: are we Getting Closer to Understanding Obesogenic Environments? *obesity reviews* 12:e95-e106.

Hawkes, C. 2008. Dietary Implications of Supermarket Development: A Global Perspective. *Development Policy Review* 26 (6): 657-692.

Hawkes, C., Chopra, M., and Friel, S. 2009. Globalization, Trade and the Nutrition Transition. In *Globalization and Health: Pathways, Evidence and Policy*, edited by R. Labonte, Schrecker, T., Packer, C. and Runnels, V. New York: Routledge.

KNBS. 2010a. 2009 Kenya Population and Housing Census. Volume IA: Population Distribution by Administrative Units.: Kenya National Bureau of Statistics.

KNBS. 2010b. Kenya - 2009 Population Census Highlights.: Kenya National Bureau of Statistics.

KNBS, and ICFMacro. 2010. Kenya Demographic and Health Survey 2008-09. Calverton, Maryland: Kenya National Bureau of Statistics and ICF Macro.

Laraia, B., Siega-Riz, A., Kaufman, J., and Jones, S. 2004. Proximity of Supermarkets is Positively Associated with Diet Quality Index for Pregnancy. *Preventive Medicine* 39 (5):869-875.

Mergenthaler, M., Weinberger, K., and Qaim, M. 2009. The food system transformation in developing countries: A disaggregate demand analysis for fruits and vegetables in Vietnam. *Food Policy* 34 (2009):426–436.

Monteiro, C., Betrazzi Levy, R. , Moreira Claro, R., Rugani Ribeiro de Castro, I., and Cannon, G. 2010. Increasing consumption of ultra-processed foods and likely impact on human health: evidence from Brazil. *Public Health Nutrition* 14 (1):5-13.

Monteiro, C., Conde, W., and Popkin, B. 2002. Is Obesity Replacing or Adding to Undernutrition? Evidence from Different Social Classes in Brazil. *Public Health Nutrition* 5 (1A):105-112.

Monteiro, C., Moura, E., Conde, W., and Popkin, B. 2004. Socieoeconomic Status and Obesity in Adult Populations of Developing Countries. A Review. *Bulletin of the World Health Organization* 82:940-946.

Moore, L., Diez Roux, A., Nettleton, J., and Jacobs, D. 2008. Associations of the Local Food Environment with Diet Quality - a Comparison of Assessments based on Surveys and Geographic Information Systems. *American Journal of Epidemiology* 167 (8):917-924.

Morland, K., Diez Roux, A., and Wing, S. 2006. Supermarkets, Other Food Stores, and Obesity. The Atherosclerosis Risk in Communities Study. *American Journal of Preventive Medicine* 30 (4):333-339.

Neven, D., and Reardon, T. 2004. The Rise of Kenyan Supermarkets and the Evolution of their Horticulture Product Procurement System. *Development Policy Review* 22 (6):669-699.

Neven, D., Reardon, T., Chege, J. , and Wang, H. 2006. Supermarkets and Consumers in Africa: The case of Nairobi, Kenya. *International Food and Agribusiness Marketing* 18 (1/2):103-123.

Pearce, J., Hiscock, R., Blakely, T., and Witten, K. 2008. The Contextual Effects of Neighbourhood Access to Supermarkets and Convenience Stores on Individual Fruit and Vegetable Consumption. *Journal of Epidemiology & Community Health* 63 (2):198-201.

Pingali, P. 2007. Westernization of Asian Diets and the Transformation of Food Systems: Implications for Research and Policy. *Food Policy* 32:281-298.

PlanetRetail. 2013. Country Report Kenya.

Popkin, B. 1997. The Nutrition Transition and its Helath Implications in Lower-income Countries. *Public Health Nutrition* 1 (1):5-21.

Popkin, B. 2004. The Nutrition Transition: An Overview of World Patterns of Change. *Nutrition Reviews* 62:S140-S143.

Popkin, B., Adair, L., and Ng, S. 2012. Global Nutrition Transition and the Pandemic of Obesity in Developing Countries. *Nutrition Reviews* 70 (1):3–21.

Powell, L., Auld, M., Chaloupka, F., O'Malley, P., and Johnsson, L. 2007. Associations Between Access to Food Stores and Adolescent Body Mass Index.

Reardon, T., Barrett, C., Berdegué, J., and Swinnen, J. 2009. Agrifood Industry Transformation and Small Farmers in Developing Countries. *World Development* 37 (11):1717-1727.

Reardon, T., Timmer, C., and Berdegué, 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.

Roemling, C., and Qaim, M. 2013. Dual Burden Households and Intra-household nutritional inequality in Indonesia. *Economics and Human Biology* 11 (4):563-573.

Rosin, O. 2008. The Economic Causes of Obesity: A Survey. *Journal of Economic Surveys* 22 (4):617–647.

Ruel, M. 2002. Is Dietary Diversity an Indicator of Food Security or Dietary Quality? A Review of Measurement Issues and Research Needs. In *FCDN Discussion Paper No. 140*: IFPRI.

Schipmann, C. , and Qaim, M. 2011. Modern Food Retailers and Traditional Markets in Developing Countries: Comparing Quality, Prices, and Competition Strategies in Thailand. *Applied Economic Perspectives and Policy* 33 (3):345–362.

Schmidhuber, J., and Shetty, P. 2005. The Nutrition Transition to 2030: Why Developing Countries are Likely to Bear the Major Burden. In *97th Seminar of the European Association of Agricultural Economists*. University of Reading, England.

Sturm, R. 2002. The Effects of Obesity, Smoking, and Drinking on Medical Problems and Costs. *Health Affairs* 21 (2):245–253.

Subramanian, S., Perkins, J., Özaltin, E., and Smith, G. 2011. Weight of Nations: a Socioeconomic Analysis of Women in Low- to Middle-Income Countries. *The American Journal of Clinical Nutrition* 93:413-421.

Tessier, S., Traissac, P., Maire, B., Bricas, N., Eymard-Duvernay, S., El Ati, J., and Delpeuch, F. 2008. Regular Users of Supermarkets in Greater Tunis Have a Slightly Improved Diet Quality. *The Journal of Nutrition* 138 (8):768-774.

Wrigley, N., Warm, D., and Margetts, B. 2003. Deprivation, Diet, and Food-retail Access: Findings from the Leeds Food Deserts' Study. *Environment and Planning A* 35 (1).