2021

TWO ESSAYS ON FOOD ENVIRONMENT, NUTRITION, AND FOOD INSECURITY

Suliman Abdulaziz Almojel
University of Kentucky, suliman_almojel@hotmail.com
Author ORCID Identifier:
https://orcid.org/0000-0001-6079-9094
Digital Object Identifier: https://doi.org/10.13023/etd.2021.351

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Suliman Abdulaziz Almojel, Student

Dr. Yuqing Zheng, Major Professor

Dr. Tyler Mark, Director of Graduate Studies

TWO ESSAYS ON FOOD ENVIRONMENT, NUTRITION, AND FOOD INSECURITY

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture, Food and Environment at the University of Kentucky

By

Suliman Abdulaziz Almojel
Lexington, Kentucky
Director: Dr. Yuqing Zheng, Associate Professor of Agricultural Economics
Lexington, Kentucky
2021

https://orcid.org/0000-0001-6079-9094

ABSTRACT OF DISSERTATION

TWO ESSAYS ON FOOD ENVIRONMENT, NUTRITION, AND FOOD INSECURITY

A healthy food environment is fundamental to good health. It contributes to the reduction of obesity and the development of healthy eating habits. In spite of this, many people in the United States (US) have been hypnotized to become obese due to the current food environment. Recently, the US has consistently ranked high in the world in terms of obesity. The rising rate is symptomatic of consuming unhealthy diets. Besides, the double-edged crisis of the US food environment and obesity poses a major threat to food security and public health. Therefore, studying the US food environment is important to sustain quality food products and healthier food habits. In order to accomplish this, appropriate analysis and estimation techniques need to be employed to formulate the overall picture of the food environment. This dissertation has applied advanced analytical tools to shed light on key factors that affect household eating habits.

This dissertation has two major essays. In essay one (Chapter 2), a Structural Equation Modeling (SEM) technique is used to examine the linear causal relationships among latent and observed variables while simultaneously accounting for measurement errors that cause endogeneity problems. In this analysis, 3,861 US households were studied to investigate the healthy food environment from their community and consumer perspectives. Healthy eating and healthy food environment indexes are used among food-secure and food-insecure households to certify the quality of their food environments. To have a comprehensive understanding of the quality of food environment and its impacts on health and food security, one needs to investigate the two dimensions of the food environment, i.e., internal and external, simultaneously. A simultaneous analysis, rather than the sequential one, is necessary to better emphasize the importance of entailing the pillars of food security and the domains of the healthy food environment. This is essential in assessing the impact of food availability, affordability, utilization, and accessibility of nutritious food around US families.

In essay two (Chapter 3), a Three-Stages Least Squares (3SLS) approach is employed to determine the likelihood of improving dietary intake by examining how the density of healthy food resources impacts the purchases. The number of healthy food sources (i.e., retail stores, specialized food stores, direct marketing, and farmers' markets) is used to quantify the amount consumed of healthful food through the distance traveled and the price. The estimation is conducted on 4,126 households in rural and urban communities. The analysis is rooted in the hypothesis that the density of healthy food stores will affect healthy food consumption through price and travel distance reduction effects.

The findings of this dissertation are: (i) sorting households according to their food security status enables us to identify key factors that influence their eating habits; (ii) the cost of healthy food remains a burden and can be seen as a quality-related product differentiation. Price changes might not be enough to make people eat a healthier diet; (iii) quality food items, a variety of food choices, and food store types are essential elements of the external food environment to improve the eating habits; (iv) a significant lack of healthy food knowledge and awareness. The problem is increased as the food insecurity situation is increased; (v) participants in SNAP had a low healthy eating score compared to non-participants; (vi) healthy food sources improve consumers' perception of food availability and accessibility. Briefly, supermarkets and supercenters increase consumers' food purchases compared to small grocery stores; (vii) the increase in food quantity in urban areas can be attributed to stores specializing in healthy foods. The same can be said for rural areas with direct marketing and farmers' markets; (viii) the incentive to choose healthier foods is indirectly affected by lower travel distances and an extensive price decrease.

Findings from this dissertation shed light on community food environments and consumer food environments. Results provide insight into what affects food security inside and outside the homes of consumers. Further, the results highlight the impact of the availability of healthy food on both urban and rural areas. This study has important implications for improving health, sustaining nutritional diets, and enhancing regional welfare in the future.

KEYWORDS: Food Security, Food Environment, Healthy Eating Index, Food Stores, Food Access, Latent Variable Model.

Suliman Abdulaziz Almojel

August 9th, 2021

TWO ESSAYS ON FOOD ENVIRONMENT, NUTRITION, AND FOOD INSECURITY

By

Suliman Abdulaziz Almojel

Dr. Yuqing Zheng

Director of Dissertation

Dr. Tyler Mark

Director of Graduate Studies

August 9th, 2021

Date

DEDICATION

This dissertation is dedicated to my senior brother, Abdullah Abdulaziz Almojel, who has been a great source of motivation, inspiration, and support. Thank you for being there for me in the absence of our late father, who died when I was 12 years old.

ACKNOWLEDGMENTS

I want to acknowledge King Saud University in Saudi Arabia for supporting my research and granting me a scholarship to pursue my graduate studies. I want to sincerely thank Dr. Yuqing Zheng, dissertation advisor, teacher, and friend who encouraged me to pursue my interests and become an independent researcher. Yuqing, thank you for your insight and guidance. I will be eternally grateful to you for taking a chance on me and helping me succeed.

I also want to acknowledge my dissertation committee members, Dr. Sayed Saghaian, Dr. Tyler Mark, Dr. Yoko Kusunose, Dr. Patrick Mooney, and Dr. Haipeng Chen. All of the members provided valuable feedback that improved the quality of the analyses and improved the overall dissertation in many important ways. I am extremely grateful to Dr. Yoko Kusunose for investing her time and effort in teaching me the intricacies of research, providing undeterred mentorship, and helping me polish my dissertation essays.

A special thank you to my colleagues for providing me with advice and guidance throughout my graduate career: Yosef Alamri, Sattam Almudarra, Jawad Alhashim, Fuad Alagsam. Moreover, I would like to acknowledge the contributions of my colleagues, Dr. Didier Alia and Dr. Mohammed Rezgar.

This dissertation would not have been possible without the support from my mother, Norah Alhumaidan, and my sisters, Dr. Wafa, Monerah, Khlood, Lamia, and Sarah. I also want to acknowledge the support provided by my brothers, Fahad, Bader, Mohammed, Hamad, and Khalid.

Finally, I want to thank my friends who went through difficult times with me, who cheered me on and supported me during each step of the journey, Mohammad Alhogel, Nasser Alghumlas, Abdullah Alomran, Ali Almarri, Mohammad Aldossary, and Saad Abu-haimed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 General Introduction

1.1 Background

Food security has recently been a serious global issue (Grote, 2018). The concept of security refers to a condition in which all people's activity, ability, and health are free of all barriers or worries at all times (FAO, 2016). Food security is defined as the capability to meet sufficient food needs by having access to affordable, nutritious, and safe food sources (Pinstrup-Andersen, 2009). Even in developed countries with high income per capita, e.g., the United States (US), national reports indicate that nearly 1 in 11 individuals are unable to eat enough food all year around. In 2018, about 11% of families were food insecure in the US (Coleman-Jensen et al., 2019). This represented a decline from a high of 15% in 2011 and the same level reported in 2007 (Coleman-Jensen et al., 2012). Approximately 12.3% (41.2 million people) of US households were once reported as food-insecure in 2016 (Coleman-Jensen et al., 2017). As food security has become an increasingly significant issue in the US, it has also become more visible across the state boundaries (Figure 2.1). Even in areas where food security is not a problem, diets are not essentially healthy.

Given the current state of food security, obesity is becoming increasingly problematic. Globally, obesity and its related health consequences are less prevalent than in the US, and direct comparisons reveal important differences. The US has the highest number of obese individuals in the world, followed by Mexico, England, and Canada. The obesity trend is projected to continue to increase until 2030 (Devaux et al., 2017). About 71% of men and 62% of women are obese in the US, compared to 38% of men and 37%

of women worldwide (WHO, 2021). The high rate of obesity among Americans can partly be attributed to the high cost of obesity-related health care (Wang et al., 2011). Although poor diets are responsible for 14% of all deaths (Afshin et al., 2019), a healthy diet can save $50 billion in health care costs and decrease deaths by 35% (Lee et al., 2019).

Regarding the high obesity rate, economists and health professionals point at improving the food environment as a mechanism to promote healthy diets. Unfortunately, the current food environment for most Americans makes them prone to obesity (Hall, 2018). People who are unable to obtain quality, affordable food products find their situation even more complicated, leading to both food insecurity and obesity. Therefore, it is necessary to examine the food environment in the US to be able to give proper policy recommendations to the corresponding policy makers to help to sustain quality food products as well as to support healthier eating habits. The following section discusses some of the most noteworthy research in the field of the food environment.

1.2 Research Gaps

In relation to food insecurity, a study conducted by Oliveira et al. (2018) argues that the root of food insecurity in developing countries is insufficient food supply, while in the US, the roots are low income and an inadequate food environment. This argument is supported by Todd et al. (2010), who indicate that families living in low-income neighborhoods and lacking healthy food resources struggle to eat adequately. They conclude that people tend to make better food choices when they have a variety of options. The arrangements are similar to those reported by Rawlins et al. (2013) and Feldman and Wolnik (2019). In their analysis of food insecurity and its spatial distribution, Ziliak and Gundersen (2016) note that a household's nutritional quality varies significantly by neighborhood, and a lack of adequate food resources leads to unhealthy products.

With regard to obesity and food insecurity, several studies have shown that insufficient food resources have been associated with increased food insecurity and obesity (Maxwell and Smith, 1992; Seligman et al., 2009). Additionally, Nielsen et al. (2017) and Carlson and Keith-Jennings (2018) observed that food-insecure families are more likely to experience various nutrition and health problems due to their limited access to food.

When it comes to promoting child and family well-being, researchers have found that inaccessibility to healthy food, results in inadequate dietary intakes in children and adults (Ghosh-Dastidar et al., 2017; Mackenbach et al., 2017). For example, Kia-Keating et al. (2018) showed that children living in weaker food environments tend to have lower accomplishment rates at school. Approximately twice as often, children in food-insecure households report having fair or poor health, and food-insecure seniors generally report more difficulties with everyday activities than food-secure seniors (Gundersen and Ziliak, 2015)

By adopting a broader perspective, Draper and Younginer (2020) indicated that healthful choices are also influenced by healthy food knowledge. Collectively, we can be concluded from these studies that the food environment plays an important role in determining overall health and wellbeing.

However, most existing studies with a common issue discuss merely one aspect of the food environment. Their main focus was on food deserts in the external dimension of the food environment and on food access in the internal dimension. In contrast to previous research, we argue that internal and external factors need to be considered simultaneously to comprehensively understand how the food environment can affect people's eating habits. This study looks specifically at the effects of the community food environment and the personal food environment on food choices.

The contribution of this dissertation is four-fold. It will shed light on the impacts of the community food environment and the personal food environment on food choice. These results will enhance understanding of the food environment and its contributions to nutrition and food insecurity. This study will also use an advanced modeling technique to identify the effect of the internal and external food environment on eating habits, with a special vision on healthy foods. Therefore, we will be able to reveal key factors that affect eating habits in the US.

Moreover, this study will provide insights into the potential benefits of improving food environments for households. Additionally, this study will provide a better understanding of how healthy food sources differ among urban and rural areas, which is useful for policy making. Generally, findings could serve as evidence for recommending policy changes to maintain nutritional diets and enhance regional household welfare as well as individuals' health.

1.3 Research Objectives

This dissertation aims to enhance understanding of the quality of food environment and examine consumer behavior regarding food allocation. The specific objectives of the dissertation are as follows.

1. Examine the influence of community and personal food environments on food choices.
2. Establish a framework to help measure the relationship between internal and external food environments and household consumption of healthy foods.
3. Determine the effect of retail store density on household consumption in the US.
4. Examine the expected effects of increased availability of healthy food sources on individuals' incentives to make healthy food choices.
5. Identify the variations in healthy food sources between rural and urban areas.

1.4 Research Framework

This dissertation contains four chapters. Chapter 1 presents a general introduction followed by a brief overview of the research topic. It states the research goal, objectives, and the analytical techniques used to achieve the objectives. Chapter 2 uses a structural equation modeling to investigate the causal effects, having endogeneity problems into account. Chapter 3 employs a Three-Stages Least Squares (3SLS) model to determine how store density affects household consumption. Finally,Chapter 4 summarizes the key findings, policy implications, limitations, and directions for future studies.

Chapter 2 How Much Does the Food Environment Matter?

Abstract

The health consequences of obesity in the US are becoming increasingly evident, making it essential to follow a healthy eating habits. This study examines the causal relationship between household healthy eating behaviors and community aspects of the food environment. Food Acquisition and Purchase Survey and Nutrition Environment Measurement data are used to construct a model of healthy consumption based on United States Department of Agriculture (USDA) guidelines and Harvard's healthy eating pyramid. An empirical framework developed in this study links the pillars of food security to the domains of a healthy food environment to derive a comprehensive understanding of the availability, affordability, accessibility, and quality of nutritious food around American families. The structural Equation Modeling (SEM) technique was used to determine the quality of the food environment and the probability of adopting healthy eating habits. The model was applied to the total sample of 3,861 households and extended to compare food insecure and food-secure families. Findings imply that the cost of healthy food and its physical presence is a burden among households. Practically, food-insecure families face the most significant impacts of high prices for healthy foods, and the issue also is packed with unexpected household expenses. Participants in SNAP had a low healthy eating score in comparison with non-participants. Findings indicate a lack of healthy food knowledge and awareness. Identifying healthy items and the time it takes to make healthy food is observed as a barrier by families with children. For individuals to improve their eating habits, quality food items, various food choices, and short distances are essential elements of the external food environment. The study identified individuals' food acquisition through internal and external food environments and its associations to health, nutrition, and food security.

KEYWORDS: Healthful Food Choice, Food Environment, Structural Equation Modeling, Healthy Eating Index, SEM

2.1 Introduction

A Healthy Food Environment plays a key role in promoting healthy eating habits. *The Food Environment* has been defined as "the physical presence of nutritious and safe food in markets that are convenient and affordable for families" (Turner et al., 2017a). It provides a basis for sufficient and optimal nutritional status. By contrast, *Malnutrition* can be defined as any stimulus that deviates an individual from achieving adequate nutrition (Hoefer and Curry, 2012). Malnutrition currently poses a serious threat to public health worldwide, coincident with obesity[1].

As mentioned in the previous chapter, the United States (US) is the world's top country struggling with obesity (Devaux et al., 2017) where its rising rate is symptomatic of consuming unhealthy diets. Accordingly, the increasing rate of obesity is more rapid among food insecure individuals. The current food environment in the US is hypnotized to make most individuals prone to obesity (Hall, 2018).

According to Committee on World Food Security (CFS), food security consists of four pillars which make food accessible available, utilized, and stable (Shaw, 2007). These four pillars include (i) accessibility, (ii) availability, (iii) utilization, and (iv) affordability. It is necessary for a society that is trying to ensure the welfare of its people to meet all four pillars of food security.

[1] A specific definition of Malnutrition has shown elusive. A possible explanation for this might be that Malnutrition varies among countries and the term has varied over time. Malnutrition is often viewed as undernutrition in developing nations, whereas it is viewed as overnutrition in developed nations.

It is worthful to declare that if no improvements are made to the food environment, these double-edged crises of the US food environment and obesity will pose a major threat to food security. Figure 2.2 illustrates the impact of the food environment across states.

As shown in Figure 2.3, the food environment is generally classified into two dimensions: external food environment, also known as community food environment, and internal food environment, also called consumer food environment. The dimension “external food environment” and “community food environment” are used mutually to indicate the supply and availability of food to families. It encompasses the combined domains of food vendor density, the spatial location of food stores and markets, retail food types, restaurant types, the monetary value of food, and attributes within products and stores (Caspi et al., 2012; Scott, 2017).

The dimension “internal food environment” and “personal food environment” are used interchangeably to mean the accessibility to allocate healthy food and affordability of a broad selection of healthy foods to consumers. It also reflects the individual perspectives about healthy food knowledge, taste, and desirability (Morland et al., 2002; Glanz et al., 2007a; Pitt et al., 2017).

Understanding the link between the internal domains, e.g., purchasing power, convenience, and desirability, and the external domains, e.g., prices/food availability, geographic and product/store properties, will provide insight into why consumers allocate their food in the manner they do. This complex linkage is at the heart of our understanding of why the current food environment in the US is hypnotized to make most individuals prone to obesity.

The remainder of the chapter is structured as follows: Section 2.2 discussed more broadly the related literature and the gaps the present study contributes to filling. The objectives and research questions of this study have been discussed in section 2.3. In section 2.5, the conceptual framework underlying the analysis is formulated. Section 2.6 presents the empirical models and analytical approach. Next, the data and descriptive results are discussed in section 2.6. The results are presented in section 2.7. Section 2.8 offers discussion, concluding remarks and identifies opportunities for future research.

2.2 Related Literature

This study sits at the intersection of two related strands of the literature. First, it contributes to the literature on food security in developed countries, focusing on dietary preferences and food consumption habits. Second, it adds to the literature on the food environment and its relationship to healthful food choices in the US. It sheds light on the influences of the external food environment and the internal food environment on food choice. This section provides details on previous studies conducted on food assistance programs and external and internal food environment domains.

2.2.1 Food Assistance Programs

The Supplemental Nutrition Assistance Program (SNAP) is administered by the US Department of Agriculture (USDA) under the Food and Nutrition Service (FNS). The program provides a monthly supplement for purchasing nutritious food through an Electronic Benefits Transfer (EBT) card similar to a bank card.

Previous research on the impact of food-aid programs is not recent, having possibly first been studied by Southworth (1945). Since it has been in operation for more than five decades, SNAP has proven effective in addressing health and food insecurity (Gundersen and Ziliak, 2015). In addition, Schanzenbach et al. (2016) indicated that households facing food insecurity are more likely to enroll in food programs. Recently, Hidrobo et al. (2014) and Todd et al. (2010) confirmed that most welfare programs have been succeeding and have played an essential role in improving the internal food environment domains among millions of Americans.

As far as SNAP is concerned, recent efforts have been devoted to how SNAP has improved the wellbeing of millions of Americans. Using data on 7,000 participating households, Beatty and Tuttle (2015) found that SNAP benefits cause families to increase food expenditure. According to Wilde (2013), the average food spent by SNAP beneficiaries is between 10% and 16%, and SNAP pays for nearly half of all food expenses for families in low-income households. The studies clearly demonstrate that SNAP has a positive impact on the internal food environment.

Regarding health, several studies stated that SNAP has played an essential role among millions of eligible, low-income families in improving health outcomes (Nord and Golla, 2009; Gundersen and Ziliak, 2015; Deb and Gregory, 2016; Bergmans et al., 2018). However, there is uncertainty about the effectiveness in promoting healthy intake. For instance, Gregory et al. (2013) measured the quality of nutritional intake among SNAP participants and non-participants. They demonstrated that nutrition quality is 2.5% lower among SNAP participants than non-participants. In the same vein, Gundersen and Oliveira (2001) and Wilde (2018) opine that food assistance benefits are still not sufficient to end their food insecurity problems for some families. Gundersen and Ziliak (2015) added that some recipients still suffer from food insecurity, which requires some adjustments.

According to Lusk and Weaver (2017), a health restriction may limit the consumption of unhealthy foods. Still, health restrictions should be implemented with caution because they may negatively impact SNAP participation. In essence, this study expects to see an increase in nutritional intake quality among SNAP participants. We explore in-home and out-of-home factors associated with food insecurity. Some of the determinants are discussed in the following subsections.

2.2.2 Studies on Internal Food Environment

Focusing on food-secure and insecure families, several studies have recently shed more light on food affordability, food utilization, and food knowledge. In their analysis on non-food and food spending, Nielsen et al. (2017) found that food-insecure homes are more likely to encounter other problems than food-related issues, for instance, housing-related financial problems, such as large unexpected bills. Carlson and Keith-Jennings (2018) have explained that healthy foods might be costly among food-insecure families. As expressed, food affordability is impacted by the family structure, and monthly expenses indirectly affect food consumption. We anticipate that large, unexpected spending impacts households so that will limit their ability to purchase healthy foods.

Regarding Nutrition and healthy behavior education, the Supplemental Nutrition Assistance Program–Education (SNAP-Ed) remains an optional program across states. Rivera et al. (2016) evaluated the effects of the SNAP-Ed policy intervention on 575 households through a randomized, controlled, parallel study design. Based on their findings, the SNAP-Ed program significantly improved the food security and health of low-income families with children. SNAP-Ed educational activities are encouraged to be included among the benefits of state governments (Naja-Riese et al., 2019). This study aims to assess the level of knowledge about healthy living among food secure and food-insecure families.

Concerning how individuals use the food they obtain, Aryal et al. (2019) and Draper and Younginer (2020) found that interacting with food intake is influenced by many factors such as desirability, the time involved in preparing healthy food, and the taste of healthy food. In addition, Charlton (2016) and Nord (2008) stated that families could convert a poor diet into a rich diet if they improve their healthy lifestyles and skills. In the same vein, Backett-Milburn et al. (2010) observed a significant correlation between parent and child behavior, including food intake, eating motivations, and satisfaction. They concluded that children raised in environments with unhealthy eating habits are likely to suffer from obesity and eating disorders.

To conclude, the literature shows that having a better internal food environment and a better diet go hand in hand. Therefore, this study expects that the internal food environment domains of healthy food knowledge, healthy food affordability, healthy food preference, and healthy food tastes differ between food secure and insecure individuals. Likewise, related domains regarding the external food environment are discussed in the following.

2.2.3 Studies on External Food Environment

Many studies have looked at the fragility of the external food environment and how it affects consumer food choices, including Blanchard and Matthews (2007); Walker et al. (2010a); Bitler and Haider (2011); Hamidi (2020). These studies point to widespread areas with limited access to healthy food. Low-income communities generally suffer from limited access to food compared to High-income and affluent neighborhoods

Numerous studies have examined the influence of external environmental factors on food choices. A lack of healthy and nutritious food resources is expected to lead to unhealthy eating habits (Dubowitz et al., 2015; Weatherspoon et al., 2015). Along with insufficient food resources, Maguire et al. (2015) studied that families often have poor diet quality. As shown in this case, Gundersen and Oliveira (2001) highlight that families may also trade-off healthy and nutritious foods for unhealthy ones because of inadequate food resources. Considering the availability of healthy foods, consumers behave differently. It will be more challenging to buy healthier foods if food resources are insufficient, such as in food deserts or food swamps.

The term *Food Deserts* was first used by Cummins and Macintyre (2002) to describe geographical locations where access to affordable and nutritious foods is limited. Conversely, The term *Food Swamps* as used by Babey et al. (2008), Kelly et al. (2011), and Cooksey-Stowers et al. (2017), refers to a geographic area with a high density of fast food/junk food and corner stores. Regarding food deserts in the US, Rhone et al. (2019) and Feldman and Wolnik (2019) argued that states in southern regions tend to have the most deficient external food environment. This is partly because of the highest levels of food deserts and low-income populations. Accordingly, people living in food desert areas exhibit a 24% higher prevalence of food insecurity than those living in areas with enough food available (Laska et al., 2015; Pitt et al., 2017).

On the other hand, food swamps are associated with higher body weight and less healthy dietary intake (Cooksey-Stowers et al., 2017). Some studies have looked at other measurements that accommodate fast food/convenience stores to healthy food outlets. According to Hendrickson et al. (2006) and Drewnowski and Specter (2004), most low-income urban areas have a higher influx of fast-food restaurants coupled with the number of corner stores that offer ready-to-eat foods than other high-income areas. These areas are often devoid of any retail stores resulting in limited access to nutritious foods (Walker et al., 2010b). Due to the limited availability of healthy foods in these neighborhoods, they can be best described as having inadequate food resources.

Several recent studies have indicated that food swamps predict neighborhood food environments better than food deserts in the US (Woodham, 2011; Luan et al., 2015; Cooksey-Stowers et al., 2017; Rummo et al., 2017). Rose et al. (2009) indicated that food swamps increase the risk of obesity around people compared to food deserts factors. Regarding a healthy food environment, we expect that the magnitude of food swamps provides a stronger predictor than food deserts. This study addresses a similar question focusing on understanding the interrelationship between the quality of the current food environment and healthful food choices.

The studies mentioned above substantiate a lack of comprehensive comparison regarding the performance of the food environment. Their focus was, generally, on one particular pillar of food security, e.g., the availability or accessibility of healthy food. Nevertheless, we looked simultaneously at the food environment's external and internal characteristics and impacted the US household's healthful food choices. It is essential to ensure the importance of the four pillars of food security, e.g., accessibility, availability, utilization, and stability, are met, alongside and simultaneously. This study aims to fill this gap in the literature to improve food security levels among American families.

2.3 Research Questions and Objectives

In addition to better external food settings, better internal food environments also lead to more nutritious eating habits. Therefore, two research questions are addressed in this study. (1) *What are the main dimensions and domains of the food environment that influence food choices and acquisition?* and (2) *Which of the domains under each dimension matter most in explaining households' healthy eating habits?*

This study aims to develop a complete understanding of food environments among US families and determine its impact on health and food security. This can only be accomplished by looking at the two dimensions of food environment simultaneously rather than sequentially. In this case, one needs to consider the pillars of food security with the domains of the healthy food environment to understand the impact of access to food, affordability, utility, and accessibility of nutritious food.

The contribution of this study is three-fold. First, the study contributes to the growing body of literature regarding food security and dietary preferences. It merges the two strands of literature about the food environment and its relation to healthful eating choices. It looks comprehensively at various in-home healthy consumption behavior and out-of-home contributing factors regarding the four food security pillars. Second, the study uses an advanced modeling technique to certify the effect of the internal and external food environment when it comes to eating healthy foods. Third, this study provides insights into the potential benefits of improving food environments in individuals.

2.4 Data Source

This study uses data from USDA's National Household Food Acquisition and Purchase Survey (FoodAPS), conducted by the USDA's Economic Research Service (ERS) and Food and Nutrition Service (FNS). The sample was selected at random using a multi-stage stratified method. The sample is nationally representative and sufficiently large to make inferences at the national level.

The FoodAPS provides information on neighborhood food environment through food store choice, geographical location, and proximity to various food venues. The data contains information on households' food acquisition and well-being through food security, health, and obesity. The individuals are asked to complete three interviews over nine days. From April 2012 to January 2013, the survey collected data on household purchases over seven days. The data recorded food-at-home and food-away-from-home consumptions by tracking store purchase and dining-out receipts and reporting the information in a food book. The original sample size was 4,826 households. The final sample size, after cleaning and eliminating observations with missed or inaccurate entries, is 3,861 households. The missing observations mainly relate to variables with health food knowledge, convenience, and desires, as well as store attributes and product attributes. There are several observations missing that cannot be predicted using other information. Therefore, the missing observation cause problems during the fitting process, such as convergence problems (Carter, 2006).

Also, the USDA's Food and Nutrient Database for Dietary Studies (FNDDS) and USDA's National Nutrient Database for Standard Reference Legacy Release (SR) are adopted to outline more detailed information regarding the description of the healthful items (Details regarding the foods used in the study are provided in Appendix 4.A).

Table 2.2 presents definitions and descriptive statistics of the variables used in the analysis. We tested whether food environment's factor structure is the same among food-secure individuals versus food-insecure households by comparing the means within groups. The average cost per unit is 0.19 cents for healthy food. The mean difference in price shows that there is no significant difference between food secure and insecure households. The average environment score index is 7.5 out of 10 for a healthy food environment, from 0 (worst) to 10 (best). The households are surrounded by two fast-food restaurants, full-service restaurants, or convenience stores as all the other sources such as supermarkets or supercenters. The density food venues categorized the number of stores within 1 mile for urban, 10 miles for rural areas[2]. We used variables that represent the reasons to choose a primary food store. From the sample, nearly 87% of the stores authorized to accept SNAP EBT, 60% of individuals shopped at their primary store due to low prices, 33% preferred stores with excellent quality items, 50% favored the store's closeness, 28% of people may have other reasons regarding the store characteristics such as a good variety of foods, and acceptance of SNAP EBT.

Geographical regions in which the household resides, including Northeast, Midwest, South, and West, play an essential role in food acquisition. In our sample, on average, 28% of households are located in rural areas, while 72% are in urban areas. Also, 17%, 25%, 23%, and 36% are from the Northeast, Midwest, West, and South, respectively.

[2] As (Cooksey-Stowers et al., 2017; Willis et al., 2020) suggested.

We also include SNAP participation and spending power along with the family size and number of kids. The reason for combining those variables into one variable is to account for the multicollinearity. In addition, the combination used to capture overall food affordability is based on previous studies such as (Carson, 2002). The spending power accounts for monthly income, taking into consideration the number of adults and children. For instance, children might require a small portion than the adults in the house. Equivalence scales usually consider a reference household size. The guide foundation is given an element of one. The element is increased by 40% for each adult and 30% for every child. For example, a family consists of a husband and wife with two children would have a spending power equivalent to $1,625 (when their income is $3,900). This formula is adapted from (Carson, 2002; Palameta and Macredie, 2005), and it is calculated as follows.

$$\text{Spending power } = \frac{\text{Household income}}{1 + 0.4(\text{adults}) + 0.3(\text{children})} \qquad 2.1$$

The variables that indicate the means of transportation show that 86% of families have access to a car, and 12% of individuals use public means of transportation to primary food stores. The impact of preference, desire, health education, knowledge, and experience for adopting healthy food eating is worth noticing. However, 65% of the total sample check the food labels, calories, or nutritional information, only 2% of consumers can identify healthy food options[3]. In addition, 44% of households indicated that healthy food is costly, 19% were too busy to take some time to prepare healthy foods, and 12% noted that healthy food tastes bad.

[3] Anastasiou et al. (2019) and Neff et al. (2019) have expatiated on this subject.

2.5 Conceptual Framework

Throughout this study, the internal and external food environments are deconstructed into sub-models to demonstrate how the internal domains, i.e., affordability, accessibility to stores, knowledge about healthy foods, desirability, and external domains, i.e., food availability, geographical location store, and product characteristics, are treated differently depending on these dimensions. This project provided an important opportunity to advance the understanding of the food environment in the US. The steps we applied to enhance our understanding of the US food environment is demonstrated below.

The framework used in this study consists of the following steps. First, to understand the inner functioning of the dimensions and their related domains followed logically from the literature. Second, to deconstruct the model by linking each domain with its dimension, as shown in Figure 2.3. The second step is important in allowing the model to be deconstructed theoretically. Hence, in order to have a reliable and valid theory, we adapted the inner functioning of the food environment from Turner et al. (2017b); (2018) and the inner functioning of the pillars of food security from the Committee on World Food Security (CFS) (Shaw, 2007). Having used the mechanism depicted in Figure 2.4, we developed a theoretical map of the external and internal food environment and their domains.

The third step consists of choosing the variables that are needed to construct each domain. This step involves a lengthy process designed to select the variables that best reflect the food environment domains. The procedure is based upon the theories of previous studies and their variable selection. Table 2.1 presents the variables embedded in the model that measures each domain of the food environment.

2.5.1 Variable Selection

Table 2.1 and Table 2.3 present the variables used to shape the external and internal food environment. It is essential to select the model and variables cautiously based on the research purpose, the analytical framework, and the theoretical support (Cover and Thomas, 2012). In this study, a long process was involved in matching the food environment construction based on the theories from previous studies (USDA, 2015; Ashworth, 2017; Turner et al., 2017a; Turner et al., 2018).

To construct the external and internal dimensions, domains were obtained from the literature as explained in the following paragraphs. Hsu et al. (2016); Ziliak and Gundersen (2016) suggested that food availability consists of a vector of variables including food outlet's density and restaurant numbers. North et al. (1999); Pitt et al. (2017) used store's features including food quality and variety of food selections, low price stores, and short distance, in the study. Food accessibility consists of variables that characterize the physical distance of food stores and fast-food restaurants in miles. Other variables include transportation means used to allocate foods (Breyer and Voss-Andreae, 2013; Pitt et al., 2017) and food Affordability (Carson, 2002; Palameta and Macredie, 2005); Convenience and Desirability consist of variables related to reasons for not consuming healthy food products, such as "healthy food taste bad," "healthy food requires extra time to cook and prepare healthy dishes," and "healthy foods cost too much." Knowledge, Skills, and

Education also were suggested by Guthrie et al. (1999); Scaglioni et al. (2008); Neff et al. (2019).

We use these data to illustrate the models as in Figure 2.5. For the external food environment, the price of food is a determining factor in the food choices. Making a change in price may not be sufficient enough to make consumers choose healthy foods (Caputo et al., 2018). Therefore, we use price as a continuous variable towards healthy foods and included a dummy variable as an indicator in the affordability to see "If healthy foods cost too much" as a technique for pricing information to study consumer perceptions toward healthy foods (Kardes et al., 2004). According to Saghaian and Reed (2004), assuming individuals are distributed uniformly, we can use the desirability domain to guide our analysis on the price. Since the mean price does not differ statistically between food-secure and food-insecure individuals, the price can be viewed as a quality-related product differentiation.

2.5.2 Harvard's Healthy Eating Pyramid

The quality of food is based on the Healthy Eating Index (HEI). The HEI is utilized by Mancino et al. (2018b), based on Harvard's Healthy Eating Pyramid (HHEP). The HHEP was used to identify food items[4] that constitute healthful foods based on Willett and Skerrett (2017). The pyramid draws from an extensive collection of epidemiological research and various approved dietary guidelines and nutritional facts to construct different food items' nutrition quality scores from the healthiest to the unhealthiest at the top (Source, 2008; Datz, 2011; Caivano and Domene, 2020).

[4] List of food items for food-at-home acquisitions is listed in Appendix 4.A.

Figure 2.8 consists of components which include cereal fiber, fruits, vegetables, nuts/soy, white to red meat, polyunsaturated to saturated fatty acids, trans fat, multivitamins, and alcohol (Willett and Skerrett, 2017; Caivano and Domene, 2020; Del Castillo et al., 2020). The HHEP has been a valuable tool for overall health professionals due to the flexible framework that allows researchers to aggregate food items to determine their healthfulness measured through the HEI[5] (Mancino et al., 2018a). In this study, the HEI score measures the dietary quality rather than quantity. The HEI evaluates healthy food acquisition concerning the internal and external food environment.

2.5.3 Retail Food Environment Index

The Retail Food Environment Index (RFEI) was used to capture healthy eating resources within grocery/convenience stores and reflect the existing food environment's quality to a given household. They can be used to measure either the nutrition quality within stores (Glanz et al., 2007a), fast-food restaurants, or the quality of a household diet (Saelens et al., 2007b; Babey et al., 2011). Adopted from CCPHA (2007) and Cooksey-Stowers et al. (2017), the RFEI to a given household is expressed as:

$$\text{RFEI} = \frac{\text{Fastfood restaurants} + \text{Convenience stores}}{\text{Total grocery stores}} \tag{2.2}$$

The setting for food shops versus fast-food restaurants and convenience stores is within 1 mile for urban and 10 miles for rural, as Cooksey-Stowers et al. (2017) and Willis et al. (2020) suggested. The RFEI value is obtained by counting the sum of the total number of fast-food restaurants and convenience stores and divided by the total number of grocery stores. The result implies how many food store outlets are around a household location.

[5] see Appendix 4.B. for more information on the HEI components we follow in this study.

For example, if a person has an RFEI of 3.0, fast-food restaurants and convenience stores are close by three times as much as large grocery stores.

In this study, the RFEI is used to discover whether households live around food swamp areas or not. In addition to the RFEI, the Nutrition Environment Measurement Survey (NEMS) tool was used to measures the nutrition environment within stores or fast-food restaurants. Several studies have used the NEMS to analyze the healthy eating environment of different neighborhoods (Krukowski et al., 2010; Hillier et al., 2012). These studies utilized this tool for community and consumer nutrition environments, the density of food stores, and fast-food restaurants. The NEMS tool was ideal for this study because it measures the food environment based on two techniques, including (1) the nutrition environment, together with the number, type, location, and food stores and (2) the consumer nutrition environment, like the availability, cost, and quality of food.[6]

2.6 Empirical Models

In this section, we will describe the reason for using a latent variable approach to build the external and internal dimensions. We also provide the method used to select the proper variables to represent each dimension. In this study, the terms '*observed variable*', '*indicator variable*', and '*measured variable*' are used interchangeably to mean data that exists and has been measured and recorded. The terms '*unobserved variable*' and '*latent variable*' are used interchangeably. They are not directly observed; rather, they are being measured according to a mathematical model (Grace, 2006; Grace et al., 2010).

[6] See Glanz et al. (2007b), Krukowski et al. (2010), and Saelens et al. (2007a) for more details on the NEMS method.

2.6.1 Latent Variable Approach

The domains in the external and internal food environment were not available in the dataset. To overcome this issue, available indicator variables in the FoodAPS were employed to reflect the external and internal food environment by using the latent variable technique (Hair et al., 1998). The latent variables can indirectly measure the domains in the external and internal food environment by using multiple indicator variables. In the area of economics, variables such as food security, hunger, poverty, or quality of life are not directly measured; they are measured through indicator variables. Another example regarding food security, a variety of indicators shapes the concept of food security, such as "a household has enough food to consume" and "short of food supply" to predict the food security according to FAO (2006). With the latent variable, analysts can expand this technique and include "market prices, income, nutritional risk factors, and other in-house related factors" to make food security indirectly observable (Barrett, 2002). In this study, we follow Pearl (2012) to build the latent variables (i.e., internal and external food environments) from the observed variables in FoodAPS.

The latent variables need multiple indicators to make the mathematical model identified. Table 2.1 identifies the latent variables with their indicators. As discussed above, the selection of all of the indicators was motivated based on the available literature (Deaton, 1989; Guthrie et al., 1999; North et al., 1999; Carson, 2002; Palameta and Macredie, 2005; Botonaki et al., 2008; Gundersen et al., 2011; Breyer and Voss-Andreae, 2013; Hsu et al., 2016; Ziliak and Gundersen, 2016; Gundersen et al., 2017; Pitt et al., 2017; Neff et al., 2019). Table 2.3 demonstrates the internal and external food environment variables. The latent variables "Product & Store Properties" and "Availability of Food" are measured indirectly by the observed variables (Least Expensive, Closest to home, Good Quality products, and other properties), and (Fast food restaurants density, Grocery store density, Food desert, and Food swamps), respectively.

A similar approach is used to construct the internal food environment factors. The internal food environment includes “Food Accessibility” measured by (physical distance, time, mode of transportation, store types), “Food Affordability” indicated by (In-house large expenses, SNAP, spending power, Out-of-pocket costs), “Convenience and Desirability“ indicated by (Taste of healthy food, time preparing healthy food, Individual reasons). “Knowledge, Skills, & Education” indicated by (Health Education & Awareness, Knowledge of Food pyramids, Employment levels, Level of Education).

2.6.2 Structural Equation Modelling

A measurement error causes endogeneity. The measurement errors cause the coefficient estimators to be inconsistent and biased toward zero, which is referred to as attenuation bias (Greene, 2012). The advantage of using an analysis that considers the latent variable will take into account the measurement error (Chin et al., 2003; Lim and Melville, 2009). the following considers the method we employ in this study.

Structural equation modeling (SEM) is a technique that simultaneously examines the interrelationships among multiple variables in a set of equations. The SEM method draws from multivariate factor analysis and multiple regression analysis (Hair et al., 1998; Byrne, 2013; Duncan, 2014). This type of analysis includes regression and path analysis, which examines causal relations between variables by using multiple regression, simultaneous econometric equations, and latent growth curve models (Goldberger, 1972; Cobas et al., 1996; Olsson et al., 2000; Grace, 2006; Irving et al., 2012).

The SEM was first noticeable, back to the path analysis work of Wright (1918, 1934) and in the 1960s by Blalock (1964) and Duncan (1966). Later, SEM was developed in econometrics by Pearl (2000, 2009) and Ullman and Bentler (2012). Additionally, Zhang et al. (2019) expanded the SEM to a more extensive set of causal effects than previously achieved with standard methods. Their contribution examined a set of interactions between one or more independent variables on one or more dependent variables simultaneously, either continuous or discrete.

The SEM has been used in economics, ecological, social, and political studies in recent decades. In econometrics and food marketing (Goldberger, 1972; Muthen, 1983; Bollen, 1989; Lai and Bessler, 2010; Bianchi, 2017; Cooper, 2017; Sarnacchiaro and Boccia, 2018); food access studies (Gustat et al., 2015); food and health economics and human health research (Muthen, 1984; Cobas et al., 1996; Botonaki et al., 2008; Kröller and Warschburger, 2009; Hsu et al., 2016); Economic Research Service (ERS) studies, agricultural futures contract, and climate change studies (Pennings and Leuthold, 2000; Alinovi et al., 2009; Grace et al., 2010; Mosheim, 2012; Smith et al., 2014; Van der Linden, 2014; Eisenhauer et al., 2015), and political science (Shook et al., 2004; Shiftan et al., 2008; Olivas et al., 2013).

There are four motives behind the use of SEM in this study. First, the SEM can handle several equations with several explanatory variables in each equation, where the dependent variable in one equation might be an independent variable in other equations and can be called either an endogenous or an exogenous variable (Smith et al., 2014).

Second, the SEM differs from the usual single equation regression model with a single dependent variable and independent variables that could only test part of the model. Third, a unique benefit of SEM is the capability to test various hypotheses in a multi-equation framework and test construct-level hypotheses at a construct level (Pearl, 2012). Fourth, the SEM handles complex interrelations with multiple dependent variables because it accounts for measurement errors. It provides more accurate diagnostics for model improvement (e.g., fix weak measures), reducing collinearity problems, and representing a model that defines the entire set of relations (Browne et al., 1993; Cudeck et al., 2001).

The SEM contains two major components: (1) the structural model and (2) the measurement model (Loehlin, 1987; Hair et al., 1998). The *structural model* represents the path model, which links independent variables to dependent variables. The structural model shows the potential causal dependencies between endogenous and exogenous variables. *The measurement model* enables us to use several indicator variables tagged with measurement errors to reflect a latent variable. In other words, it connects the latent variables with their indicators considering the measurement errors.

2.6.3 Measurement Error and Reliability

The SEM accounts for the measurement errors to provide much more accurate estimates of the interactions between constructs. The interrelationship between the indicator variables and the latent variable is modeled through regression models to account for these measurement errors (Sobel, 1982; Browne et al., 1993; Bollen, 1998).

Nevertheless, we aim to reduce measurement errors. When the interaction among latent variables are measured, the interrelationship is clear from the measurement error because the error has been estimated and omitted, giving only a common variance, and this is one of the benefits of using the SEM (Finch et al., 1997; Shook et al., 2004; Kenny, 2019).

Reliability and measurement error are inversely related. High reliability is linked with lower measurement error, which means constructing latent variables reveals more of the variance in each indicator, e.g., more of a result is defined.

A regression coefficient is made of two parts: the true structural coefficient, which estimates the relation between dependent and independent variables, and the reliability of an indicator variable. However, the causal relationships in SEM between latent constructs will estimate the true structural coefficient $\hat{\beta}_{SEM}$ instead of the observed regression coefficient $\hat{\beta}$ compared to those found with more straightforward approaches. The effect of measurement error can be explained in the regression $\beta = \beta_{SEM} * \rho_{Reliability}$. From this expression, if the reliability of the indicator variable ρ_x is 100% (e.g., no measurement error at all), the observed correlation (and the resulting regression coefficient $\hat{\beta}$) will underestimate the true relationship. That is, the SEM automatically corrects and takes into account the amount of measurement errors and estimate what the relationship would be if there were no measurement errors[7] (see Caspi et al. (2012); Irving et al. (2012); Penney et al. (2014); Cobb et al. (2015); Pitt et al. (2017) for more details).

Figure 2.5 presents the structure of SEM employed in this study. Two different shapes, including oval and rectangle, are demonstrated. Each oval shape denotes a latent variable connected with three to four observed variables, i.e., rectangle shapes. Figure 2.6

[7] see Appendix 4.A for more details on the measurement errors

specifies the components of structural and measurement models. The measurement model shown in the left part illustrates each latent variable with its indicators. The structural model contains the relationships among the six latent variables, shown on the right side. An arrow at either end means a relationship between the variables exists, and any shapes that have one-way arrows aiming at them are dependent variables in the model.

2.6.4 Model Framework

Following Hair et al. (1998); Pennings and Leuthold (2000); Cudeck et al. (2001); Grace (2006); Duncan (2014), the structural equation model can be expressed as:

$$\mathbf{Y} = \alpha + \mathbf{B}Y + \mathbf{\Gamma}X + \zeta \qquad (2.3.a)$$

$$\mathbf{x} = \tau_x + \mathbf{\Lambda}_x\xi + \delta \qquad (2.3.b)$$

where $\mathbf{Y}$ in Equation $2.1.a$ denotes the healthy food environment which consists of a vector of all endogenous variables $\mathbf{Y} = \begin{bmatrix} y \\ \eta \end{bmatrix}$; y represents the observed endogenous variables, and η is a vector of m latent endogenous variables, where η_1 and η_2 denote the external and internal food environments, respectively (see Table 2.3). $\alpha = [\alpha_i]$ is the vector of constants for the endogenous variables; $\mathbf{B} = [\beta_{ij}]$ is an $m \times m$ matrix of coefficients on endogenous variables η's predicting other variables $\mathbf{\Gamma} = [\gamma_{ij}]$ is an $m \times n$ matrix of coefficients relating ξ to η as $(q \times p)$ of exogenous to endogenous variables; the vector of all exogenous variables is $X = \begin{bmatrix} \mathbf{x} \\ \xi \end{bmatrix}$, ξ is a vector of n latent exogenous variables. $\mathbf{x} = (x_1, \cdots, x_{24})'$ is a vector of x indicators; $\zeta = (\zeta_1, \cdots, \zeta_q)'$ denotes the vector of errors where $Cov(X, \zeta) = 0$. Equation $2.1.b$ is the measurement model, where $\tau_{\mathbf{x}}$ is vector of p intercepts for x indicators; $\mathbf{\Lambda}_x$ matrix of loadings $(n \times p)$ corresponding to the latent exogenous variables, and. $\delta = (\delta_1, \cdots, \delta_p)'$ is a vector of residuals for exogenous variables; $\epsilon = (\epsilon_1, \cdots, \epsilon_q)'$ is a vector of residuals for endogenous variables where ϵ is uncorrelated with δ.

The mean vector of the endogenous variables, the variance matrix of the endogenous variables, and the covariance matrix between the endogenous to exogenous variables are displayed, respectively, in Equations 2.1. c, 2.1. d, and 2.1. e, as:

$$\mu_Y = \mathrm{E}(Y) = (\mathrm{I} - \mathbf{B})^{-1}(\mathbf{\Gamma}_\kappa + \alpha) \qquad (2.3.\,c)$$

$$\Sigma_{YY} = \mathrm{Var}(Y) = (\mathrm{I} - \mathbf{B})^{-1}\ (\mathbf{\Gamma}\Phi\mathbf{\Gamma}' + \Psi)\{(\mathrm{I} - \mathbf{B})^{-1}\}' \qquad (2.3.\,d)$$

$$\Sigma_{YX} = Cov(Y, X) = (\mathrm{I} - \mathbf{B})^{-1}\ \mathbf{\Gamma}\Phi \qquad (2.3.\,e)$$

where Σ stands for the sample variance-covariance matrix; Φ is $n \times n$ matrix of exogenous latent ξ; Ψ is $m \times m$ matrix of ζ. Allow $\kappa = [\kappa_j] = \mathrm{E}[X], \Phi = [X] = Var(X)$, and $\Psi = [\psi_{ij}] = Var(\zeta)$. The mean vector and the variance matrix are:

$$\mu = \mathrm{E}(Z) = \binom{\mu_Y}{\kappa} \qquad (2.3.\,f)$$
$$\Sigma = \mathrm{Var}(Z) = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma'_{YX} & \Phi \end{pmatrix}$$

where the vector of all variables is represented as $Z = \begin{bmatrix} Y \\ X \end{bmatrix}$ (See Appendix 4.C for details on the SEM and the system of equations).

2.6.5 The Maximum Likelihood Estimation

Besides the importance of using the robust estimators in our estimation to cluster the standard errors to correct for heteroskedasticity, there is also a significant interest in adopting the maximum likelihood estimation (Lim and Melville, 2009). The maximum likelihood estimation is one of the most popular methods to empirically assess complex theories (Jöreskog, 1970; Rossi, 2018). The SEM uses the maximum likelihood estimation to compare the elements of the observed variance-covariance matrix to that expected, given the specification of the model:

$$\theta = \begin{pmatrix} vec(\mathbf{B}) \\ vec(\mathbf{\Gamma}) \\ vech(\Psi) \\ vech(\Phi) \\ \alpha \\ \kappa \end{pmatrix} \qquad (2.4.a)$$

where θ the vector of unique model parameters in B, and $\varGamma$. vec is vec operator, and $vech$ is the half-vectorization. Under the multivariate normal distribution, the log-likelihood for θ is:

$$log\, L(\theta) = -\frac{w}{2}\left[k\, log(2\pi) + log\{det(\varSigma_0)\} + tr(D{\varSigma_0}^{-1})\right] \qquad (2.4.b)$$

where k is the number of observed variables, $\varSigma_0$ is the submatrix of $\varSigma$ corresponding to the indicator variables, and tr refers to the trace of a matrix. The standardized parameter estimates are $\tilde{\beta}_{ij} = \hat{\beta}_{ij}\sqrt{\frac{\hat{\sigma}_{ii}}{\hat{\sigma}_{jj}}}$, $\tilde{\gamma}_{ij} = \hat{\gamma}_{ij}\sqrt{\frac{\hat{\Phi}_{ii}}{\hat{\sigma}_{jj}}}$, where $\hat{\sigma}_{ii}$ is the ith diagonal element of $\hat{\Sigma}_{YY}$.

2.7 Results

Table 2.4 reports results for the case where SEM is applied to the total sample size. Results for the classified two groups based on their food security status are reported in Table 2.5. The casual relationships between the latent variable and its indicators are presented in Table 2.6. This study examines the impact of internal factors, including affordability, convenience, and desirability of healthy food, and external factors including food prices, availability of healthy food, location, store, and product properties, on the likelihood of individuals adopting healthier eating habits according to the HEI score.

For each model fit, we checked whether the results align with the recommended statistical criteria and overall significance. The fit of the model is assessed using different types of fit indices and evaluated by the Tucker Lewis Index (TLI), Root Mean Squared Error of Approximation $(RMSEA)$, and Standardized Root Mean Square Residual $(SRMR)$[8]. The CFI value is 0.908 and implies a good model fit, based on Hu and Bentler (1999) and Bentler (1990) implication that a value of 0.9 or greater implies a good model-data fit. The results of the $RMSEA$ is 0.034 and indicate that the fitted model does estimates the population covariance matrix per degree of freedom[9]. The $SRMR$ result is equal to 0.041, which is considered an acceptable fit when the value is smaller than 0.08. The overall models are excellent, and the goodness-of-fit results show that the SEM captured important associations among the variables, according to Hu and Bentler (1999) and Kline (2011).

Except for food availability, all domains of internal and external food environment are statistically significant in explaining the probability of adopting healthy eating habits. Concerning the external environment, the price of healthy food items considerably influences the external food environment. As expected, the sign of the coefficient of price is negative. Households will be less likely to eat healthily if healthy food becomes expensive. Nonetheless, the coefficients for the whole sample and food security status are

[8] See Appendix 3.C for more details on the model validation and cutoff criteria.
[9] (a value below 0.08 implies a close fit as advised by Hair et al. (1998); Hu and Bentler (1999); Shiftan et al. (2008)

identical. With the full sample, one of the indicators indicating "if healthy food is costly" is significant at 1% level, and among food-insecure households and food-secure households. The effect of price is greater when a household is a food insecure since the size coefficient of affordability lowers the possibility of adopting healthy eating habits, subsequently causing the HEI score to decrease by 0.6%.

Consumer preference is highly influenced by the latent variables representing product and store properties (significant at 1% level). It is expected that high-quality products and a variety of food choices will have a positive impact on the HEI. According to all sample sizes, high-quality products and a wide variety of food choices account for 2.4% of the latent variable's impact, increasing the likelihood of adopting healthy eating as measured by the HEI score by 0.316%.

Healthy foods affect an individual's food environment. For the full sample and according to the food security status, food accessibility was significant at the 1% level except for food-insecure individuals. The size coefficient of individuals experiencing food insecurity is the highest. While food insecure individuals have the lowest size coefficient.

When it comes to food accessibility, food outlet density has been significant at 1%. A food center specializing in fruit, vegetables, meat, poultry, or seafood is an indicator of easy access to a healthy diet. The addition of 1% of these stores will increase the accessibility of healthy foods by 2.5%, contributing to an increase of 3.8% in the HEI score.

Like food access, it is observed that a healthy food environment is positively affected by food affordability. An increase of 1% in food affordability will increase HEI scores by 1.02%. Households experiencing food insecurity have poor affordability (0.5%), while households experiencing high food security have 1.3% more purchasing power than food insecure individuals.

Regarding SNAP participation and healthy eating, a healthy diet cannot be guaranteed when taking part in the program. The finding shows that the HEI for SNAP participants is lower by 0.45% than non-participants. We did not expect this finding to be in favor of non-participants. As far as affordability of healthy food is concerned, the findings show that unexpected large-in-house expenses caused the HEI to fall by 1.04 percent.

Health education and awareness is significant at the 1% level for the total sample and food secure households. Food insecure households were significant at 10%. When compared with high education individuals, low educational levels directly reduce the latent variable healthy food knowledge and awareness by 7.2%. Comparatively, participating in nutrition education events, seeking nutrition information on the internet, or attending nutrition events increases knowledge about healthy food by 10%.

Considering the previous indicators, increased knowledge of healthy foods is the most critical aspect of a healthy eating pattern. Increased knowledge of healthy foods leads to a 6.3% improvement in consumer HEI scores. It is expected that preparing healthy dishes and the taste of healthy foods will increase the HEI by 7.4% for in-house members. People with low food security will experience a greater impact of 8% than those with a high level of food security.

2.8 Discussion and Conclusions

In this study, we examined how the food environment affects healthy eating habits. The primary objective of this study was to explore the relationship between the food environment's internal and external characteristics and households' healthy eating behaviors in the US. The FoodAPS data was employed in this paper, and the sample size consisted of 3,861 households. A Healthy Eating Index based on Harvard's healthy eating pyramid and neighborhood food environment index was employed to investigate the potential of adapting healthy eating habits from the current healthy food environment. The SEM method was used to quantify the household variables to identify the unobserved food environment attributes on adopting healthy food consumption habits.

Based on the empirical results, a significant effect of incentive on adopting healthy eating behaviors can be observed. According to the total sample, all the domains of the internal food environment had a positive influence on healthy eating habits. In addition, the size of the coefficient estimate within the internal food environment is considerably larger than the coefficient within the external food environment. This emphasizes the importance of the internal food environment in relation to the external food environment.

The coefficient estimate of food secure households versus food-insecure households is larger for most variables except desirability when comparing between food security groups. Therefore, food-secure households are more responsive to improving the external and internal food environment when it comes to adopting healthy eating habits. A discussion of this finding and its relevance to previous research will follow.

For food insecure individuals, the price of healthy food products was an essential factor in making healthy food choices. In addition, they allocate healthy food within the desirable range that maximizes their utility (Saghaian and Reed, 2004). As the desirability domain is the largest among food-insecure individuals, the impact of prices is more significant. As Hendrickson et al. (2006) and Drewnowski and Specter (2004) indicated, price is an essential factor for consumers in making food choices. In addition, the results emphasize that consumers care more about the quality of food items, regardless of their demographic characteristics. This result aligns with findings from Cummins et al. (2014) and Hendrickson et al. (2006).

Contrary to expectations, this study did not find a positive correlation between SNAP participation and the HEI score. The HEI of SNAP participants is 0.45% lower than that of non-participants. However, these findings are in accordance with Gregory et al. (2013), who pointed out that SNAP participants have a lower HEI score, with a 2.5% lower than non-participants. It might be the case that purchasing healthier food items is costly, and those items can significantly impact a household's overall HEI score. With these findings, we realize better that Carlson and Keith-Jennings (2018), Nielsen et al. (2017), and Prathap (2018) observations confirmed the existence of non-food hardship among families. This observation is supported by Schanzenbach et al. (2016), who stated that participation in SNAP is common among low and very low food security individuals because SNAP improves their purchasing power.

A healthy eating pattern begins with knowledge about healthy foods. The latent variable 'healthy food knowledge and awareness' is directly affected by low education levels. This impact is large among food-insecure households. This result confirms that food education/awareness makes people more aware of their options. These conclusions, which Draper and Younginer (2020), Reel and Badger (2014), and Guthrie et al. (1999) discussed in the literature, added weight to our argument that awareness of food nutritive value aims to enhance the in-house food environment. It is necessary to develop a program that targets households with low education levels to improve their food security status[10] (Hite, 2019). Besides, more nutrition information on product labels significantly impacts the product properties, which shows a great indication to a stockholder to consider, as suggested by Dallas et al. (2015) and Neuhofer et al. (2020).

Regarding the latent variable defining the convenience and desirability of healthy food, questions about the time and effort involved in making healthy meals and the taste of healthy foods show a 1.1% decrease. As a result of convenience and the desire to consume healthy foods, the HEI of a person will increase by 7.4% based on the total sample. The impact on individuals who are food insecure is 8% greater than those who are food secure. This remark agrees with Backett-Milburn et al. (2010) and Scaglioni et al. (2008), who pointed out that some families with children struggle to familiarize themselves with healthy food dishes due to tastes.

[10] Through participation in healthy eating plans that are consistent with the Dietary Guidelines for Americans (HHS, 2015). A State may participate in SNAP Education (SNAP-Ed) because it remains an optional program.

2.9 Policy Implications and Limitation

In this study, we examined the causal relationships among latent and observed variables while simultaneously accounting for measurement errors that lead to problems with endogeneity. We studied 3,861 US households to understand the healthy food environment from the perspective of their community and consumers. A healthy eating index and a healthy food environment index were used to evaluate the quality of food environments among food-secure and food-insecure households.

Several practical implications can be drawn from the findings of this study. Policy makers should encourage SNAP recipients to adopt a healthy lifestyle and to eat healthy foods. Continued efforts among states are needed to make nutrition education to SNAP beneficiaries more accessible.

A policymaker should consider adjusting the relative price of healthy foods for consumers with limited access to healthy foods. An important practical implication is that a SNAP-authorized retailer requires selling healthy foods for home consumption and preparation.

Despite some key results and comprehensive efforts, this study has limitations and caveats that need to be noted. The study was limited by data regarding the domain "marketing and regulation" within the external food environment dimension. It contains information regarding promotional information, branding, advertising, sponsorship, and policies.

The SNAP-Ed program could be examined in greater detail to provide interesting findings related to obesity and food insecurity. Future qualitative study, e.g., Difference-in-Difference estimation, to examine the impact of SNAP-Ed program implementation on food security and obesity rates in a state government.

2.10 Tables and Figures for Chapter 2

Table 2.1 Definitions of latent and its indicator variables

Latent variables	Indicator variables description
External Food Environment	
Food Availability* (Hsu et al., 2016; Ziliak and Gundersen, 2016)	Variables reflect each outlet's density, including food desert and food swamp.
Prices and Monetary value (Deaton, 1989; Gundersen et al., 2011; 2017)	Price per unit for healthy food products.
Store/Product properties* (North et al., 1999; Pitt et al., 2017)	Store's features include food quality and variety of food selections, low price stores, and short distance.
Regions	Census regions such as South, West, Northeast, and Midwest; Urban/Rural.
Internal Food Environment	
Food Accessibility* (Breyer and Voss-Andreae, 2013; Pitt et al., 2017)	Variables characterize the physical distance of food stores and fast-food restaurants in miles. Transportation means used to allocate foods.
Food Affordability* (Carson, 2002; Palameta and Macredie, 2005)	Variables represent the household income, purchase power, expenditure, financial condition, and participation in food assistance programs.
Convenience and Desirability* (Guthrie et al., 1999; Botonaki et al., 2008)	Reasons for not consuming healthy food products. Such as if healthy food taste bad, healthy food requires extra time to cook and prepare healthy dishes.
Knowledge, Skills and Education* (Guthrie et al., 1999; Scaglioni et al., 2008; Neff et al., 2019)	Health education, knowledge, and experience for adopting healthy food eating to improve nutrition and well-being. Participation in USDA programs such as follow MyPlate and MyPyramid guidelines.

Variables marked with “*” are latent and unobserved.

Table 2.2 Definitions and descriptive statistics of the variables used in the analysis.

Variables	Definitions	Full Sample (N=3,861)	High food security (N=2,539)	Low food security (N=1,322)
External Food Environment				
PRICE *	Price per unit of Healthful food categorized by HHEP	0.19	0.19	0.18
ENV_SCORE ***	Nutrition Environment score index	7.57	7.6	7.52
FOOD_SWAMP ***	The density of fast food/junk food and corner stores	2.35	2.31	2.5
FOOD_ STORE *	The density of Food Stores	4.04	4.3	3
STORE_PRO1 ***	Prices are lower than others	0.6	0.57	0.64
STORE_PRO2 ***	The store is nearby	0.5	0.52	0.47
STORE_PRO3 **	Store/Product feature a Good and high-quality product	0.33	0.34	0.31
STORE_PRO4 ***	Shop for another reason such as hours of operation, accept EBT	0.28	0.3	0.24
RURAL ***	Rural = 1, 0 otherwise	0.27	0.29	0.24
WEST ***	West region = 1, 0 otherwise	0.23	0.21	0.25
NORTH ***	Northern region = 1, 0 otherwise	0.17	0.19	0.14
SOUTH ***	South region = 1, 0 otherwise	0.36	0.33	0.41

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 2.2 Continued

Variables	Definitions	Full Sample (N=3,861)	High food security (N=2,539)	Low food security (N=1,322)
MIDWEST ***	Midwest region = 1, 0 otherwise	0.25	0.27	0.2
Internal Food Environment				
HEI_SCORE ***	The Healthy Eating Score from 1% to 100%.	51.14	51.81	49.8
SNAP ***	SNAP participated= 1, 0 otherwise	0.30	0.20	0.50
SPEND_PWR ***	The purchasing power of household within income and family size	2185	2630	1330
LARGE_EXP ***	If household had a large or unexpected expense =1, 0 otherwise	0.05	0.004	0.14
OVERWEIGHT	Individual's weight: Overweight =1, 0 otherwise	0.45	0.44	0.47
OBESE	Individual's weight: Obese=1, 0 otherwise	0.49	0.48	0.50
NOT_OVERWEIGHT	Individual's weight: not overweight =1, 0 otherwise	0.55	0.552	0.553
PRE_COMP_STO ***	If Individual shopped for food at a convenience store=1, 0 otherwise	0.34	0.32	0.37
PRE_BIGBOX ***	If Individual shopped for food at a big box store=1, 0 otherwise	0.43	0.44	0.41
PRE_WAREHOUS ***	If Individual shopped for food at a wholesale club=1, 0 otherwise	0.22	0.26	0.16

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 2.2 Continued

Variables	Definitions	Full Sample (N=3,861)	High food security (N=2,539)	Low food security (N=1,322)
PRE_SPECIAL_STO ***	If household shopped for food at specialized food stores =1, 0 otherwise	0.45	0.6	0.3
EDU_LOW ***	Level of education: High school or less =1, 0 otherwise	0.34	0.28	0.47
EDU_HI ***	Level of education: Bachelor and higher =1, 0 otherwise	0.3	0.38	0.14
EDU_MED ***	Level of education: Some College =1, 0 otherwise	0.36	0.35	0.39
PYRAMID ***	Knowledge of Food pyramids =1, 0 otherwise	0.27	0.29	0.24
HEALTHY_AWARE	If health Educated & Awareness to eat healthy =1, 0 otherwise	0.02	0.012	0.006
GROC_LIST ***	Individual shops with a grocery list =1, 0 otherwise	0.70	0.73	0.66
H_COSTLY ***	If healthy foods cost too much=1, 0 otherwise.	0.44	0.3	0.72
H_TASTE ***	If healthy foods do not taste good=1, 0 otherwise.	0.12	0.09	0.18
TIME_PREPAR***	Healthy foods preparation requires a lot of time =1, 0 otherwise	0.19	0.17	0.23

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 2.3 The abbreviation of indicators and latent variables.

Abb.	Variable Description	Abb.	Variable Description
η_1	Internal Food Environment**	η_2	External Food Environment**
ξ_1	Knowledge, Skills, & Education*	ξ_5	Product & Store Properties*
x_1	Health Education & Awareness	x_{15}	Least Expensive
x_2	Employment levels	x_{16}	Closest to home
x_3	Knowledge of Food pyramids	x_{17}	Good Quality products
x_4	Level of Education	x_{18}	Other properties
ξ_2	Convenience & Desirability*	ξ_6	Food Availability*
x_5	Taste issues	x_{19}	Food Swamp density
x_6	Individual reasons	x_{20}	Food Environment score
x_7	Time issues	x_{21}	Grocery Store Density
		x_{22}	Food Deserts density
ξ_3	Food Affordability*		
x_8	Purchase power	x_{23}	Monetary value
x_9	SNAP participation	x_{24}	Regions
x_{10}	In-house large expenses	Y	Healthy Eating Index Score**
x_{11}	Out-of-pocket costs		
ξ_4	Food Accessibility*		
x_{12}	Distance to food outlets		
x_{13}	Types of stores		
x_{14}	Transportation means		

* Latent exogenous variables

** Latent endogenous variables

$\mathbf{x} = (x_1, \cdots, x_{24})'$ are indicator variables

Table 2.4 Regression results of the Structural equation model for the total sample.

	Probability of adapting healthy eating habits from the current healthy food environment						
Latent variables	External Food Environment			Internal Food Environment			
Total sample (N= 3,861)	Food Availability	Prices	Store Properties	Food Access	Affordability	Health Education	Desirability
Coefficients	0.183	-0.351***	0.161***	0.195***	0.522***	0.3224***	0.381***
Std. Error	1.187	0.0046	1.48	0.638	0.855	0.547	0.814
Fit Statistics	$\chi_m{}^2$(d.f.) =1271.04 (233)		RMSEA = 0.034	SRMR = 0.041		CFI = 0.908	TLI = 0.882

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001.

Table 2.5 Results of the SEM for groups classified according to their food security.

	Probability of adapting healthy eating habits from the current healthy food environment						
Latent variables	External Food Environment			Internal Food Environment			
High Food Security (N= 2,539)	Food Availability	Prices	Store Properties	Food Access	Affordability	Health Education	Desirability
Coefficients	0.194	-0.347***	0.175***	0.278**	0.671**	0.301***	0.365***
Std. Error	1.590	0.00459	1.654	0.884	1.641	0.567	1.131
Low Food Security (N= 1,322)	Food Availability	Prices	Store Properties	Food Access	Affordability	Health Education	Desirability
Coefficients	0.206	-0.347***	0.130***	0.161	0.268**	0.192*	0.412***
Std. Error	1.644	0.00459	2.257	1.70	0.915	0.78	0.88
Fit Statistics	χ_m^2(d.f.) =2917.68 (510)		RMSEA = 0.040	SRMR = 0.079		CFI =0.852	TLI =0.826

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001.

Table 2.6 Results of the Measurement Models of latent variables and the indicators.

Latent variable	λ	**Latent variable**	λ
Knowledge, Skills, & Education*		Product & Store Properties*	
Health Edu. & Awareness	0.0109* (2.02)	Least Expensive	-0.297** (-3.28)
Level of Education	-0.728*** (-6.19)	Closest to home	0.228** (2.79)
Knowledge of Food pyramids	0.0373 (1.46)	Good Quality products	0.222** (2.82)
Employment levels	1	Other properties	1
Food Accessibility*		Food Availability*	
Distance to food outlets	0.191 (0.52)	Food Swamp density	-2.632* (-2.02)
Types of stores	0.223*** (5.35)	Food Environment score	0.313 (0.58)
Transportation means	1.050 (1.82)	Food Deserts density	0.168 (0.00)
Other shops	1	Grocery Store Density	1
Food Affordability*			
SNAP participation	- 0.45*** (-11.56)		
In-house large expenses	-0.109*** (-4.69)		
Out-of-pocket costs	-0.461*** (-5.67)		
Purchase power	1		
Convenience & Desirability*			
Taste preference	-0.0526* (-2.55)		
Time prepares healthy food	-0.111** (-2.74)		
Individual reasons	1		

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. t-value in parentheses.

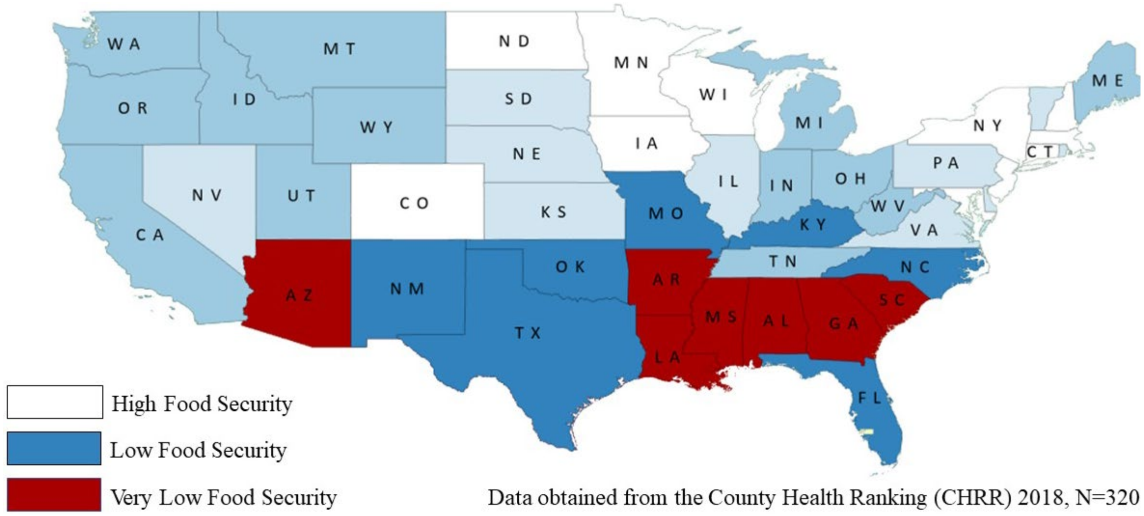

Figure 2.1 Spatial distribution of food insecurity in the US

Note: To better understand food insecurity in the US, we used data from County Health Rankings (CHR, 2019) collected from 230 counties across all 50 states. The levels of food insecurity are generally higher in the southern region, and this disparity becomes more pronounced over time among children and older people. Mississippi, Louisiana, Alabama, Arizona, and Arkansas, had the highest levels of food insecurity, while Virginia, North Dakota, Minnesota, Colorado, Iowa, Kansas, and New York had the lowest levels of food insecurity.

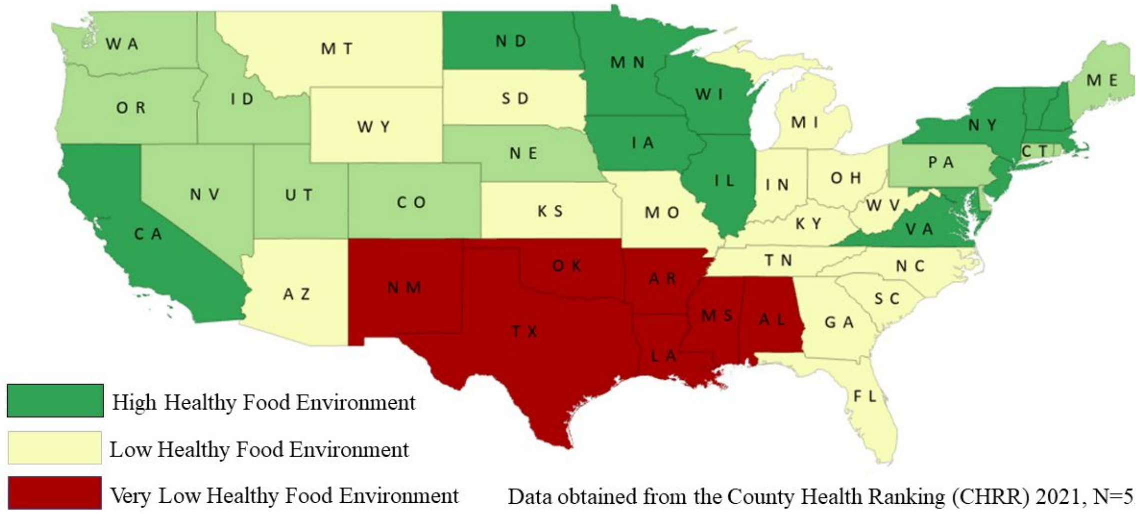

Figure 2.2 The geographic status of healthy food environments in the US

Note: Southern regions generally have less of a healthy food environment. Concerning diet and food environment measures, South Dakota, Georgia, Mississippi, and Arizona lack healthy food environments. In contrast, Virginia, New York, North Dakota, and Kansas have the highest food environments (CHR, 2021).

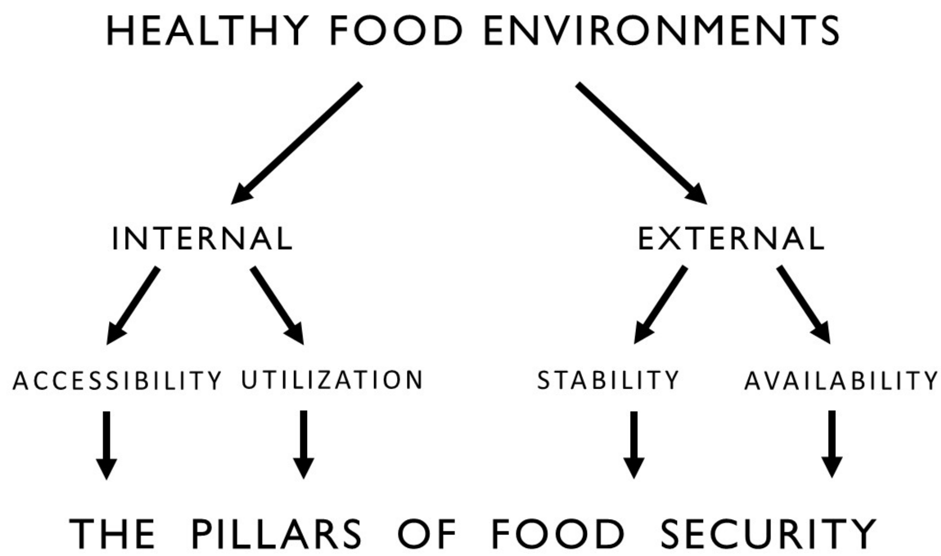

Figure 2.3 The path from the healthy food environment to the pillars of food security.

Note: A healthy food environment is foremost to establish high food security as demonstrated between the domains and the four pillars of food security (accessibility, availability, utilization, and stability).

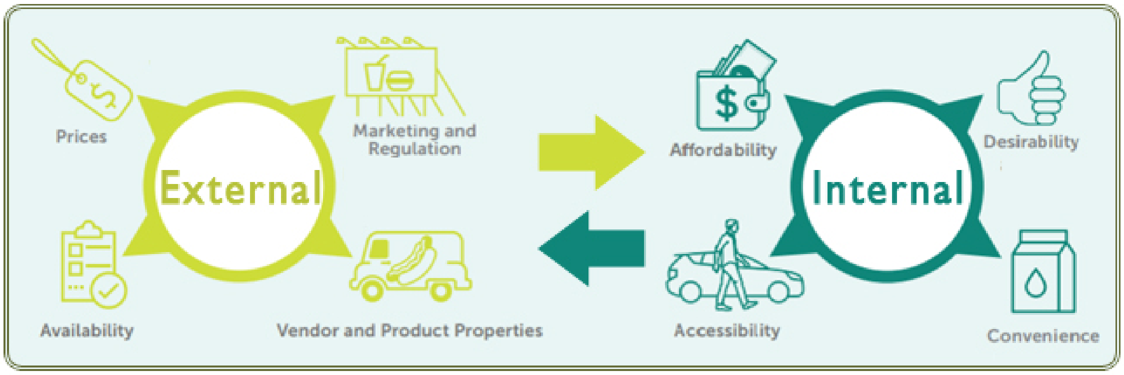

Figure 2.4 Food environment domains and dimensions.

Note: This Figure illustrates the interface of the food environment within the broader food system and the interaction of external/personal domains with people's food purchases. Source: (Turner et al., 2017b)

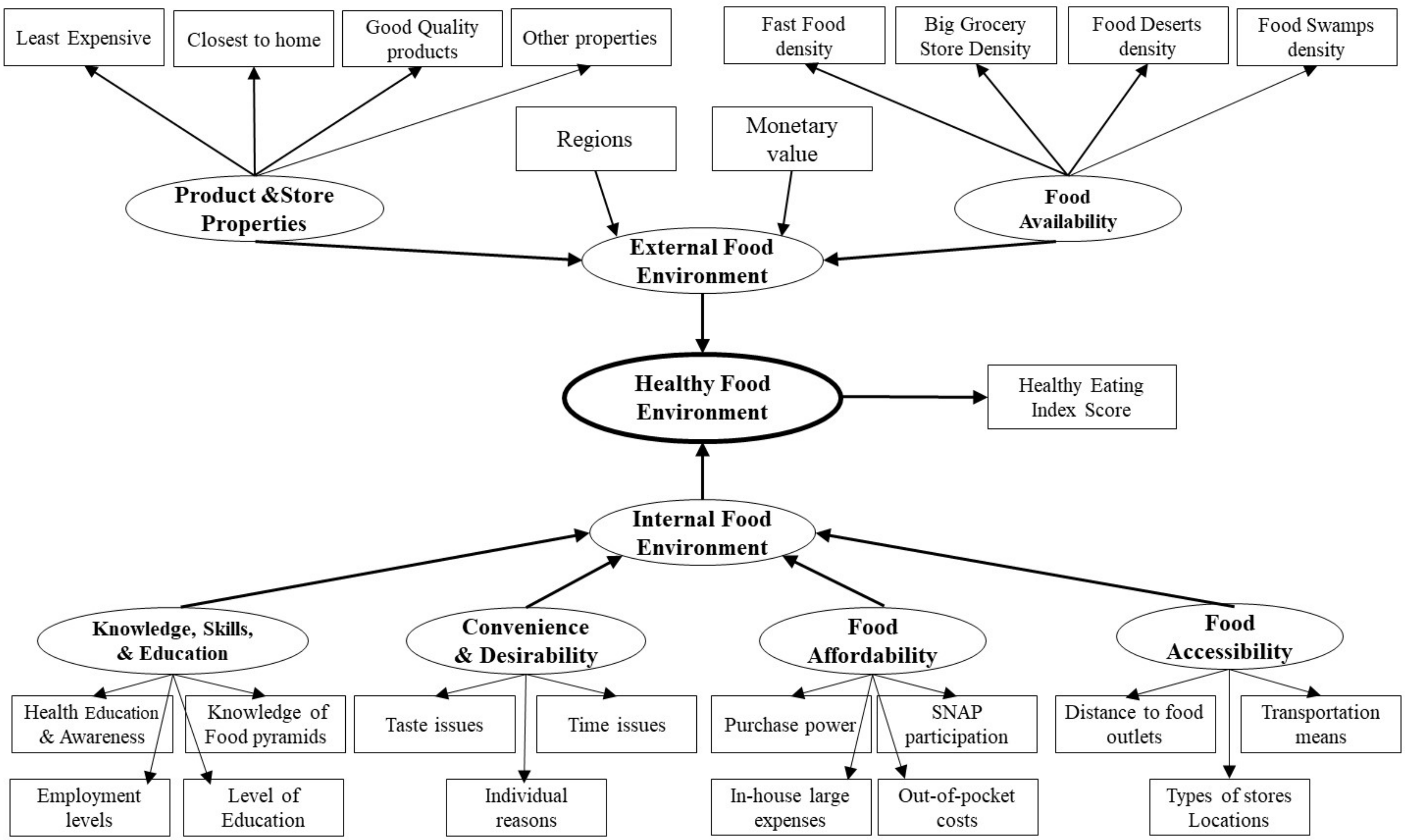

Figure 2.5 Diagram of SEM of the external and internal food environment.

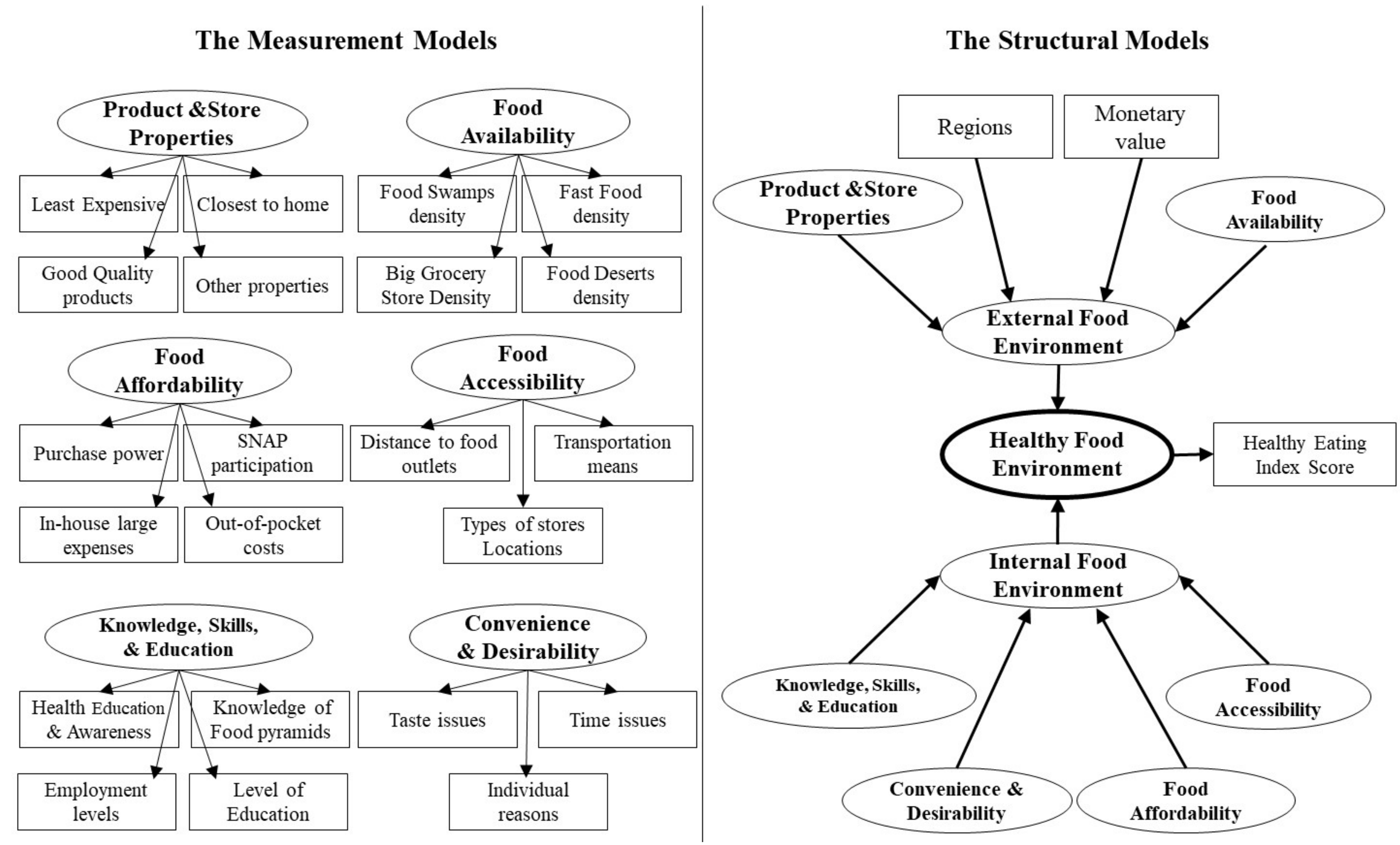

Figure 2.6 The Structural and Measurement models in SEM.

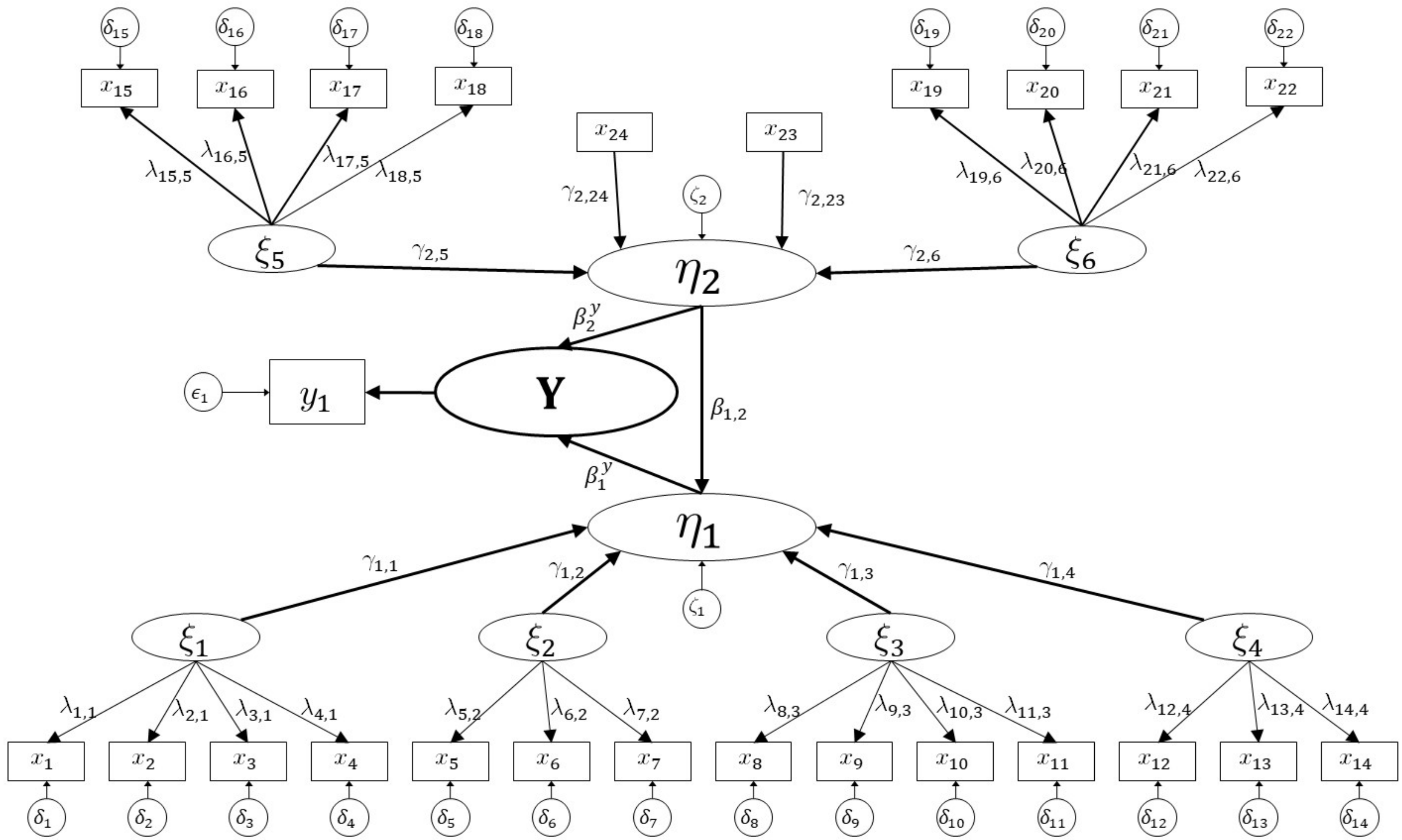

Figure 2.7 Illustrative model of relationships among the internal and external factors.

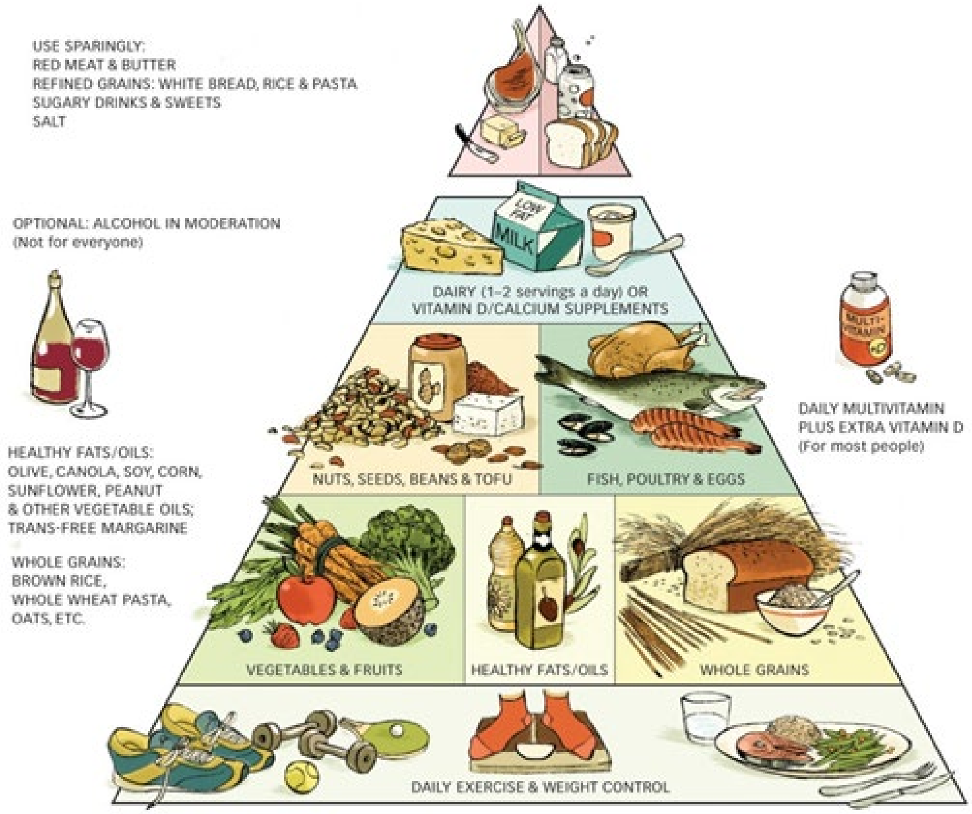

Figure 2.8 The Harvard's Healthy Eating Pyramid

Note: The HHEP also includes other healthy lifestyle elements such as multivitamin supplements, vitamin D, weight-loss programs, sports, and small alcoholic beverages.

Source (Willett and Skerrett, 2017)

Chapter 3 Does The Local Density of Food Retail Outlets Affect Access to Healthy Food?

Abstract

Lack of access to healthy foods is one of the major obstacles to forming healthy eating habits. Economists are increasingly paying attention to how these barriers affect residents' food purchase behavior. This study analyzed how better food access affects food consumption by separating the demand-side effect from the supply-side effect. Specifically, we studied the influence of the density of retail stores on residents' produce consumption through a price-reduction effect and a travel-distance reduction effect. We used data from the USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). We employed a Three-Stage Least-Squares estimator (3SLS) in a total sample of 4,126 American households. Findings show that the healthy food sources stores improve consumers' perception of food availability and accessibility. Reduced prices and reduced travel distances indirectly influence the choice of healthier foods. In particular, supermarkets or supercenters increase consumers' food purchases when compared to small grocery stores. Furthermore, direct marketing farmers and farmer markets facilitate the purchase of food by rural consumers. Findings indicate that large grocery stores and specialized food stores play a significant role in purchasing healthy foods. In rural areas, direct marketing markets are more dense than in urban areas. Car accessibility has been found to have a significant impact on consumers within 15 miles of a healthy store. This study shows that household food quantity responds to significant increases in retailer numbers, mediated by differences in travel distance and prices. Results provide insight into what affects food security inside and outside the homes of consumers. Further, the results highlight the impact of the availability of healthy food on both urban and rural areas. This study has important implications for improving health, sustaining nutritional diets, and enhancing regional welfare in the future.

KEYWORDS: Food Stores, Price Reduction Effect, Distance Reduction Effect, Three-Stage Least-Squares Estimator, 3SLS, Two-Stage Least-Squares Estimator, 2SLS

3.1 Introduction

US retailers have been experiencing changes over the past few decades. Consumer preferences and competition have been significant contributors to these changes. Among the most notable traits of the US industry is the growing power of large national chains (Powell et al., 2007; 2013). The food retail growth rate has been accelerated because food sales had risen rapidly from 35% in 1990 to 65% in 2019, and their combined sales have exceeded $410.4 billion (USDA, 2019). Despite this rapid growth, about 55 million Americans do not have access to large grocery stores[1] (Rhone et al., 2017), raising their likelihood of food insecurity by 24 % (Laska et al., 2015).

The availability of adequate disposable income determines the consumption of food. Therefore, food-aid programs, such as the Supplemental Nutrition Assistance Program (SNAP), significantly impact the retail food sector. The literature has revealed that SNAP has played an essential role in improving food expenditure among millions of low-income individuals and families (Nord and Golla, 2009; Bergmans et al., 2018).

However, some studies have found that recipients were more likely to use food assistance benefits to purchase less healthy items (Mancino et al., 2009; McGuire et al., 2010; Beatty and Tuttle, 2015). They argued that although food assistance has a higher effect on caloric intake, it might adversely affect diets that are low in nutrients. The evidence suggests that, in general, food assistance programs have been associated with

[1] See Xin et al. (2019) , Elbel et al. (2020), and Fulfrost and Howard (2006) for details about the obesity-related impact of convenience and small grocery stores.

higher food sales and market share due to the increased purchasing power of low-income families and may result in fewer opportunities to improve healthy food choices.

For decades, food stores have played an essential role in supplying food to SNAP recipients. In return, SNAP generates income for many food retailers and allows small companies to expand. Among the top supermarkets in the retail food sector, Wal-Mart and Kroger comprise over 80% of all food sales and generate more than $15 billion in revenue from USDA food support programs (Clemmitt, 2018; Orleck, 2018).

Generally, retail stores have a significant impact on the community where they are located. Lowering prices and decreasing travel distance may indirectly improve access to healthy food. This study determines the likelihood of improving dietary intake by examining how the density of healthy food resources impacts purchases.

The remainder of this paper is organized as follows. Section 3.2 explores the literature more generally and discussed the gaps that this research contributes to filling. The objectives of this study have been discussed in section 3.3. In section 3.4, we examine the data and descriptive results. Analytical and empirical models are discussed in section 3.5. In section 3.6, our results and discussion are presented. Finally, in section 3.7, we summarize and identify areas for policy implication and future research.

3.2 Literature Review

3.2.1 Neighborhood Disparities

One strand of literature argues that different geographical areas have a different density of food stores. In light of this statement, different studies point out that rural and low-income neighborhoods tend to have fewer large food stores (Bitler and Haider, 2011).

When it comes to healthy food options, a study by Drewnowski et al. (2012) extended their research to high-income and low-income regions. Their findings indicated that one way to improve economic access to healthy food sources could be by improving physical access to food stores. *Healthy food source*s can broadly be defined as large grocery stores, specialized food stores, direct marketing farmers, and farmers' markets. Regarding large and small grocery stores as well as direct marketing farmer and farmers' markets, Taylor and Villas-Boas (2016) found that households tended to shop at large grocery stores, such as superstores and supermarkets, rather than direct marketing farmers or farmers' markets due to travel distance.

Recently, the high quality of local foods has attracted consumers in the US, and consumers can now purchase locally produced food directly from farmers (Plakias et al., 2020). In a Community Supported Agriculture (CSA) program, buyers purchase shares in a farm before planting and then receive a portion of what the farm produces each week. A share of produce provides meat, enough fresh fruits, and vegetables (Brown and Miller, 2008).

Other types of the healthy food source are specialized food stores. The term "*specialized food store*" can describe fruit and vegetable markets, meat markets, and fish and seafood markets that mostly sell healthy food (Gale, 1997). When it comes to the density of food retailers offering healthier options, Yan et al. (2015) found that in urban areas, specialized food stores tend to be widely spread. While in rural areas, a study by Liese et al. (2007) stated that outlets with more affordable and healthy food options began to be replaced by convenience stores or small grocery stores that sell less nutritious food.

Consequently, the location of retail food stores is determined by a number of factors, including consumer demand, price, and location. A dense network of farmers' markets and direct marketing farms is expected to impact rural consumers more than urban areas, while specialized food stores will affect urban consumers. This study investigates whether direct marketing and farmers' markets can play a role at the community level and whether specialized food stores provide healthier food than small grocery stores.

3.2.2 Food Proximity and Mobility

Proximity to food stores is defined as how close the food stores are to consumers (Cleveland et al., 2015). For consumers, it is more convenient for them to take fewer trips to a food store. An introductory study conducted by Blaylock (1989) examined the shopping frequency from a US population survey through a constrained utility maximization model. His findings stated that factors, such as the distance to grocery stores and family size, had statistically significant effects on shopping frequency. Ghosh-Dastidar et al. (2014) found that food consumption is affected by travel distance, and this effect is more substantial for households with lower income. This finding is consistent with results from Blaylock (1989).

Some states have been influenced by food proximity, causing the population to crave supermarket stores to enhance their diet. Ver Ploeg (2010) stated that transport infrastructure is the defining characteristic of small towns and rural areas with limited access to fresh food. Families living far from supermarkets have difficulty meeting their food needs. There was a wide variation in food purchase patterns and healthy food preferences among families with limited mobility in the included studies. For example, Hilmers et al. (2012) established that vehicle ownership affects food purchases compared with public transportation. This claim is consistent with Block and Kouba (2006).

In addition, living near larger stores can help consumers make better food choices Gustat et al. (2015) and decrease the odds of being obese and overweight (Bell et al., 2013). Several studies, including Gustat et al. (2015), Morland et al. (2002), and Zenk et al. (2013), argued that residents of neighborhoods with large grocery stores ate more fruit and vegetables than residents of neighborhoods with low supermarket density. Rose and Richards (2004) noted that the spatial aspect made faraway products undesirable because consumer preference towards food changes with increased travel distance.

Regarding local markets and distance, farmers' markets contribute to people's health because they offer fresh, delicious food and are relatively affordable to their communities (O'Hara and Benson, 2019). In addition to the marketplace, local economies can also be built by direct selling by producers. The increasing number of farmers' markets provides a short distance for local people. It provides healthy food to people living in low-income areas. The density of direct marketing farmers and farmers’ markets in urban and rural consumer markets is expected to vary.

In light of this, residents of areas with limited access to fresh foods may often need to travel to other neighborhoods to find fresh foods in addition to those provided by small stores. Generally, findings from the existing literature indicate that transportation is a challenging issue for many consumers. Thus, we expect that a reduced travel distance coupled with the ownership of a vehicle positively affects consumers' health.

3.2.3 Price Competition

A second strand of the studies examines how new stores affect local farming enterprises. Rural and urban areas may differ in farmers’ markets and direct marketing. Farmers' markets are health-promoting because they sell fresh, delicious food at affordable prices (Brown and Miller, 2008). Producers can also build local economies by selling directly to consumers as an alternative to using the marketplace.

For low-income neighborhoods without access to healthful foods, farmers' markets provide fresh, local fruits and vegetables which are cheap and fresh. Thus, the price of fresh food products has a significant impact on purchase trends. As more food stores are in an area, existing stores compete with one another. An increase in food supply decreases prices, having other factors being constant (Smith, 1776). Consequently, the reduction in price as a result of increased competition can boost consumption. Dong and Lin (2009) estimated that a 10% reduction in fruit and vegetable prices would give low-income families a 5.2% increase in fruit and a 4.9% increase in vegetable consumption. Due to the decline in prices, studies have found that retailers adapt their supply chains to break even (Gutman, 2002; Heinrich and Betts, 2003; Lee, 2004). Hamilton (2018) conducted a parallel study with similar market conditions in the US to show that food venues are likely to continue cutting prices to keep up with the competition. Thus, the price of fresh food products has a significant impact on purchase trends.

Overall, previous studies indicate that store density in different locations can influence the supply and demand of local foods through various factors such as competitive pricing, food production, purchasing patterns, and store density. Furthermore, consumers in rural and urban areas may differ in the benefits due to proximity and price. In this study, we examine the impact of retail store density on healthy food consumption through price reductions and travel distance reduction effects among rural and urban consumers.

3.3 Objectives

This study examines the impact of retail store density on consumption among residents in the US. A retail food store affects the food purchase of a resident through several mechanisms. One is *a price reduction effect*. When a large number of food retailers compete with each other, typically lower food prices will follow. The second is *a travel distance reduction effect*. When a new store opens nearby, customers naturally switch over to the new store. Figure 3.1a and Figure 3.1b provide detailed explanations on the impacts of the number of food stores on consumer healthy food purchases.

We investigate if consumers' healthy food choices can be indirectly affected by lowering prices and reducing travel distances. This study provides information on whether the density of large, small, specialized food stores and farmers' markets would be more effective at encouraging produce consumption. For instance, if a travel-distance reduction is more effective than a price reduction, new stores are more likely to affect consumption in the long run. If the impact of price reduction is more effective than the travel-distance reduction, subsidies, on the other hand, help consumers and producers in the short run. The purpose of this study is to determine the effect of retail store density on US consumption. Our study investigates whether lower prices and shorter travel distances have an indirect impact on consumers' food choices.

3.4 Data and Descriptive Results

3.4.1 Dataset Source

We used data from USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). The data collection is funded by the US Department of Agriculture's Economic Research Service (ERS) and Food and Nutrition Service (FNS). Detailed demographic and socio-economic attributes were collected for each family and its members, including SNAP households. The FoodAPS included nationally representative data from 4,826 American households who completed the survey between April 2012 and January 2013. The total final modified size is 4,126 households selected in this analysis, with actual expenditures recorded on the five food items (dairy, fresh meat, grain, fruits, and vegetables).

The survey was designed to illustrate SNAP households and nonparticipant households in different income groups: low-income and higher-income households. Besides, the survey has information on each family on purchases from thirty-three categories grouped under eight broad food types such as grains, vegetables, fruit, milk products, meat and beans, prepared meals, other or un-coded foods, and the means of payment by all household individuals for seven days (ERS, 2016). Therefore, the FoodAPS survey is ideal for this project because it provides information on large, small, specialized food stores, farmers' markets, geographical location, and proximity concerning the food venue types within a certain radius to construct the model developed in this study.

3.4.2 Descriptive Analysis

This section presents descriptive statistics of the variables used in the empirical analysis. Table 3.1 shows the definition and descriptive statistics of the sample generated from the FoodAPS dataset. We separate the sample to rural and urban consumers. The dependent variable represents the quantity purchased on dairy, fresh meat, grain, fruits, and vegetables. The average quantity of food purchase is 2.17 pounds. The benchmark prices[2] are used to compare price levels across stores and distances. The price is \$3.2 per unit, on average. In order to minimize the price variation, its practical unit value is based on the quantity ratio representing household purchase decisions of specific food groups (dairy, fresh meat, grain, fruits, and vegetables). We used the price of food purchased to determine whether household food quantity changes differently when the travel distance of food stores increases.

Families' socioeconomic variables such as income and SNAP participation are included in the analysis to account for the outcomes of each group. The average income of a household is \$3,840. About 32% of the sample population participated in SNAP, and 85% had access to a car. The geographic region in which a household resides (South, Northeast, Midwest, and West) was a significant impact among individuals living in rural

[2] By following Deaton (1989); Gundersen et al. (2011); Gundersen et al. (2012) we derived the price from the quantity purchased and the total expenditure.

or urban areas. In addition, The North is home to 17% of the population, the Midwest has 24%, the West has 22%, and the South has 35%.

Considering that people generally select their grocery stores based on location, residents of urban areas are more likely to drive more than 2.5 miles to a food venue, on average, compared to the entire population of 3.8 miles, and those living in rural areas drive 7.5 miles. Regarding the density of healthy food stores, the number of food stores by type located within 1, 5, and 15 miles of the residence if urban; 10, 15, and 30 miles for rural areas, is selected based on USDA (2015) and previous researches (Breyer and Voss-Andreae, 2013; Ziliak and Gundersen, 2016; Rhone et al., 2019). Store types were categorized according to the North American Industry Classification System (NAICS) (USDA, 2019).

To achieve the non-collinearity condition, the number of supermarket and supercenter variables were highly correlated based on the Variation Inflation Factor test (VIF is greater than 0.8). To reduce the multicollinearity, we combined nearest supermarket and supercenter into a single variable LG; small and combination grocer into SG variable. LG consists of stores that carry a broad selection of fruits, vegetables, dairy products, fresh meats, and other groceries. SG variable denotes the grocery stores with a cash register or two, and where fruit and vegetables may or may not be available, as well as grocery stores offering milk, grains, drinks, and snacks.

In addition, we generated two variables about the direct marketing farmer/farmer market, specialized food store according to the size, and the nearest to the household. CSF and SPE variables consist of direct marketing farmer/farmers' markets and food stores specializing in the sale of fruits, vegetables, meat, poultry, or seafood products. We see high variation between the number of small grocery stores among rural and urban groups, and the variation is large in rural areas. Large grocery stores, direct marketing farmers, farmers' markets, and specialized grocery stores were statistically significant. However, the number of direct marketing farmers and farmers' markets were not large among groups.

3.5 Methods

We developed an econometric model consisting of three equations. The model is applied to our cross-sectional household-level data.[3]

$$PRICE = f_P(STORES, X) \tag{3.1.a}$$

$$TRAVEL = f_T(STORES, V, X) \tag{3.1.b}$$

$$FOOD = f_F\left(\underbrace{PRICE}_{\substack{\text{Price}\\\text{effect}}}, \underbrace{TRAVEL}_{\substack{\text{Distance}\\\text{effect}}}, STORES, V, X\right) \tag{3.1.c}$$

where $PRICE$ and $TRAVEL$ are two dependent (endogenous) variables. $PRICE$ denotes the price of food per unit; $TRAVEL$ represents the travel distance of primary food stores to households' homes; V indicated the transportation methods; $STORES$ is an exogenous variable represents the number of each food stores' type near a household home within a certain radius (1, 5 or 15 miles) for urban, (10, 15, and 30 miles) for rural. The variable $STORES$ consists of LG, SG, CSF and SPE which represent medium/large, supermarket, supercenter, combination, direct marketing farmer/farmer market, and specialized food stores in fruits, vegetables, meat, poultry, or seafood products; P, T, and F subscripts stand for the price, travel, and food functions respectively. X is a set of

[3] Subscript for household is suppressed in the equations.

explanatory variables that represent household socio-economic and demographic characteristics.

With Equations $3.1.a$ and $3.1.b$, we allow the number of food retailers per type to affect fresh food acquisition indirectly through an effect on the price and the travel distance to the food store. Equation $3.1.c$ is the equation where stores number is expected to affect food consumption directly and indirectly through prices and travel distance. Figure 3.1 illustrates the changes in demand and shifts in the quantity demanded of healthy foods. Based on our fundamental hypothesis, more food retailers increase the quantity of healthy food items through price-reduction effect (based on *a movement along* the demand curve) and distance-reduction effect due to switch to another closer food store (based on *an outward shift* of the demand curve). As a result, a lower price or a shorter travel distance to the new food store encourages customers to purchase food. We expect households in rural areas are more likely to live nearby fewer food stores. In addition, we allow the distance effect to interact with the ownership of vehicles and income to determine whether this effect is more substantial.

3.5.1 Three-Stage Least Squares Estimator

To determine the impacts of the price and distance reduction effects on the purchases of healthy food among rural and urban consumers, we utilize a system of equations to capture the variation resulting from the geographic location of the households. We started the analysis by using the Ordinary least squares (OLS) estimator and performed a Hausman's specification test (Hausman, 1978) and a Breusch–Pagan test (Breusch and Pagan, 1980) to evaluate whether the OLS yield consistent and efficient estimate (Greene, 2012). We test the null hypotheses that the OLS estimator is consistent and efficient for any existing correlation between the error terms and Equations $3.1.\,a$, $3.1.\,b$. and $3.1.\,c$ are independent equations. The null hypotheses were rejected and therefore, the OLS estimator is biased and inconsistent due to the estimated parameters and standard errors of the estimates are interpreted wrongly (Cameron and Trivedi, 2010).

In this study, we used instrumental variables estimation methods to handle endogeneity and heteroskedasticity. Three-Stage Least Squares (3SLS), developed by Zellner and Theil (1962), provides accurate estimates for conditionally homoscedastic residuals. The 3SLS is a combination of multivariate regression Two-Stage Least Squares (2SLS) and Seemingly Unrelated Regressions (SUR) for linear regression models with different equations, including different dependent variables and independent variables. The advantages of 3SLS are that the 3SLS utilizes all of the information available in Equations $3.1.\,a$, $3.1.\,b$., and $3.1.\,c$. . It provides accurate estimates for correlation between $PRICE$ and $TRAVEL$ variables and the error terms of the outcome variable $FOOD$. By following Greene (2012) and Gallant and Jorgenson (1979), the system of equations is postulated as follows:

$$y_i = X_i\ \beta_i + u_i = Z_i\ \beta_{1i} + Y\ \beta_{2i} + u_i \qquad (3.2.a)$$

where X_I is an $n \times k_I$ matrix of explanatory variables that can be separated as $X_i = [Z_i Y_i]$. Now Z_I is an $n \times k_{1i}$ matrix of variables that are assumed to be exogenous, and Y_I is an $n \times k_{2i}$ matrix of endogenous variables, where $k_{1i} + k_{2i} = k_i$. The $k_i - vector$ β_I of parameters can be separated as $[\beta_{1i} \vdots\ \beta_{2i}]$ to fit with the partitioning of X. The g endogenous variables y_1 to y_g are assumed to be jointly generated by g equations of the form $3.2.\mathrm{a}$. The number of exogenous variables that appear anywhere in the system is l, which implies that $k_{1i} \leq l$ for all i_2. To allow for the error terms assumption in Equraion $3.2.\mathrm{a}$ to have expected value $E(u_i) = 0$ and for the equations $E(u_i u_j) = \Sigma$, where Σ is a positive definite matrix, so that $E(uu') = \Sigma$. The matrix Σ is called the new covariance matrix. The error terms u_i arranged into an $n \times g$ matrix U and follows from (3) that $E(U_t{}'U_t) = \frac{1}{n}, E(U_t{}'\ U) = \Sigma$. If we combine equations $3.2.\mathrm{a}$, we obtain the classical SUR model that specifies the set of all exogenous variables as X and results in $\hat{z}_i = X(X'X)^{-1}X'z_I = Y_i X_1$ for each i as follow:

$$\hat{z}_i = \begin{bmatrix} Y_i\ X_1 & 0 & \dots & 0 \\ 0 & Y_i\ X_2 \dots & & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots Y_i\ X_3 & \end{bmatrix} \qquad (3.2.\mathrm{b})$$

where $\hat{Z}$ has the instrumented values for all regressors. Using the previous formulation, the estimation procedure of 3SLS consists of three stages. *Stage 1:* We regress the endogenous variables on the independent variables in Equation $3.1.\mathrm{a}$, $3.1.\mathrm{b}$, and $3.1.\mathrm{c}$. We predict the residuals (regression-predicted values) of the endogenous variables and then set them up as independent variables (e.g., instrument variables). Due to the instrumental variables, generalized least squares (GLS) (Aitken, 1935) estimator could be established for all the parameters of the system as follow:

$$\hat{\beta}_{3SLS} = \left\{\hat{Z}'(\Sigma^{-1} \otimes\ I\ )Z\right\}^{-1} \hat{Z}'(\Sigma^{-1} \otimes\ I\ )\ y \qquad (3.2.c)$$

where I is an identity matrix and $\otimes$ is Kronecker's product. We obtain the 2SLS value for each equation in the system, and the spread of errors can be eliminated. *Stage 2:* The second stage is to predict the residual. Next, the individual equations are estimated by 2SLS using their expected (updated) values in place of the endogenous variables from stage one (e.g., optimal instrument or weighting matrix). Then, we determine the residuals for each equation based on the 2SLS regressions. Due to the different parameter vectors of each equation and the correlation between error terms, we must combine the SUR method with the instrumental variable's 2SLS method. Thus, the residuals of the 2SLS model are used to estimate the cross-equation error variance-covariance matrix. To obtain a consistent estimator for Σ, we formed from the residuals of 2SLS estimates of each equation in the system. The residuals can be computed alternately and identically from the estimates formed by taking Σ to be an identity matrix. The full system of coefficients enables constraints to be used when the residuals are estimated. If U the matrix of residuals from these estimates, a consistent estimate of the matrix Σ is going to be $\hat{\Sigma} = \frac{U'U}{n}$ where n is the number of observations. *Stage 3:* We estimate the system equations using the Feasible Generalized Least Squares (FGLS) (cross-equation) variance-covariance matrix of the error terms. The variance-covariance matrix of the 3SLS estimator is similar to the 2SLS formula mixed with the GLS correction as

$$\hat{V}_{\beta} = \left\{ \bar{Z}' \left(\hat{\Sigma}^{-1} \otimes \ I \ \right) \bar{Z}' \right\}^{-1} \tag{3.2.d}$$

where $\bar{Z}' = diag[X\Pi_j, X_j]$ Estimation of this matrix requires bracketed inverse matrix in Equation$(3.2.c)$, and $\bar{Z}$ can be estimated with $\hat{Z}'$.

Box-Cox logarithm transformation was performed to normalize the distribution of the variables $FOOD$, $STORES$, and $TRAVEL$. The effect of log transformation is presented in Figure 3.3 (Box and Cox, 1964). Focusing on the key variables, we can rewrite Equations $3.1.a$, $3.1.b$, and $3.1.c$ in logarithm form as:

$$\ln(PRICE) = \beta_{S_{\pi},P} \ln(STORES) + \beta_{X,P}\, ln(X) \ + \beta_{D,P}\, D + u_P \tag{3.3.a}$$

$$\ln(TRAVEL) = \beta_{S_{\pi},T} \ln(STORES) + \beta_{V,T}\, V + \beta_{X,T}\, ln(X) \ + \beta_{D,T}\, D + u_T \tag{3.3.b}$$

$$\ln(FOOD) = \beta_0 + \beta_{P,F} \ln(PRICE) + \beta_{T,F} \ln(TRAVEL) + \beta_{D,F}\, D \tag{3.3.c}$$

$$+ \beta_{S_{\pi},F} \ln(STORES) + \beta_{V,F}\, V + \beta_{X,F}\, ln(X) + \ u_F$$

where u_P, u_T, and u_F are error terms for $PRICE$, $TRAVEL$, and $FOOD$ equations. X is a set of continuous variables that represent household socio-economic characteristics; D denotes dummy variables; V stands for car access; $\pi = LG, SG, CSF$, and SPE stands for the store type, where LG is large grocery stores (i.e., supermarket, supercenter), SG is a small grocery store (i.e., combination stores, convenience store). The direct(DE), indirect (IE), and total effects (TE) as presented in Figure 3.2, are calculated as:

$$TE_{S_{\pi},F} = IE_{S_{\pi},i} + DE_{S_{\pi},F} \tag{3.4.a}$$

where $i = P, T$ represent the price and travel-distance effects, respectively. The effect of store type is $DE_{S_{\pi},i}$; the indirect effect of store density on food consumption through price is $IE_{S_{\pi},P} = \ \beta_{S_{\pi},P} \ \cdot \ \beta_{P,F}$; and the indirect effect of store density on food consumption through travel distance is $IE_{S_{\pi},T} = \beta_{S_{\pi},T} \cdot \beta_{P,T}$; the coefficient of $\beta_{S_{\pi},P}$ denotes the effect of store type on price; $\beta_{S_{\pi},T}$ stands for the effects of store type on travel distance; $\beta_{S_{\pi},F}$ is the direct effect of store type on food consumption; $\beta_{P,F}$ and $\beta_{T,F}$ are each the effects of price and travel distance on food consumption. We rewrite the Equation $3.4.a$ to demonstrate the total effect of store number on food consumption through price and travel distance as:

$$TE_{S_{\pi},F} = [\beta_{S_{\pi},P} \cdot \beta_{P,F}] + [\beta_{S_{\pi},P} \cdot \beta_{T,F}] + \beta_{S_{\pi},F} \qquad (3.4.b)$$

3.6 Results and Discussion

This section presents summarized results for the 3SLS estimate. Multiple tests were performed to determine the goodness of fit. The non-normality tests after the 3SLS estimation for a single and overall system of equations were applied by using Jarque and Bera (1987) and Geary (1970) tests. Skewness and Kurtosis tests were computed using Srivastava LM Skewness Test (Nelson, 1998). Hence, the non-normality as indicated by the result of the tests could not be rejected, showing that the non-normality for the equations is not problematic. The heteroskedasticity was checked by using Breusch and Pagan (1980), Pagan and Hall (1983), and Engle (1982) tests in a linear regression model (LM). We accept the null hypothesis that the error variances are all equal. Hence, the homoscedasticity condition is met, and the 3SLS yields consistent and efficient estimation. The dependent variable is $FOOD$ denotes the quantity of dairy, meat, fruits, grains, and vegetables.

The 3SLS was estimated for small, large grocery stores, specialized food stores for meat, fish, fruits, and vegetables, and direct marketing/farmers' markets, which are classified respectively, into four types LG, SG CSF, and SPE. We allow the travel distance effect to interact with variable CAR. $SOUTH$, $INCOME$, and $SNAP$ were included in all models. The impacts of increasing the number of retail stores on food consumption are examined in seven models. Model 1 shows the results for the full sample regarding the number of food resources nearby the households and was estimated for a sample of 4,126 households. Models 2, 3, and 4 show the results for consumers in urban areas with radius buffers 1, 5, and 15 miles for a sample of 2,988 households, respectively. Models 5, 6, and 7 present the consumers in rural areas with radius buffers 10, 15, and 30 miles and were estimated on a sample of 1,138 households, respectively. Columns (1) and (3) demonstrate the cross-equation of $PRICE$ and $TRAVEL$, respectively. Indirect Effects (I.E.), the direct effects (D.E.), and Total Effects (T.E.) are listed as columns (2), (4), (5), and (6), respectively.

We focus on interpreting the results for Model 1 of the total sample, Model 2 for urban consumers living nearby stores within 1 mile, and Model 5 for rural consumers within 10 miles of food stores. We will compare the findings when increasing the buffer. The estimates of the coefficients of price, SNAP, and income were statistically significant at the 1% level, but the South region was statistically significant at the 5% level for consumers in rural areas within 30 miles of a store.

The price is significant at the 1% level for the total sample. If the price of healthy goods is halved, it is expected to increase the consumption by 4.6% for the consumer in urban areas, as demonstrated in Column (5) of Table 3.3, 3.4, and 3.5.

As expected, the impact of SNAP participation on consumers is significant at the 1% level. Consumers within 15 miles of healthy food stores in rural areas are expected to increase their consumption by approximately 1.10% compared to those who live within 30 miles (Table 3.7, Column (5)). The income is significant at the 1% level, and the impact on consumers living within 5 miles of food stores is large compared to rural. A 1% increase in income results in a 0.9% increase in the consumption of healthy foods.

Regarding the impact of travel distance on healthy food consumption, the sign of travel distance is only significant at a 10% level among consumers in rural areas within 10 miles radius. A 1% increase in travel distance for consumers in rural areas raises food consumption by 0.09%.

There was a significant indirect effect of car accessibility on consumers in urban areas within 15 miles of healthy food stores (Table 3.5 Column (6)). The indirect effect of car access is positive and significant at the 1% level for consumers in urban areas. Accordingly, for consumers living close to stores within 10 miles of rural areas, the coefficient from car access was significant at 1%. For consumers residing near stores within 15 and 30 miles, the coefficient is positive and significant at 10%.

In terms of the indirect impact of the density of healthy food stores on consumer consumption, there was a significant correlation between the number of large grocery stores and consumption of healthy food. For the total sample in Table 3.2 Column (6), a 1% increase in the number of large grocery stores results in an increase of 0.78% in the consumption of healthy foods. The difference between the rural and urban groups was significant. When compared to consumers in urban areas, food consumption in rural areas within 10 miles of stores will decline by 0.02%. The sign is significant at the 5% level. When the density of large grocery stores increases, the price will decrease, as expected.

According to the total sample, an increase of 1% in the number of small grocery stores is expected to indirectly increase the consumption of healthy foods by 0.15%. In urban areas within 10 miles of nearby stores, food consumption will increase by 0.13% relative to rural areas.

Similarly, direct marketing and farmers' market variable shows a significant sign for consumers in rural and urban areas. Within 10 miles of rural areas, a 1% increase in direct marketing is expected to raise healthy food consumption by 1.63%, as demonstrated in Table 3.6 Column (6). The expected consumption will increase by 1.22% when the direct marketing farmer is within 30 miles, as shown in Table 3.8 Column (6).

Furthermore, the indirect effect of the density of stores specializing in meat, poultry, seafood, fruit, and vegetables is significant at the 1% level in the total sample and 10% for consumers in urban areas. The indirect effect indicates that an increase of 1% in the number of specialty food shops near the home is expected to increase healthy food consumption by 0.73%, as presented in Table 3.2 Column (6). However, there is an expected decrease by 0.97% among consumers living within 10 miles of these stores. No impact has been observed among consumers in rural areas.

In summary, the findings show that when the density of large groceries increases, the indirect effect of healthy food consumption is expected to increase more than small grocery stores. These findings align with Taylor and Villas-Boas (2016), who found that large grocery stores tend to be preferred by households due to their accessibility.

The study has shown that food prices are expected to decrease in an increasing number of stores. As the miles buffer increased, rural areas experienced a greater decrease than urban areas. This result is in line with Abratt and Goodey (1990); Powell et al. (2007) findings, which found that when the number of grocery stores increased, the market would become more competitive, as was reported in earlier studies by (Powell et al., 2013). This result is consistent with findings from Gutman (2002); Heinrich and Betts (2003); Lee (2004).

Contrary to expectations, the travel distance for a consumer in urban areas was not statistically significant. The finding shows that as travel distance increases, rural residents are likely to purchase more food. Gustat et al. (2015) and Block and Kouba (2006) also observed that consumer preference towards food changes when travel distance is reduced.

The results show that the impact of car access was significant among consumers in rural areas. However, consumers in urban areas show more response in increasing the healthy food consumption when a car is available. The impact is larger among consumers in urban areas within 15 miles. This outcome is consistent with Ver Ploeg (2010), who demonstrated that families living far from supermarkets have difficulty meeting their food needs. He noted that access to fresh food is limited in rural areas and small towns with poor transportation infrastructure.

The findings indicated that specialized food stores show a negative correlation with healthy food consumption. This finding contradicts previous studies by Yan et al. (2015), who found that Specialty food stores tend to be more widespread in urban areas. The South region was statistically significant at the 5% level for consumers in rural areas within 30 miles of a store. These results reflect Dutko (2012), who reported that the South region has the highest level of food deserts.

3.7 Conclusions and Policy implication

We investigated how the density of healthy food resources impacts the purchases of healthy foods by utilizing the Three-Stages Least Squares (3SLS) method. Based on distance traveled and price, the number of healthy food sources, i.e., supermarkets, specialized food stores, direct marketing, and farmers' markets, was used to quantify the consumption of healthful foods. We examined the impact of retail store density on the consumption pattern of US residents using the USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). We estimated 4,126 households in both rural and urban areas. The analysis was based on the hypothesis that the density of healthy food stores will reduce the number of trips to the store, and it affect healthy food consumption.

Results show that a price increase reduces healthy food consumption by US households. An increase in travel distance will increase the amount of food purchased for consumers living in rural areas. In urban areas, however, travel distances among US consumers are negative and statistically insignificant. The density of large grocery stores and specialized food stores indirectly increases the consumption of healthy foods. The availability of cars has a positive impact on US consumers living in urban areas. SNAP participants in rural areas within 30 miles consume fewer healthy food products than urban populations. Results from the study indicate that specialized food stores have a positive impact on the total sample. Direct marketing and farmers' markets significantly affect US consumers in rural areas. Farmers' markets do not have a significant impact on urban individuals. Car access showed a significant impact on consumers in an urban area with far food stores. Consumers in rural areas differ from those in urban areas in terms of SNAP benefits and food consumption.

To conclude, the uniqueness of this study is allowing stores to affect food consumption through a direct as well as an indirect effect. The stores' density improves consumers' perception of food availability and accessibility. The increased number of large grocery stores and direct marketing/farmers' markets increase food consumption. The impact on consumption is getting smaller as the distance is increasing. Also, income, SNAP participation, car access, and the South region were statistically significant in influencing healthy food consumption.

The objective of these interventions is to increase the availability of nutritious food to give consumers more choices. Therefore, it would seem that ownership of a car positively impacts food purchases for consumers in rural areas (10-30 miles) and urban areas (15 miles). According to this result, residents without transportation need to access public transportation regularly to purchase healthy foods. Therefore, policies intended to boost the purchase and availability of healthy foods for consumers should consider providing regular transportation, especially for those in rural areas who do not own their cars.

A limitation of this study is that the use of cross-section data allows us only to determine the impact of the density of stores on household consumption of healthy food. This limits our study from determining the impact of a new store opening on household consumption over time. Panel data and panel data modeling techniques are required to assess the impact of new store openings on household consumption choices.

In light of this, future studies can benefit from using panel data to determine the impact of new store openings on household consumption of healthy foods. Deliveries of groceries may be an alternative worth considering for those without reliable transportation or who need access to other food sources

3.8 Tables and Figures for Chapter 3

Table 3.1 Definitions and descriptive statistics of the variables

Variables	Definitions	Full Sample (N=4,126)	Urban (N=2,988)	Rural (N=1,138)
FOOD	The dependent variable denotes the quantity of dairy, meat, fruits, grains, and vegetables (lb.)	2.17	2.16	2.19
TRAVEL***	Driving distance, in miles, between residence and the source of healthy grocery stores or (miles)	3.87	2.51	7.47
PRICE*	Price per unit ($)	3.20	3.19	3.26
INCOME	Household average income in 1000 ($)	3.84	3.87	3.74
SNAP	Respondent is SNAP participated= 1, 0 otherwise	0.32	0.32	0.30
VEHICLE***	Respondent has a car access=1, 0 otherwise	0.85	0.83	0.91
LG***	Number of nearest supermarket/supercenters	2.26	2.32	2.10
SM***	Number small and combination grocery stores	6.64	3.17	15.78
CSF*	Number of nearest direct marketing farmers, or farmers' markets.	0.03	0.03	0.04
SPE***	Number of nearest food stores specializing in the sale of fruits, vegetables, meat, poultry, or seafood products.	0.03	0.03	0.05
FARM_MARKET***	Respondent ever shop at direct marketing farmer, or farmer market in season =1, 0 otherwise	0.48	0.43	0.59

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 3.1 Continued

Variables	Definitions	Full Sample (N=4,126)	Urban (N=2,988)	Rural (N=1,138)
SOUTH***	South region = 1, 0 otherwise	0.35	0.32	0.47
NORTH***	Northern region = 1, 0 otherwise	0.17	0.18	0.16
MIDWEST***	Midwest region = 1, 0 otherwise	0.24	0.24	0.26
WEST***	West region = 1, 0 otherwise	0.22	0.27	0.12
STO_CO_1m***	Number of combination grocery/other stores within 1 mile	2.28	2.99	0.41
STO_MLG_1m***	Number of medium & large grocery stores within 1 mile	1.46	1.99	0.09
STO_SM_1m***	Number of supermarkets within 1 mile	0.92	1.22	0.12
STO_SS_1m***	Number of superstores within 1 mile	0.81	1.08	0.10
STO _CO_5m***	Number of combination grocery/other stores within 5 miles	33.50	44.37	5.15
STO _MLG_5m***	Number of medium & large grocery stores within 5 miles	19.37	26.55	0.55
STO _SM_5m***	Number of supermarkets within 5 miles	13.25	17.68	1.61
STO _SS_5m***	Number of superstores within 5 miles	12.65	16.98	1.28
STO_CO_10m***	Number of combination grocery/other stores within 10 miles	100.80	133.34	15.53

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 3.1 Continued

Variables	Definitions	Full Sample (N=4,126)	Urban (N=2,988)	Rural (N=1,138)
STO_MLG_10m***	Number of medium & large grocery stores within 10 miles	72.83	99.91	1.77
STO_SM_10m***	Number of supermarkets within 10 miles	40.57	53.98	5.39
STO _SS_10m***	Number of superstores within 10 miles	37.90	50.80	3.99
STO _CO_15m***	Number of combination grocery/other stores within 15 mi	173.53	225.71	36.53
STO _MLG_15m***	Number of medium & large grocery stores within 15 mi	116.59	158.85	5.65
STO _SM_15m***	Number of supermarkets within 15 mi	68.74	90.19	12.42
STO _SS_15m***	Number of superstores within 15 mi	65.08	86.10	9.92
STO _MLG_30m***	Number of medium & large grocery stores within 30 mi	201.90	266.31	32.95
STO _SM_30m***	Number of supermarkets within 30 mi	146.00	186.26	43.26
STO _SS_30m***	Number of superstores within 30 mi	144.60	184.86	39.04

Student's t-tests are used to evaluate whether two groups have a significant difference between the means ***, **, * denote significant difference at 1%, 5%, and 10% levels, respectively.

Table 3.2 Results of stores density near households for the total sample (N=4,126)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG GROCERY	-0.732***	0.289***	0.105*	0.008	0.484*	0.781***
	(0.020)		(0.045)		(0.232)	
SG GROCERY	-0.325***	0.129***	0.146***	0.011	0.007	0.147***
	(0.009)		(0.008)		(0.010)	
CSF	-0.042	0.017	0.064**	0.005	-0.097	-0.075
	(0.063)		(0.023)		(0.109)	
SPE	-0.237*	0.094	-0.064	-0.005	0.637***	0.726***
	(0.103)		(0.077)		(0.190)	
CAR			0.744***	0.056	0.319	0.375
			(0.042)		(0.331)	
TRAVEL					0.076	
					(0.040)	
PRICE					-0.396***	
					(0.076)	
INCOME					0.954***	
					(0.151)	
SNAP					0.462***	
					(0.135)	
SOUTH					0.086	
					(0.285)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.3 Results of stores density in urban areas within a 1-mile radius (N=2,988)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_1_mile	0.009	-0.004	-0.083***	0.001	0.029***	0.025***
	(0.008)		(0.001)		(0.003)	
SG_1_mile	-0.202***	0.094***	0.035***	0	0.039*	0.132***
	(0.009)		(0.006)		(0.019)	
CSF_1_mile	-0.252	0.116	-0.071	0.001	-0.06	0.057
	(0.486)		(0.493)		(1.337)	
SPE_1_mile	0.367*	-0.17*	-0.27**	0.003	-0.808*	-0.974*
	(0.166)		(0.086)		(0.372)	
CAR			0.698***	-0.009	0.422	0.413
			(0.029)		(0.351)	
TRAVEL					-0.012	
					(0.055)	
PRICE					-0.462***	
					(0.078)	
INCOME					0.961***	
					(0.262)	
SNAP					0.577***	
					(0.145)	
SOUTH					0.073	
					(0.250)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.4 Results of stores density in urban areas within a 5-mile radius (N=2,988)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_5_miles	-0.004***	0.0017**	-0.0005	0	0.016	0.018*
	(0.001)		(0.002)		(0.008)	
SG_5_miles	-0.006***	0.0027***	-0.0003	0	-0.012*	-0.0091
	(0.001)		(0.002)		(0.006)	
CSF_5_miles	-0.31***	0.143***	0.0036	-0.0002	-0.425***	-0.282***
	(0.045)		(0.100)		(0.078)	
SPE_5_miles	-0.422	0.195	-0.147	0.008	0.933	1.136
	(0.542)		(0.137)		(1.197)	
CAR			0.645***	-0.034	0.339	0.305
			(0.032)		(0.372)	
TRAVEL					-0.052	
					(0.043)	
PRICE					-0.462***	
					(0.057)	
INCOME					0.986*	
					(0.391)	
SNAP					0.556***	
					(0.122)	
SOUTH					0.178	
					(0.191)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.5 Results of stores density in urban areas with a 15-mile radius (N=2,988)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_15_miles	0.0012*** (0.000)	-0.0005*	-0.001*** (0.000)	0	0.006* (0.002)	0.0054*
SG_15_miles	-0.0027*** (0.000)	0.0013***	0.0007*** (0.000)	0	-0.0042* (0.002)	-0.003
CSF_15_miles	-0.314*** (0.067)	0.147***	0.018 (0.048)	-0.0004	-0.337*** (0.057)	-0.191***
SPE_15_miles	-0.564*** (0.138)	0.263*	-0.066 (0.141)	0.0016	1.048 (1.371)	1.313
CAR			0.634*** (0.013)	-0.015	0.403* (0.172)	0.388***
TRAVEL					-0.024 (0.104)	
PRICE					-0.467*** (0.090)	
INCOME					0.995*** (0.182)	
SNAP					0.56*** (0.148)	
SOUTH					0.317 (0.253)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.6 Results of stores density in rural areas within a 10-mile radius (N=1,138)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_10_miles	0.0055	-0.0023	-0.0024	-0.0002	-0.0224*	-0.0249**
	(0.006)		(0.007)		(0.011)	
SG_10_miles	-0.035***	0.0145**	0.0232***	0.0021	0.0038	0.0204***
	(0.007)		(0.006)		(0.006)	
CSF_10_miles	-0.122	0.0504	0.2835	0.026	1.5494**	1.6259**
	(0.166)		(0.256)		(0.481)	
SPE_10_miles	-0.8681**	0.3589**	0.4617***	0.0424	-0.3832	0.0181
	(0.306)		(0.115)		(0.921)	
CAR			0.8609***	0.079*	0.3675*	0.4465***
			(0.014)		(0.171)	
TRAVEL					0.0918*	
					(0.045)	
PRICE					-0.4135***	
					(0.101)	
INCOME					0.9829***	
					(0.159)	
SNAP					0.5182***	
					(0.099)	
SOUTH					0.0901	
					(0.294)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.7 Results of stores density in rural areas within a 15-mile radius (N=1,138)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_15_miles	0.0083**	-0.0034***	-0.0063***	-0.0006*	-0.0005	-0.0045
	(0.003)		(0.002)		(0.025)	
SG_15_miles	-0.019***	0.0077***	0.014***	0.0013*	-0.004	0.0053
	(0.001)		(0.001)		(0.013)	
CSF_15_miles	-0.096	0.04	0.207**	0.019	1.235	1.294
	(0.283)		(0.067)		(0.918)	
SPE_15_miles	-0.886***	0.367**	0.473***	0.044	-0.1	0.311
	(0.176)		(0.130)		(0.332)	
CAR			0.85***	0.079	0.367	0.447*
			(0.022)		(0.215)	
TRAVEL					0.093	
					(0.050)	
PRICE					-0.414***	
					(0.067)	
INCOME					0.983***	
					(0.239)	
SNAP					0.512***	
					(0.119)	
SOUTH					0.105	
					(0.114)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

Table 3.8 Results of stores density in rural areas within a 30-mile radius (N=1,138)

Variable	Equation 3.3. a		Equation 3.3. b		Equation. 3.3. c	T.E. of Stores
	β	I.E.	β	I.E.	D.E.	
	Price		Travel		Food	
	(1)	(2)	(3)	(4)	(5)	(6)
LG_30_miles	0.006***	-0.002***	-0.0064***	-0.0006	-0.002	-0.0049
	(0.001)		(0.001)		(0.010)	
SG_30_miles	-0.007***	0.003***	0.0069***	0.0007	-0.0001	0.0037
	(0.001)		(0.000)		(0.006)	
CSF_30_miles	0.042	-0.017	0.087	0.008	1.235***	1.226**
	(0.077)		(0.145)		(0.369)	
SPE_30_miles	-0.833*	0.345*	0.449*	0.042	-0.118	0.269
	(0.383)		(0.198)		(0.415)	
CAR			0.822***	0.078	0.368**	0.446*
			(0.028)		(0.119)	
TRAVEL					0.094	
					(0.122)	
PRICE					-0.414***	
					(0.024)	
INCOME					0.985***	
					(0.054)	
SNAP					0.517***	
					(0.086)	
SOUTH					0.067**	
					(0.021)	

Note: p <0.10, *p <0.05, **p < 0.01, ***p < 0.001. Standard Error in parentheses. I.E., D.E., and T.E. represent indirect effects, direct effects, and the total effects. All the variables are in log form except for CAR, $SNAP$, and $SOUTH$.

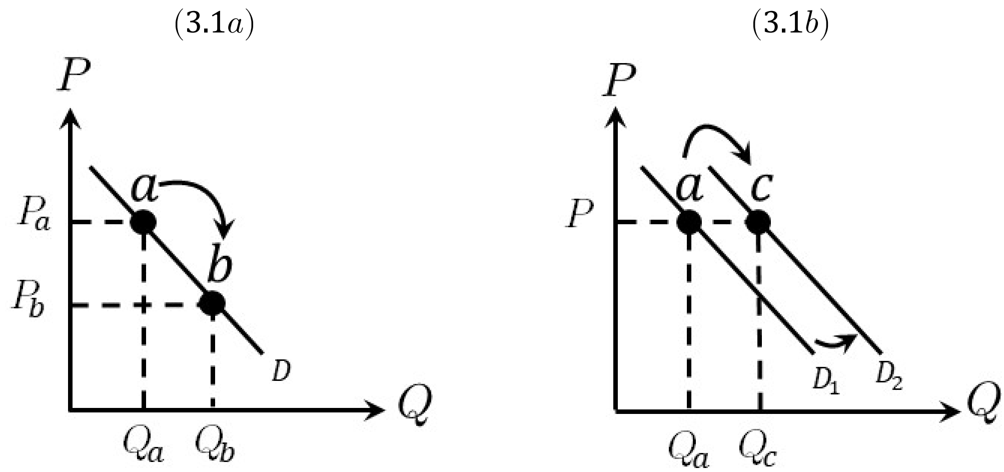

Figure 3.1 The impact of the number of food retailers on healthy food demand.

Note: Figure 3.1a shows the impact of the price-reduction effect caused by increasing the number of food retailers. The movement along the existing demand curve (D) from (point a to b). With a shift in supply, the price of healthy foods will decrease from P_a to P_b, and the quantity demanded will increase from Q_a to Q_b. Figure 3.1b illustrates the impact of the travel distance reduction effect. The initial demand curve D_1 shifts to become D_2 caused by the density of food retailer. The price of healthy food P is constant while the quantity demanded will increase from Q_a to Q_c (point a to c).

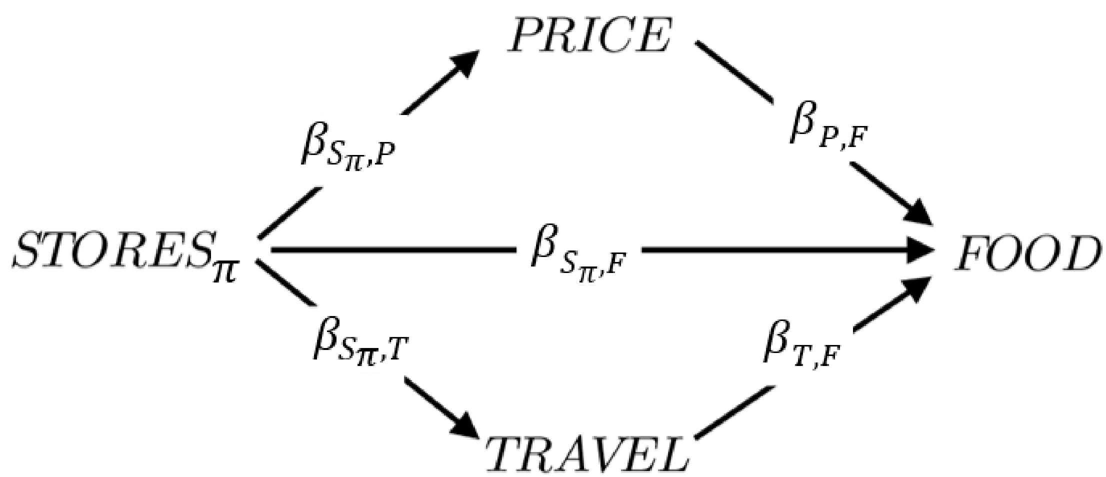

Figure 3.2 The indirect effect of store density on food consumption.

Note: $\beta_{S_{\pi},P}$ denotes the effect of $STORES$ on $PRICE$; $\beta_{S_{\pi},T}$ stands for the effect of $STORES$ on $TRAVEL$; $\beta_{S_{\pi},F}$ is the direct effect of $STORES$ on $FOOD$; $\beta_{P,F}$ denotes the effect of $PRICE$ on $FOOD$; $\beta_{T,F}$ stands for the effect of $TRAVEL$ on $FOOD$.

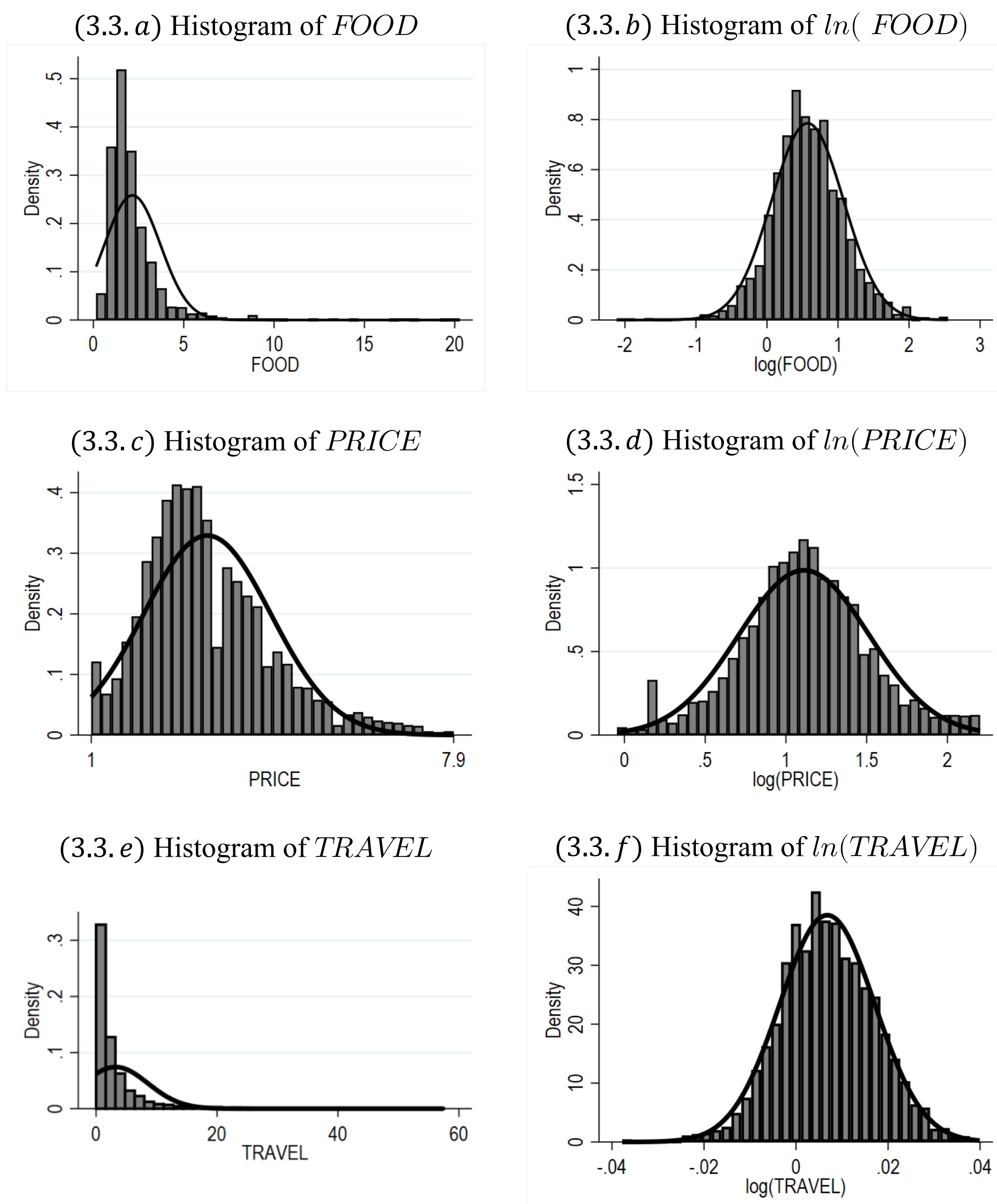

Figure 3.3 Histogram of the variables with the effect of log transformation

Chapter 4 General Summary and Conclusions

4.1 Overview

This dissertation focused on the impact of the food environment and store density on household food consumption in the US. Accordingly, the study applied two main analytical frameworks to shed light on consumer food choice. First, the Structural Equation Modeling (SEM) was used to measure the effect of internal and external food environments on household healthy food consumption in the US. Second, a Three-Stage Least Square Equation (3SLS) modeling is used to determine the impact of store density on household consumption in the US.

Contributions of this dissertation included estimating the effect of internal and external food environments on household consumption behavior of healthy food (Chapter 2). This study also contributes to the literature on food security in developed countries, focusing on dietary preferences and food consumption habits (Chapters 2 and 3). Further, this study adds to the literature on the food environment and its relationship to healthful food choices in the US (Chapters 2 and 3).

This dissertation discussed the complicated nature of the healthy food environment in the US in two essays. The first essay looked at how in-home healthy consumption and out-of-home factors are expected to influence obesity and food insecurity. Specifically, we looked at internal and external food environments to explain the effect between food environments and healthy eating habits.

In the second essay, we quantified the change of food availability through the density of healthy food sources, i.e., supercenter, supermarket, specialized food store, direct marketing market, and farmers' markets, on the consumption of healthy products. Such expansion in the density of healthy food sources increases the likelihood of improving overall nutrition among urban and rural consumers.

4.2 Summary of Key Findings

Key results from Chapter 2 are as follows. The importance of the internal food environment is relative to the external food environment. Except for food availability, all internal and external food environment characteristics are statistically significant in explaining the probability of adopting healthy eating habits. Price is significant and identical for all models. Store and product properties positively impact healthy eating habits, and consumers consider high-quality stores, making households more likely to eat healthily.

Participating in nutrition education events, the desire to learn about healthy food, or searching the internet for nutrition information increases healthy food knowledge and awareness. Moreover, the impact of knowledge of healthy foods was higher among food-secure households compared to food-insecure individuals. Accessibility to healthy food was significant, and the size was larger compared to food-insecure households. Compared to food insecure individuals, the impact of convenience, i.e., taste, incentive, i.e., timing, and the desire to consume healthy foods, is more significant among food-secure households than food-insecure individuals. Participation in the SNAP program does not guarantee a higher HEI score. Indicators such as healthy food items are expensive, and households with significant non-food expenses may decrease HEI.

Key results from Chapter 3 are that the density of healthy food stores indirectly affects healthy food consumption in rural and urban areas. The density of healthy food sources improves consumers' perception of food availability and accessibility. Consumers' healthy food choices are indirectly affected by lowering prices and reducing travel distances. The number of supermarkets or supercenter stores increases consumers' food purchases compared to small grocery stores. Specialized food stores in meat, poultry, seafood, and vegetables and fruits have a significant impact on driving consumers to choose healthier food options. Urban and rural consumers were separated in the analysis to see the overall impact of density of healthy food sources on food consumption through travel distance and price.

Also, income, SNAP participation, and the South region were statistically significant in influencing healthy food consumption. These effects were most apparent within 1-5 miles of urban areas and a range of 10-15 miles of rural areas. Such result dissipates when the distance to the store is increased to 25 miles

4.3 Policy implication

This study provides insights and guidance about the potential benefits of improving food environments in the economy and individuals. The findings serve as an evidence base for recommending policies to sustain nutritional diets and enhance regional households' welfare and individuals' healthful choice and wellbeing.

Based on these results, guidance and insights can be provided. This study has shown important findings. A healthy eating index and a healthy food environment index were used to evaluate the quality of food environments among food-secure and food-insecure households. To emphasize the connection between the pillars of food security and healthy eating environments, a comprehensive assessment of the impact of food availability, affordability, utilization, and accessibility in US families was studied.

As part of further investigation of the availability domain, US households reduce healthy food consumption when prices rise. Consumption of food in rural areas will increase as travel distance increases. In urban areas, however, travel distances are statistically insignificant among US consumers. Large grocery stores and specialized food stores indirectly increase the consumption of healthy foods. US urban consumers benefit from the availability of cars.

Rural SNAP participants consume fewer healthy food products than urban ones. Specialized food stores positively impact the total sample, while direct marketing and farmers' markets have a significant impact on US consumers in rural areas. Farmers' markets have little impact on urban residents. In an urban area with far food stores, car access had a noticeable impact on consumers. Consumers in rural areas differ from those in urban areas in terms of SNAP benefits and food consumption.

This study has several practical implications. In order to encourage SNAP recipients to adopt a healthy lifestyle and eat healthy foods, it is recommended that policymakers continue and elevate the efforts of Supplemental Nutrition Assistance Program Education (SNAP-Ed). In order to provide nutrition education for SNAP beneficiaries, states must maintain their efforts. A policymaker should also consider adjusting the relative price of healthy foods for consumers with limited access to healthy foods. An important practical implication is that a SNAP-authorized retailer requires to sell healthy foods for home consumption and preparation.

This intervention increases the availability of nutritious food and gives consumers more options. Owning a car may positively impact food purchasing for consumers in rural areas (10-30 miles) and urban areas (15 miles). Residents without access to transportation must use public transportation regularly in order to purchase healthy foods. To boost the purchase and availability of healthy foods, policies should consider providing better public transportation, especially for those in rural areas without car access. The policymakers could also consider more ways for direct marketing to consumers.

Despite some key results and comprehensive efforts, this study has limitations and caveats that need to be noted. The study was limited by data regarding the domain “marketing and regulation” within the external food environment dimension. It contains information regarding promotional information, branding, advertising, sponsorship, and policies. Another limitation of this study is that the use of cross-section data allows us only to determine the impact of the density of stores on household consumption of healthy food.

This limits our study from determining the impact of a new store opening on household consumption over time. The impact of a new shop opening on household consumption must be assessed using panel data and panel data modeling techniques.

For future research, the SNAP-Ed program could be examined in greater detail to provide interesting findings related to obesity and food insecurity. A future qualitative study, e.g., Difference-in-Difference estimation, could examine the impact of SNAP-Ed program implementation on food security and obesity rates in a state government. In addition, a future study could benefit from using panel data to assess the impact of new store openings on household consumption of healthy foods. Grocery delivery might be an option worth considering for future studies for those lacking reliable transportation.

Appendices

Appendix 4.A

List of food items for food-at-home acquisitions

Item	Code*	Item	Code**
Whole-grain breads, rolls, etc.	10101	Milk	10
Whole-grain rice and pasta	10102	Dairy drinks substitutes	14
Whole-grain breakfast cereals	10103	Cheese	16
Whole-grain flour, bread mixes, frozen dough	10104	Yogurt	18
Fresh starchy vegetables	20101	Meats	20
Frozen starchy vegetables	20102	Poultry	22
Canned starchy vegetables	20103	Seafood	24
Fresh tomatoes	20201	Eggs	25
Canned tomatoes	20203	Plant-based protein foods	28
Fresh dark green vegetables	20301	Rice, pasta, cooked grains	40
Frozen dark green vegetables	20302	Breads, rolls, tortillas	42
Canned dark green vegetables	20303	Fruits	60
Fresh red and orange vegetables	20401	Vegetables	64
Frozen red and orange vegetables	20402	White potatoes	68
Canned red and orange vegetables	20403	100% juice	70
Fresh beans, lentils, legumes	20501	Coffee and tea	73
Canned beans, lentils, legumes	20503	Plain water	77
Fresh other/mixed vegetables	20601	Baby foods	90
Frozen other/mixed vegetables	20602	Baby juice and water	92
Canned other/mixed vegetables	20603	Infant formulas	94
Fresh whole fruit	30101	Protein and nutritional powders	98
Frozen whole fruit	30102	Chicken, turkey, game birds	50202
Canned whole fruit	30103	Chicken, turkey, game birds	50203
Dried whole fruit	30104	Fresh fish and seafood	50301
100% fruit and vegetable juices	30201	Frozen fish and seafood	50302
Whole milk	40101	Canned fish and seafood	50303
Whole milk yogurt	40103	Raw nuts and seeds	50401
Low-fat or skim milk	40201	Processed nuts/seeds and spreads	50402
Low-fat or skim milk yogurt	40203	Unsweetened coffee and tea	70302
All unprocessed cheese	40301	Water	70306
Fresh beef, pork, veal, lamb, game	50101	Vitamins and meal supplements	70601
Frozen beef, pork, veal, lamb, game	50102	Baby food	70701
Fresh chicken, turkey, game birds	50201	Infant formula	70801

* Source of food pattern was from USDA's National Nutrient Database for Standard Reference Legacy Release (SR), ** USDA's Food and Nutrient Database for Dietary Studies (FNDDS)

Appendix 4.B

Healthy Eating Index components and scoring standards

Component	Maximum Score[a] %	Standard for Maximum Score[b] equivalent per 1,000 kcal
HEI	100	Total HEI Score from 0 to 100
Total Fruit	5	≥ 0.8 cup
Whole Fruit	5	≥ 0.4 cup
Total Vegetables	5	≥ 1.1 cup
Greens and Beans	5	≥ 0.2 cup
Whole Grains	10	≥ 1.5 oz
Dairy	10	≥1.3 cup
Total Protein Foods	5	≥ 2.5 oz
Seafood and Plant Proteins	5	≥ 0.8 oz
Fatty Acids	10	≥ 2.5 oz
Refined Grains	10	≤ 1.8 oz
Sodium	10	≤ 1.1 gram.
Empty Calories	20	≤ 19% of energy

[a] Note: Standard for minimum score is 0 except for Refined Grains (≥4.3), Sodium (≥2.0), and Empty Calories (≥50% of energy).

[b] Sources (Guenther et al., 2013; Krebs-Smith et al., 2018)

Appendix 4.C

The SEM Formulation and Methods to Construct the Latent Variable Model

As Bollen (1987) explained in detail, we rewrite the structural equation model in Equation 2.1. a and 2.1. b as:

$$\eta = \mathbf{B}\eta + \boldsymbol{\Gamma}\xi + \zeta \tag{C. 1}$$

$$\mathbf{x} = \boldsymbol{\Lambda}_x \xi + \delta \tag{C. 2}$$

where $E(\delta) = 0$ and uncorrelated with ζ. In the internal food environment η_1, the unidentified measurement model for the four latent exogenous variables $\xi = (\xi_1, \xi_2, \xi_3, \xi_4)'$ with their indicators $(x_1, \dots, x_{14})'$ shows the form of the factor loadings matrix $\boldsymbol{\Lambda_1}$, considering the matrix representation of Figure 2.7, as:

$$\boldsymbol{\Lambda_1} = \begin{bmatrix} \lambda_{1,1} & 0 & 0 & 0 \\ \lambda_{2,1} & 0 & 0 & 0 \\ \lambda_{3,1} & 0 & 0 & 0 \\ \lambda_{4,1} & 0 & 0 & 0 \\ 0 & \lambda_{5,2} & 0 & 0 \\ 0 & \lambda_{6,2} & 0 & 0 \\ 0 & \lambda_{7,2} & 0 & 0 \\ 0 & 0 & \lambda_{8,3} & 0 \\ 0 & 0 & \lambda_{9,3} & 0 \\ 0 & 0 & \lambda_{10,3} & 0 \\ 0 & 0 & \lambda_{11,3} & 0 \\ 0 & 0 & 0 & \lambda_{12,4} \\ 0 & 0 & 0 & \lambda_{13,4} \\ 0 & 0 & 0 & \lambda_{14,4} \end{bmatrix} \tag{C. 3}$$

where the arrow goes from column j to row i

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{bmatrix} = \begin{bmatrix} \tau_{x_1} \\ \tau_{x_2} \\ \tau_{x_3} \\ \tau_{x_4} \\ \tau_{x_5} \\ \tau_{x_6} \\ \tau_{x_7} \\ \tau_{x_8} \\ \tau_{x_9} \\ \tau_{x_{10}} \\ \tau_{x_{11}} \\ \tau_{x_{12}} \\ \tau_{x_{13}} \\ \tau_{x_{14}} \end{bmatrix} + \begin{bmatrix} \lambda_{1,1} & 0 & 0 & 0 \\ \lambda_{2,1} & 0 & 0 & 0 \\ \lambda_{3,1} & 0 & 0 & 0 \\ \lambda_{4,1} & 0 & 0 & 0 \\ 0 & \lambda_{5,2} & 0 & 0 \\ 0 & \lambda_{6,2} & 0 & 0 \\ 0 & \lambda_{7,2} & 0 & 0 \\ 0 & 0 & \lambda_{8,3} & 0 \\ 0 & 0 & \lambda_{9,3} & 0 \\ 0 & 0 & \lambda_{10,3} & 0 \\ 0 & 0 & \lambda_{11,3} & 0 \\ 0 & 0 & 0 & \lambda_{12,4} \\ 0 & 0 & 0 & \lambda_{13,4} \\ 0 & 0 & 0 & \lambda_{14,4} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \\ \delta_{12} \\ \delta_{13} \\ \delta_{14} \end{bmatrix} \tag{C. 4}$$

Now, we can specify the structural model of the internal food environment η_1 as:

$$\eta_1 = \alpha_1 + \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \gamma_{1,3} & \gamma_{1,4} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \beta_{1,2}\,\eta_2 + \zeta_1 \tag{C. 5}$$

where $\mathbf{\Gamma}$ matrix denotes the relationship between exogenous and endogenous variables, $\gamma_{1,1}$ $\gamma_{1,2}$ $\gamma_{1,3}$ and $\gamma_{1,4}$ are the regression path of the latent endogenous variable η_1 with the exogenous variable $\xi_1, \xi_2, \xi_3,$ and ξ_4. $\beta_{1,2}$ represents the coefficient of the external food environment η_2. For the external food environment η_2 , the matrix $\boldsymbol{\Lambda_2}$ for is as:

$$\boldsymbol{\Lambda_2} = \begin{bmatrix} \lambda_{15,5} & 0 \\ \lambda_{16,5} & 0 \\ \lambda_{17,5} & 0 \\ \lambda_{18,5} & 0 \\ 0 & \lambda_{19,6} \\ 0 & \lambda_{20,6} \\ 0 & \lambda_{21,6} \\ 0 & \lambda_{22,6} \end{bmatrix} \tag{C. 6}$$

$$\begin{bmatrix} x_{15} \\ x_{16} \\ x_{17} \\ x_{18} \\ x_{19} \\ x_{20} \\ x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} \tau_{x_{15}} \\ \tau_{x_{16}} \\ \tau_{x_{17}} \\ \tau_{x_{18}} \\ \tau_{x_{19}} \\ \tau_{x_{20}} \\ \tau_{x_{21}} \\ \tau_{x_{22}} \end{bmatrix} + \begin{bmatrix} \lambda_{15,5} & 0 \\ \lambda_{16,5} & 0 \\ \lambda_{17,5} & 0 \\ \lambda_{18,5} & 0 \\ 0 & \lambda_{19,6} \\ 0 & \lambda_{20,6} \\ 0 & \lambda_{21,6} \\ 0 & \lambda_{22,6} \end{bmatrix} \begin{bmatrix} \xi_5 \\ \xi_6 \end{bmatrix} + \begin{bmatrix} \delta_{15} \\ \delta_{16} \\ \delta_{17} \\ \delta_{18} \\ \delta_{19} \\ \delta_{20} \\ \delta_{21} \\ \delta_{22} \end{bmatrix} \tag{C. 7}$$

Then, the structural model of the external food environment η_2 is:

$$\eta_2 = \alpha_2 + \begin{bmatrix} \gamma_{2,5} & \gamma_{2,6} \end{bmatrix} \begin{bmatrix} \xi_5 \\ \xi_6 \end{bmatrix} + \begin{bmatrix} \gamma_{2,23} & \gamma_{2,24} \end{bmatrix} \begin{bmatrix} x_{23} \\ x_{24} \end{bmatrix} + \zeta_2 \tag{C. 8}$$

where $\gamma_{2,5}$ $\gamma_{2,6}$ $\gamma_{2,23}$ and $\gamma_{2,24}$ are the regression path of the latent endogenous variable η_2 with the exogenous variable ξ_5, ξ_6 and observed variables x_{23}, and x_{24} for the price and regions. To specify the interrelationship of the first endogenous variable to the other, the structural matrices corresponding to the model in Figure 2.7,

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \gamma_{1,3} & \gamma_{1,4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_{2,5} & \gamma_{2,6} & \gamma_{2,23} & \gamma_{2,24} \end{bmatrix} \tag{C. 9}$$

$$\mathbf{B} = \begin{bmatrix} 0 & \beta_{1,2} \\ 0 & 0 \end{bmatrix} \tag{C. 10}$$

We substitute $(B.\,9)$ and $(B.\,10)$ in the observed effect for ξ on η and observe $\mathbf{\Gamma}$.

Appendix 4.D

Model Validation

It is essential to perform global fit measures in the SEM to provides information about the model fit. We aim to obtain the best fit for the sample data and the best fit to describe the population as a whole. The coefficient of determination for the equation-level goodness is processed as:

$$R_i^2 = 1 - \frac{\hat{\psi}_{ii}}{\hat{\sigma}_{ii}} \tag{D. 1}$$

The goodness-of-fit and the effect sizes are assessed through coefficients for each level-equation and for the global model (these coefficients assess the fraction of the variance explained by indicators), multiple correlations (mc) and Bentler- Raykov multiple correlations (mc^2) Bentler and Raykov (2000). These last two coefficients report how each dependent variable relates to the model's linear prediction. The Bentler–Raykov squared multiple correlations for the ith endogenous variable is computed as:

$$mc_i^2 = \frac{\widehat{Cov}(y_i, \hat{y}_i\ )}{\sqrt{\hat{\sigma}_{ii}\widehat{Var}(\hat{y}_i)}} \tag{D. 2}$$

where $\hat{\sigma}_{ii}$ is a diagonal element of $\hat{\Sigma}$, and $\widehat{Var}(\hat{y}_i)$ is a diagonal element of

$$\widehat{Var}(\hat{Y}) = \left(1 - \widehat{B}\right)^{-1} \hat{\Gamma}\hat{\Phi}\hat{\Gamma}' \left\{\left(1 - \widehat{B}\right)^{-1}\right\}' + \left\{\left(1 - \widehat{B}\right)^{-1} - I\right\} \hat{\Psi} \left\{\left(1 - \widehat{B}\right)^{-1} - I\right\}'$$

The coefficient of determination (CD) is also utilized to evaluate a model's best fitting to the sample data. The CD value demonstrates the percentage of variation identified in the endogenous constructs by the exogenous constructs (e.g., the CD for the equation level, while the R^2 For the whole model). The overall coefficient of determination is calculated as:

$$CD = 1 - \frac{det(\hat{\Psi})}{det(\hat{\Sigma})} \tag{D. 3}$$

The saturated model with degrees of freedom (df):

$$df_s = \binom{p+q+1}{2} + p + q \tag{D. 4}$$

where p is the number of observed endogenous variables, and q is the number of observed exogenous variables. Based on the observed variables, a reduced covariance matrix is fitted. All variables are uncorrelated in the baseline model if there are no endogenous variables (Browne and Cudeck, 1992). In the baseline model, the degree of freedom is expressed as:

$$df_b = \begin{cases} 2q & , if\ p = 0 \\ 2p + q + \binom{q+1}{2} & , if\ p > 0 \end{cases} \tag{D. 5}$$

The likelihood-ratio test of the baseline contrasted with saturated models is computed with degrees of freedom $df_{bs} = df_s - df_b$ as:

$$\chi^2_{bs} = 2(logL_s - logL_b) \tag{D. 6}$$

The likelihood-ratio test of the specified model against the saturated model is computed with degrees of freedom $df_{ms} = df_s - df_m$ as:

$$\chi^2_{ms} = 2\ \left(logL_s - logL(\hat{\theta})\right) \tag{D. 7}$$

It is also relevant to calculate the two information criteria that can be applied to demonstrate how the model fits the data set and compare the fit of different models. The first criterion computed is the Akaike Information Criterion(AIC), which estimates the amount of information lost by a given model (Akaike, 1974). The AIC suggests that the less information a model loses, the higher its quality (e.g., the smaller values imply a better fit). The is AIC described as:

$$AIC = -2\,log\,L\left(\hat{\theta}\right) + 2df_m \tag{D. 8}$$

The other criterion for model selection, which is closely related to (AIC), is the Bayesian Information Criterion known as (BIC) is based in part on the likelihood function (Schwarz, 1978). The Bayesian information criterion (BIC) is described as:

$$BIC = -2\,log\,L\left(\hat{\theta}\right) + Ndf_m \tag{D. 9}$$

The root mean squared error of approximation$(RMSEA)$ proposed by (Browne et al., 1993) is computed with the 90% confidence interval as:

$$RMSEA = \left\{\frac{(\chi^2_{ms} - df_{ms})G}{N\ df_{ms}}\right\}^{\frac{1}{2}}, RMSEA_{90\%\ C.I.}\left(\sqrt{\frac{G\lambda_L}{N\ df_{ms}}}, \sqrt{\frac{G\lambda_U}{N\ df_{ms}}}\right) \tag{D. 10}$$

where λ_L and λ_U are the non-centrally parameters corresponding to a noncentral chi-squared distribution with df_{ms} degrees of freedom in which chi-squared has a cumulative distribution function equal to 0.95 and 0.90, respectively. G is the number of groups. The $p - value$ test of close fit with the null hypothesis (H_0: $RMSEA \leq 0.05$) Browne et al. (1993) is computed as:

$$p = 1 - Pr\,(\chi^2 < \chi^2_{ms}\ |\lambda, df_{ms}), \qquad \lambda = (0.05)^2\ (N-1)df_{ms} \tag{D. 11}$$

where χ^2 is distributed noncentral chi-squared with noncentral parameter λ with df_{ms} degrees of freedom. Under baseline comparison, the comparative fit index (CFI) and the

Tucker–Lewis index (TLI) are employed to compare the fit of a hypothesized model to the baseline model (i.e., the worst fit model) (Bentler, 1990). They are computed, respectively, as:

$$CFI = 1 - \left[\frac{(\chi^2_{ms} - df_{ms})}{max\{(\chi^2_{bs} - \ df_{bs}), (\chi^2_{ms} - \ df_{ms})\}}\right] \tag{D. 12}$$

$$TLI = \frac{(\chi^2_{bs}/\ df_{bs}) - (\chi^2_{ms}/df_{ms})}{(\chi^2_{bs}/\ df_{bs}) - 1} \tag{D. 13}$$

where k is the number of observed variables. According to Mueller and Hancock (2008), if means are not in the fitted model, the standardized root mean squared residual $(SRMR)$ is computed as:

$$SRMR = \left\{\frac{2\sum_{i=1}^{k}\sum_{j\leq i} r_{ij}^2}{k(k+1)G}\right\}^{\frac{1}{2}} \tag{D. 14}$$

where r_{ij} is the standardized covariance residual.

$$r_{ij} = \ \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} - \ \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}} \tag{D. 15}$$

If means are in the fitted model, $SRMR$ is computed as

$$SRMR = \left\{\frac{2\sum_{i=1}^{k}\left(m_i^2 + \sum_{j\leq i} r_{ij}^2\right)}{k(k+3)G}\right\}^{\frac{1}{2}} \tag{D. 16}$$

where m_i is the standardized mean residual $m_i = \ \frac{\bar{z}_i}{\sqrt{s_{ii}}} - \ \frac{\hat{\mu}_i}{\sqrt{\hat{\sigma}_{ii}}}$.

Reference

Abratt, R., & Goodey, S. (1990). Unplanned buying and in-store stimuli in supermarkets. *Managerial and Decision Economics., 11*(2), 111-121.

Afshin, A., Sur, P., Fay, K., Cornaby, L., Ferrara, G., Salama, J., Mullany, E., Abate, K., Abbafati, C., & Abebe, Z. (2019). Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet, 393*(10184), 1958-1972.

Aitken, A. (1935). Note on selection from a multivariate normal population. *Proceedings of the Edinburgh Mathematical Society, 4*(2), 106-110.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control, 19*(6), 716-723.

Alinovi, L., Mane, E., & Romano, D. (2009). Measuring household resilience to food insecurity: application to Palestinian households. *EC-FAO Food Security Programme Rom*.

Anastasiou, K., Miller, M., & Dickinson, K. (2019). The relationship between food label use and dietary intake in adults: A systematic review. *Appetite, 138*, 280-291.

Aryal, J., Mottaleb, K., & Rahut, D. (2019). Untangling gender differentiated food security gaps in Bhutan: An application of exogenous switching treatment regression. *Review of Development Economics Journal, 23*(2), 782-802.

Ashworth, P. (2017). *Leptogenicity of the Food Environment and Food Choice Behaviour in Leisure Centres.* Sheffield Hallam University,

Babey, S., Diamant, A., Hastert, T., & Harvey, S. (2008). Designed for disease: the link between local food environments and obesity and diabetes.

Babey, S., Wolstein, J., & Diamant, A. (2011). Food environments near home and school related to consumption of soda and fast food.

Backett-Milburn, K., Wills, W., Roberts, M., & Lawton, J. (2010). Food, eating and taste: Parents' perspectives on the making of the middle class teenager. *Social Science and Medicine Journal, 71*(7), 1316-1323.

Barrett, C. (2002). Food security and food assistance programs. *Handbook of agricultural economics, 2*, 2103-2190.

Beatty, T., & Tuttle, C. (2015). Expenditure Response to Increases in In-Kind Transfers: Evidence from the Supplemental Nutrition Assistance Program. *American Journal of Agricultural Economics, 97*(2), 390-404. doi:10.1093/ajae/aau097

Bell, J., Mora, G., Hagan, E., Rubin, V., & Karpyn, A. (2013). Access to healthy food and why it matters: A review of the research. *Philadelphia, PA: The Food Trust*.

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin Journal, 107*(2), 238.

Bentler, P., & Raykov, T. (2000). On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology, 85*(1), 125.

Bergmans, R., Berger, L., Palta, M., Robert, S., Ehrenthal, D., & Malecki, K. (2018). Participation in the Supplemental Nutrition Assistance Program and maternal depressive symptoms: Moderation by program perception. *Social Science and Medicine Journal, 197*, 1-8.

Bianchi, C. (2017). Exploring urban consumers' attitudes and intentions to purchase local food in Chile. *Journal of Food Products Marketing, 23*(5), 553-569.

Bitler, M., & Haider, S. (2011). An economic view of food deserts in the united states. *Journal of Policy Analysis and Management, 30*(1), 153-176.
Blalock, H. (1964). 1964 Causal inferences in non-experimental research. Chapel Hill, NC, University of North Carolina Press.
Blanchard, T., & Matthews, T. (2007). Retail Concentration, Food Deserts, and Food Disadvantaged Communities
in Rural America. *Research Initiation Program at Mississippi State University*, 201-215.
Blaylock, J. (1989). An economic model of grocery shopping frequency. *Applied Economics, 21*(6), 843-852.
Block, D., & Kouba, J. (2006). A comparison of the availability and affordability of a market basket in two communities in the Chicago area. *Public Health Nutrition Journal, 9*(7), 837-845.
Bollen, K. (1987). Total, Direct, and Indirect Effects in Structural Equation Models. *Journal of Sociological Methodology, 17*, 37-69. doi:10.2307/271028
Bollen, K. (1989). *Structural Equations with Latent Variables*.
Bollen, K. (1998). Structural equation models. Enciclopedia of Biostatistics. P. Armitage & T. Colton. In: Sussex, England: John Wiley.
Botonaki, A., Natos, D., & Mattas, K. (2008). Exploring convenience food consumption through a structural equation model. *Journal of Food Products Marketing, 15*(1), 64-79.
Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26*(2), 211-243.
Breusch, T., & Pagan, A. (1980). The Lagrange multiplier test and its applications for the error components model with incomplete panels. *Review of Economic Studies Journal, 47*(1), 239-253.
Breyer, B., & Voss-Andreae, A. (2013). Food mirages: Geographic and economic barriers to healthful food access in Portland, Oregon. *Health and Place Journal, 24*, 131-139.
Brown, C., & Miller, S. (2008). The impacts of local markets: a review of research on farmers markets and community supported agriculture (CSA). *American Journal of Agricultural Economics, 90*(5), 1298-1302.
Browne, M., Cudeck, R., Bollen, K., & Long, J. (1993). Testing structural equation models.
Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological methods & research, 21*(2), 230-258.
Byrne, B. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*: routledge.
Caivano, S., & Domene, S. (2020). Consensus among experts on healthy eating and diet quality index. *Ciência & Saúde Coletiva, 25*, 2551-2560.
Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using stata* (Vol. 2): Stata press College Station, TX.
Caputo, V., Lusk, J. L., & Nayga, R. M. (2018). Choice experiments are not conducted in a vacuum: The effects of external price information on choice behavior. *Journal of Economic Behavior & Organization, 145*, 335-351. doi:https://doi.org/10.1016/j.jebo.2017.11.018

Carlson, S., & Keith-Jennings, B. (2018). SNAP Is Linked with Improved Nutritional Outcomes and Lower Health Care Costs. *Center on Budget and Policy Priorities, 17*.

Carson, J. (2002). Family spending power. *Perspectives on Labour and Income, 10*, 5-13.

Carter, R. L. (2006). Solutions for Missing Data in Structural Equation Modeling. *Research & Practice in Assessment, 1*, 4-7.

Caspi, C., Sorensen, G., Subramanian, S., & Kawachi, I. (2012). The local food environment and diet: a systematic review. *Health and place, 18*(5), 1172-1187.

CCPHA. (2007). Searching for healthy food: the food landscape in California cities and counties. In: California Center for Public Health Advocacy Davis (CA).

Charlton, K. (2016). Food security, food systems and food sovereignty in the 21st century: A new paradigm required to meet Sustainable Development Goals. *Nutrition & Dietetics, 73*(1), 3-12.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information systems research, 14*(2), 189-217.

CHR, C. H. R. (2019). *University of Wisconsin Population Health Institute. County Health Rankings Key Findings 2019.*

CHR, C. H. R. (2021). *University of Wisconsin Population Health Institute. County Health Rankings Key Findings 2021.* Retrieved from: https://www.countyhealthrankings.org/

Clemmitt, M. (2018). Issue: Minimum Wage Minimum Wage.

Cleveland, D., Carruth, A., & Mazaroli, D. (2015). Operationalizing local food: goals, actions, and indicators for alternative food systems. *Agriculture and Human Values, 32*(2), 281-297.

Cobas, J., Balcazar, H., Benin, M., Keith, V., & Chong, Y. (1996). Acculturation and low-birthweight infants among Latino women: a reanalysis of HHANES data with structural equation models. *American journal of public health, 86*(3), 394-396.

Cobb, L., Appel, L., Franco, M., Jones-Smith, J., Nur, A., & Anderson, C. (2015). The relationship of the local food environment with obesity: a systematic review of methods, study quality, and results. *Obesity, 23*(7), 1331-1344.

Coleman-Jensen, A., Matthew, P., Christian, A., & Anita, S. (2017). Household Food Security in the United States in 2016. [Economic Research Service.]. (ERR-237).

Coleman-Jensen, A., Nord, M., Andrews, M., & Carlson, S. (2012). Household Food Security in the United States in 2011.

Coleman-Jensen, A., Rabbitt, M., Gregory, C., & Singh, A. (2019). Household Food Security in the United States in 2018.

Cooksey-Stowers, K., Schwartz, M., & Brownell, K. (2017). Food swamps predict obesity rates better than food deserts in the United States. *International journal of environmental research and public health, 14*(11), 1366.

Cooper, B. (2017). What drives compliance? An application of the theory of planned behaviour to urban water restrictions using structural equation modelling. *Applied Economics, 49*(14), 1426-1439.

Cover, T., & Thomas, J. (2012). *Elements of information theory*: John Wiley & Sons.

Cudeck, R., Jöreskog, K., Sörbom, D., & Du Toit, S. (2001). *Structural equation modeling: Present and future: A Festschrift in honor of Karl Jöreskog*: Scientific Software International.

Cummins, S., Flint, E., & Matthews, S. (2014). New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity. *Health Affairs, 33*(2), 283-291.

Cummins, S., & Macintyre, S. (2002). "Food deserts"—evidence and assumption in health policy making. *Bmj, 325*(7361), 436-438.

Dallas, S., Liu, P., & Ubel, P. (2015). Potential problems with increasing serving sizes on the nutrition facts label. *Appetite, 95*, 577-584.

Datz, T. (2011). Harvard serves up its own "plate". *Harvard Gazette, 14*.

Deaton, A. (1989). *Saving and liquidity constraints* (0898-2937). Retrieved from

Deb, P., & Gregory, C. (2016). *Who Benefits Most from SNAP? A Study of Food Security and Food Spending*. Retrieved from

Del Castillo, M., Iriondo-DeHond, A., Iriondo-DeHond, M., Gonzalez, I., Medrano, A., Filip, R., & Uribarri, J. (2020). Healthy eating recommendations: good for reducing dietary contribution to the body's advanced glycation/lipoxidation end products pool? *Nutrition Research Reviews*, 1-16.

Devaux, M., Goryakin, Y., Cecchini, M., Huber, H., & Colombo, F. (2017). OECD Obesity Update 2017. In: OECD: Paris, France.

Dong, D., & Lin, B. (2009). *Fruit and vegetable consumption by low-income Americans: would a price reduction make a difference?* Retrieved from

Draper, C., & Younginer, N. (2020). P131 Identifying SNAP Participant Experiences and Environments to Inform Tailored SNAP-Ed Programming. *Journal of Nutrition Education and Behavior, 52*(7), S78.

Drewnowski, A., Aggarwal, A., Hurvitz, P., Monsivais, P., & Moudon, A. (2012). Obesity and supermarket access: proximity or price? *American journal of public health, 102*(8), e74-e80.

Drewnowski, A., & Specter, S. (2004). Poverty and obesity: the role of energy density and energy costs. *The American journal of clinical nutrition, 79*(1), 6-16.

Dubowitz, T., Zenk, S., Ghosh-Dastidar, B., Cohen, D., Beckman, R., Hunter, G., Steiner, E., & Collins, R. (2015). Healthy food access for urban food desert residents: examination of the food environment, food purchasing practices, diet and BMI. *Public Health Nutrition Journal, 18*(12), 2220-2230.

Duncan, O. (1966). Path analysis: Sociological examples. *American journal of Sociology, 72*(1), 1-16.

Duncan, O. (2014). *Introduction to structural equation models*: Elsevier.

Dutko, P. (2012). Food deserts suffer persistent socioeconomic disadvantage. *Choices, 27*(3), 1-4.

Eisenhauer, N., Bowker, M., Grace, J., & Powell, J. (2015). From patterns to causal understanding: structural equation modeling (SEM) in soil ecology. *Pedobiologia, 58*(2-3), 65-72.

Elbel, B., Tamura, K., McDermott, Z., Wu, E., & Schwartz, A. (2020). Childhood Obesity and the Food Environment: A Population-Based Sample of Public School Children in New York City. *Obesity, 28*(1), 65-72.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987-1007.

ERS. (2016). Nutrient Coding Overview National Household Food Acquisition and Purchase Survey (FoodAPS). from U.S. Department of Agriculture

FAO. (2006). Assessing the impact of food aid on recipient countries: a survey. *Food and Agriculture Organization Publications, Rome*.

FAO. (2016). The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. *Food and Agriculture Organization Publications, Rome*.

Feldman, B., & Wolnik, D. (2019). *Farmers markets as a strategy for improving food environments*.

Finch, J., West, S., & MacKinnon, D. (1997). Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models. *Structural Equation Modeling A Multidisciplinary Journal, 4*(2), 87-107.

Fulfrost, B., & Howard, P. (2006). Mapping the markets: the relative density of retail food stores in densely populated census blocks in the central coast region of California.

Gale, F. (1997). Direct farm marketing as a rural development tool. *Rural America/Rural Development Perspectives, 12*(2221-2019-2615), 19-25.

Gallant, A., & Jorgenson, D. (1979). Statistical inference for a system of simultaneous, non-linear, implicit equations in the context of instrumental variable estimation. *Journal of Econometrics 11*(2-3), 275-302.

Geary, R. C. (1970). Relative efficiency of count of sign changes for assessing residual autoregression in least squares regression.

Ghosh-Dastidar, B., Cohen, D., Hunter, G., Zenk, S., Huang, C., Beckman, R., & Dubowitz, T. (2014). Distance to store, food prices, and obesity in urban food deserts. *American journal of preventive medicine, 47*(5), 587-595.

Ghosh-Dastidar, B., Hunter, G., Collins, R., Zenk, S., Cummins, S., Beckman, R., Nugroho, A., Sloan, J., & Dubowitz, T. (2017). Does opening a supermarket in a food desert change the food environment? *Health and place, 46*, 249-256.

Glanz, K., Sallis, J., Saelens, B., & Frank, L. (2007a). Nutrition Environment Measures Survey in stores (NEMS-S): development and evaluation. *American journal of preventive medicine, 32*(4), 282-289.

Glanz, K., Sallis, J., Saelens, B., & Frank, L. (2007b). Nutrition Environment Measures Survey in stores (NEMS-S): development and evaluation. *American journal of preventive medicine, 32*(4), 282-289. doi:10.1016/j.amepre.2006.12.019

Goldberger, A. (1972). Structural equation methods in the social sciences. *Econometrica, Journal of the Econometric Society*, 979-1001.

Grace, J. (2006). *Structural equation modeling and natural systems*: Cambridge University Press.

Grace, J., Anderson, T., Olff, H., & Scheiner, S. (2010). On the specification of structural equation models for ecological systems. *Ecological Monographs, 80*(1), 67-87.

Greene, W. (2012). *Econometric Analysis (6th ed.)*: Journal of Boston: Pearson Education.

Gregory, C., Ver Ploeg, M., Andrews, M., & Coleman-Jensen, A. (2013). *Supplemental Nutrition Assistance Program (SNAP) participation leads to modest changes in diet quality*. Retrieved from

Grote, U. (2018). IFPRI: Global Food Policy Report 2018. *Food Security, 10*(6), 1665-1667. doi:10.1007/s12571-018-0844-x

Guenther, P., Casavale, K., Reedy, J., Kirkpatrick, S., Hiza, H., Kuczynski, K., Kahle, L., & Krebs-Smith, S. (2013). Update of the healthy eating index: HEI-2010. *Journal of the Academy of Nutrition and Dietetics, 113*(4), 569-580.

Gundersen, C., Brown, J., Engelhard, E., & Waxman, E. (2011). Map the meal gap: Technical brief. *Chicago, IL: Feeding America*.

Gundersen, C., Kreider, B., & Pepper, J. (2017). Partial identification methods for evaluating food assistance programs: A case study of the causal impact of SNAP on food insecurity. *American Journal of Agricultural Economics, 99*(4), 875-893.

Gundersen, C., & Oliveira, V. (2001). The Food Stamp Program and Food Insufficiency. *American Journal of Agricultural Economics, 83*(4), 875-887. doi:https://doi.org/10.1111/0002-9092.00216

Gundersen, C., Waxman, E., Engelhard, E., & Del Vecchio, T. (2012). Map the meal gap 2012: Ohio food insecurity estimates at the county level. *Feeding America*.

Gundersen, C., & Ziliak, J. (2015). Food Insecurity And Health Outcomes. *Health Affairs, 34*(11), 1830-1839. doi:10.1377/hlthaff.2015.0645

Gustat, J., O'Malley, K., Luckett, B., & Johnson, C. (2015). Fresh produce consumption and the association between frequency of food shopping, car access, and distance to supermarkets. *Preventive Medicine Journal*

2, 47-52.

Guthrie, J., Derby, B., & Levy, A. (1999). What people know and do not know about nutrition. *America's eating habits: Changes and consequences*, 243-290.

Gutman, G. (2002). Impact of the rapid rise of supermarkets on dairy products systems in Argentina. *Development Policy Review, 20*(4), 409-427.

Hair, J., Anderson, R., Tatham, R., & William, C. (1998). Multivariate data analysis. In (pp. 454): Upper Saddle River, NJ: Prentice Hall.

Hall, K. (2018). Did the Food Environment Cause the Obesity Epidemic? *Obesity (Silver Spring, Md.), 26*(1), 11-13. doi:10.1002/oby.22073

Hamidi, S. (2020). Urban sprawl and the emergence of food deserts in the USA. *Journal of Urban Studies, 57*(8), 1660-1675.

Hamilton, S. (2018). *Supermarket USA: Food and Power in the Cold War Farms Race*: Yale University Press.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica, 46*(6), 1251-1271. doi:10.2307/1913827

Heinrich, C., & Betts, B. (2003). *Adapt or die: transforming your supply chain into an adaptive business network*: John Wiley & Sons.

Hendrickson, D., Smith, C., & Eikenberry, N. (2006). Fruit and vegetable access in four low-income food deserts communities in Minnesota. *Agriculture and Human Values., 23*(3), 371-383.

HHS. (2015). *US Department of Health and Human Services and U.S. Department of Agriculture*

Dietary guidelines for Americans 2015-2020: Government Printing Office.

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., & Moreira, V. (2014). Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador. *Journal of Development Economics, 107*, 144-156.

Hillier, A., McLaughlin, J., Cannuscio, C., Chilton, M., Krasny, S., & Karpyn, A. (2012). The impact of WIC food package changes on access to healthful food in 2 low-income urban neighborhoods. *Journal of Nutrition Education and Behavior, 44*(3), 210-216.

Hilmers, A., Hilmers, D., & Dave, J. (2012). Neighborhood disparities in access to healthy foods and their effects on environmental justice. *American Journal of Public Health, 102*(9), 1644-1654.

Hite, A. (2019). A Material-Discursive Exploration of "Healthy Food" and the Dietary Guidelines for Americans.

Hoefer, R., & Curry, C. (2012). Food Security and Social Protection in the United States. *Journal of Policy Practice, 11*(1-2), 59-76. doi:10.1080/15588742.2011.624062

Hsu, S., Chang, C., & Lin, T. (2016). An analysis of purchase intentions toward organic food on health consciousness and food safety with/under structural equation modeling. *British Food Journal.*

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling A Multidisciplinary Journal, 6*(1), 1-55.

Irving, B., John, W., & Wayne, V. (2012). Structural Equation Modeling. In *Handbook of Psychology, Second Edition*.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, 163-172.

Jöreskog, K. (1970). A general method for analysis of covariance structures. *Biometrika Jounal, 57*(2), 239-251.

Kardes, F. R., Cronley, M. L., Kellaris, J. J., & Posavac, S. S. (2004). The Role of Selective Information Processing in Price-Quality Inference. *Journal of Consumer Research, 31*(2), 368-374. doi:10.1086/422115

Kelly, B., Flood, V., & Yeatman, H. (2011). Measuring local food environments: an overview of available methods and measures. *Health and Place Journal, 17*(6), 1284-1293.

Kenny, D. (2019). Enhancing Validity in Psychological Research. *Journal of American Psychologist., 74*(9), 1018.

Kia-Keating, M., Nylund-Gibson, K., Kia-Keating, B., Schock, C., & Grimm, R. (2018). Longitudinal Patterns of Self-Regulation among Ethnic Minority Children Facing Poverty. *Journal of Child and Family Studies., 27*(2), 398-411.

Kline, R. (2011). *Convergence of structural equation modeling and multilevel modeling*: na.

Krebs-Smith, S., Pannucci, T., Subar, A., Kirkpatrick, S., Lerman, J., Tooze, J., Wilson, M., & Reedy, J. (2018). Update of the healthy eating index: HEI-2015. *Journal of the Academy of Nutrition and Dietetics, 118*(9), 1591-1602.

Kröller, K., & Warschburger, P. (2009). Maternal feeding strategies and child's food intake: considering weight and demographic influences using structural equation

modeling. *International Journal of Behavioral Nutrition and Physical Activity., 6*(1), 78.
Krukowski, R., West, D., Harvey-Berino, J., & Prewitt, T. (2010). Neighborhood impact on healthy food availability and pricing in food stores. *Journal of Community Health, 35*(3), 315-320.
Lai, P., & Bessler, D. (2010). *On aggregation bias in structural demand models*. Retrieved from
Laska, M., Caspi, C., Pelletier, J., Friebur, R., & Harnack, L. (2015). Peer reviewed: Lack of healthy food in small-size to mid-size retailers participating in the Supplemental Nutrition Assistance Program, Minneapolis–St. Paul, Minnesota, 2014. *Preventing Chronic Disease CDC, 12*.
Lee, H. (2004). The triple-A supply chain. *Harvard Business Review. , 82*(10), 102-113.
Lee, Y., Mozaffarian, D., Sy, S., Huang, Y., Liu, J., Wilde, P., Abrahams-Gessel, S., Jardim, T., Gaziano, T., & Micha, R. (2019). Cost-effectiveness of financial incentives for improving diet and health through Medicare and Medicaid: A microsimulation study. *PLoS Medicine Journal, 16*(3), e1002761.
Liese, A. D., Weis, K. E., Pluto, D., Smith, E., & Lawson, A. (2007). Food store types, availability, and cost of foods in a rural environment. *Journal of the American Dietetic Association, 107*(11), 1916-1923.
Lim, S., & Melville, N. P. (2009). Robustness of structural equation modeling to distributional misspecification: Empirical evidence & research guidelines. *Available at SSRN 1375251*.
Loehlin, J. (1987). *Latent variable models*: hillsdale, nj: erlbaum.
Luan, H., Law, J., & Quick, M. (2015). Identifying food deserts and swamps based on relative healthy food access: a spatio-temporal Bayesian approach. *International Journal of Health Geographics, 14*(1), 37.
Lusk, J. L., & Weaver, A. (2017). An experiment on cash and in-kind transfers with application to food assistance programs. *Food Policy, 68*, 186-192. doi:https://doi.org/10.1016/j.foodpol.2017.02.005
Mackenbach, J., Burgoine, T., Lakerveld, J., Forouhi, N., Griffin, S., Wareham, N., & Monsivais, P. (2017). Accessibility and affordability of supermarkets: associations with the DASH diet. *American journal of preventive medicine, 53*(1), 55-62.
Maguire, E., Burgoine, T., & Monsivais, P. (2015). Area deprivation and the food environment over time: A repeated cross-sectional study on takeaway outlet density and supermarket presence in Norfolk, UK, 1990–2008. *Health and Place Journal 33*, 142-147.
Mancino, L., Todd, J., & Lin, B. (2009). Separating what we eat from where: measuring the effect of food away from home on diet quality. *Journal of Food Policy, 34*(6), 557-562.
Mancino, L., Todd, J., & Scharadin, B. (2018a). *USDA's National Household Food Acquisition and Purchase Survey: Methodology for Imputing Missing Quantities To Calculate Healthy Eating Index-2010 Scores and Sort Foods Into ERS Food Groups*. Retrieved from
Mancino, L., Todd, J. E., & Scharadin, B. (2018b). *USDA's National Household Food Acquisition and Purchase Survey: Methodology for Imputing Missing Quantities*

To Calculate Healthy Eating Index-2010 Scores and Sort Foods Into ERS Food Groups. Retrieved from
Maxwell, S., & Smith, M. (1992). *Household food security: a conceptual review*. Retrieved from Rome and New York:
McGuire, S., Todd, J., Mancino, L., & Lin, B. (2010). The impact of food away from home on adult diet quality. ERR-90. *US Department of Agriculture*, 442-443.
Morland, K., Wing, S., & Roux, A. (2002). The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *American journal of Public Health., 92*(11), 1761-1768.
Mosheim, R. (2012). *A quarterly econometric model for short-term forecasting of the US dairy industry*. Retrieved from
Mueller, R., & Hancock, G. (2008). *Best Practices in Quantitative Methods* (Vol. 488508). Clemson University, USA: SAGE Publications, Inc.
Muthen, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics., 22*(1-2), 43-65.
Muthen, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika Journal, 49*(1), 115-132.
Naja-Riese, A., Keller, K. J., Bruno, P., Foerster, S. B., Puma, J., Whetstone, L., MkNelly, B., Cullinen, K., Jacobs, L., & Sugerman, S. (2019). The SNAP-Ed Evaluation Framework: demonstrating the impact of a national framework for obesity prevention in low-income populations. *Translational behavioral medicine, 9*(5), 970-979.
Neff, R., Spiker, M., Rice, C., Schklair, A., Greenberg, S., & Leib, E. (2019). Misunderstood food date labels and reported food discards: A survey of US consumer attitudes and behaviors. *Waste Management & Research, 86*, 123-132.
Nelson, L. S. (1998). The Anderson-Darling test for normality. *Journal of Quality Technology, 30*(3), 298.
Neuhofer, Z., McFadden, B., Rihn, A., Wei, X., Khachatryan, H., & House, L. (2020). Can the updated nutrition facts label decrease sugar-sweetened beverage consumption? *Economics & Human Biology., 37*, 100867.
Nielsen, R., Seay, M., & Wilmarth, M. (2017). Does Prior Government Assistance Reduce Food or Housing Assistance among Low-Income and Food Insecure Households? *Journal of Consumer Affairs.*
Nord, M. (2008). *Household Food Security in the US (2006)*: Diane Publishing.
Nord, M., & Golla, A. (2009). *Does SNAP Decrease Food Insecurity? Untangling the Self-Selection Effect*: Economic Research Service.
North, A., Hargreaves, D., & McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology., 84*(2), 271.
O'Hara, J., & Benson, M. (2019). Where have all the direct-marketing farms gone? *Journal of Agriculture, Food Systems, and Community Development, 9*(1), 31-37.
Olivas, R., Díaz, M., & Bernabéu, R. (2013). Structural Equation Modeling of lifestyles and consumer attitudes towards organic food by income: a Spanish case study. *International Journal of Agriculture and Natural Resources., 40*(2), 265-277.
Olsson, U., Foss, T., Troye, S., & Howell, R. (2000). The Performance of ML, GLS, and WLS Estimation in Structural Equation Modeling Under Conditions of

Misspecification and Nonnormality. *Structural Equation Modeling A Multidisciplinary Journal, 7*(4), 557-595.
Orleck, A. (2018). *We Are All Fast-Food Workers Now: The Global Uprising Against Poverty Wages*: Beacon Press.
Pagan, A., & Hall, A. (1983). Diagnostic tests as residual analysis. *Econometric Reviews, 2*(2), 159-218.
Palameta, B., & Macredie, I. (2005). *Property taxes relative to income*: Statistiscs Canada.
Pearl, J. (2000). Causality: models, reasoning, and inference. *Cambridge University Press., 521*(77362), 8.
Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys Journal, 3*, 96-146.
Pearl, J. (2012). *The causal foundations of structural equation modeling*. Retrieved from
Penney, T., Almiron-Roig, E., Shearer, C., McIsaac, J., & Kirk, S. (2014). Modifying the food environment for childhood obesity prevention: challenges and opportunities. *Proceedings of the Nutrition Society., 73*(2), 226-236.
Pennings, J., & Leuthold, R. (2000). The role of farmers' behavioral attitudes and heterogeneity in futures contracts usage. *American Journal of Agricultural Economics, 82*(4), 908-919.
Pinstrup-Andersen, P. (2009). Food security: Definition and Measurement. *Food Security Journal, 1*(1), 5-7.
Pitt, E., Gallegos, D., Comans, T., Cameron, C., & Thornton, L. (2017). Exploring the influence of local food environments on food behaviours: a systematic review of qualitative literature. *Public Health Nutrition Journal, 20*(13), 2393-2405.
Plakias, Z. T., Demko, I., & Katchova, A. L. (2020). Direct marketing channel choices among US farmers: evidence from the Local Food Marketing Practices Survey. *Renewable Agriculture and Food Systems, 35*(5), 475-489. doi:10.1017/S1742170519000085
Powell, L., Chriqui, J., Khan, T., Wada, R., & Chaloupka, F. (2013). Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes. *Obesity reviews, 14*(2), 110-128.
Powell, L. M., Slater, S., Mirtcheva, D., Bao, Y., & Chaloupka, F. J. (2007). Food store availability and neighborhood characteristics in the United States. *Preventive medicine, 44*(3), 189-195.
Prathap, V. (2018). SNAP: The Supplemental Nutrition Assistance Program and Its Future. Retrieved from https://pedsadvocacy.com/2018/03/02/snap-the-supplemental-nutrition-assistance-program-and-its-future/
Rawlins, E., Baker, G., Maynard, M., & Harding, S. (2013). Perceptions of healthy eating and physical activity in an ethnically diverse sample of young children and their parents: the DEAL prevention of obesity study. *Journal of Human Nutrition and Dietetics., 26*(2), 132-144.
Reel, J., & Badger, B. (2014). From food deserts to food swamps: health education strategies to improve food environments in urban areas. *Journal of Obesity & Weight Loss Therapy., 4*(Special Issue).

Rhone, A., Ver Ploeg, M., Dicken, C., Williams, R., & Breneman, V. (2017). *Low-income and low-supermarket-access census tracts, 2010-2015*. Retrieved from

Rhone, A., Ver Ploeg, M., Williams, R., & Breneman, V. (2019). *Understanding Low-Income and Low-Access Census Tracts Across the Nation: Subnational and Subpopulation Estimates of Access to Healthy Food*. Retrieved from

Rivera, R. L., Maulding, M. K., Abbott, A. R., Craig, B. A., & Eicher-Miller, H. A. (2016). SNAP-Ed (Supplemental Nutrition Assistance Program–Education) increases long-term food security among Indiana households with children in a randomized controlled study. *The Journal of nutrition, 146*(11), 2375-2382.

Rose, D., Bodor, J., Swalm, C., Rice, J., Farley, T., & Hutchinson, P. (2009). Deserts in New Orleans? Illustrations of urban food access and implications for policy. *University of Michigan National Poverty Center/USDA Economic Research Service Research, Ann Arbor, MI*.

Rose, D., & Richards, R. (2004). Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutrition Journal, 7*(8), 1081-1088.

Rossi, R. (2018). *Mathematical statistics: an introduction to likelihood based inference*: John Wiley & Sons.

Rummo, P., Guilkey, D., Ng, S., Meyer, K., Popkin, B., Reis, J., Shikany, J., & Gordon-Larsen, P. (2017). Does unmeasured confounding influence associations between the retail food environment and body mass index over time? The Coronary Artery Risk Development in Young Adults (CARDIA) study. *International Journal of Epidemiology., 46*(5), 1456-1464.

Saelens, B., Glanz, K., Sallis, J., & Frank, L. (2007a). Nutrition Environment Measures Study in restaurants (NEMS-R): development and evaluation. *American Journal of Preventive Medicine., 32*(4), 273-281. doi:10.1016/j.amepre.2006.12.022

Saelens, B. E., Glanz, K., Sallis, J. F., & Frank, L. D. (2007b). Nutrition Environment Measures Study in restaurants (NEMS-R): development and evaluation. *American journal of preventive medicine, 32*(4), 273-281.

Saghaian, S. H., & Reed, M. R. (2004). Demand for quality-differentiated beef in Japan. *German Journal of Agricultural Economics, 53*(670-2016-45745), 344-351.

Sarnacchiaro, P., & Boccia, F. (2018). Some remarks on measurement models in the structural equation model: an application for socially responsible food consumption. *Journal of Applied Statistics., 45*(7), 1193-1208.

Scaglioni, S., Salvioni, M., & Galimberti, C. (2008). Influence of parental attitudes in the development of children eating behaviour. *British Journal of Nutrition., 99*(S1), S22-S25.

Schanzenbach, D., Bauer, L., & Nantz, G. (2016). *Twelve facts about food insecurity and SNAP*: Brookings Institution.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics., 6*(2), 461-464.

Scott, P. (2017). Global panel on agriculture and food systems for nutrition: food systems and diets: facing the challenges of the 21st century. In: Springer.

Seligman, H., Laraia, B., & Kushel, M. (2009). Food insecurity is associated with chronic disease among low-income NHANES participants. *The Journal of Nutrition., 140*(2), 304-310.

Shaw, D. (2007). World Food Summit, 1996. In *World Food Security* (pp. 347-360): Springer.
Shiftan, Y., Outwater, M., & Zhou, Y. (2008). Transit market research using structural equation modeling and attitudinal market segmentation. *Journal of Transport Policy, 15*(3), 186-195.
Shook, C., Ketchen Jr, D., Hult, G., & Kacmar, K. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal, 25*(4), 397-404.
Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations: Volume One. In: London: printed for W. Strahan; and T. Cadell, 1776.
Smith, R., Davis, A., Jordan, N., Atwood, L., Daly, A., Grandy, A., Hunter, M., Koide, R., Mortensen, D., & Ewing, P. (2014). Structural equation modeling facilitates transdisciplinary research on agriculture and climate change. *Crop Science., 54*(2), 475-483.
Sobel, M. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Journal of Sociological Methodology, 13*, 290-312.
Source, T. N. (2008). *Healthy eating plate and healthy eating pyramid. Harvard School of Public Health*.
Southworth, H. M. (1945). The Economics of Public Measures to Subsidize Food Consumption. *Journal of Farm Economics, 27*(1), 38-66. doi:10.2307/1232262
Taylor, R., & Villas-Boas, S. B. (2016). Food store choices of poor households: A discrete choice analysis of the National Household Food Acquisition and Purchase Survey (FoodAPS). *American Journal of Agricultural Economics, 98*(2), 513-532.
Todd, J., Jessica, E., Mancino, L., & Lin, B. (2010). The impact of food away from home on adult diet quality.
Turner, C., Aggarwal, A., Walls, H., Herforth, A., Drewnowski, A., Coates, J., Kalamatianou, S., & Kadiyala, S. (2018). Concepts and critical perspectives for food environment research: A global framework with implications for action in low-and middle-income countries. *Global Food Security., 18*, 93-101.
Turner, C., Kadiyala, S., Aggarwal, A., Coates, J., Drewnowski, A., Hawkes, C., Herforth, A., Kalamatianou, S., & Walls, H. (2017a). Concepts and methods for food environment research in low and middle income countries. *Technical Brief, Agriculture, Nutrition and Health Academy Food Environments Working Group on Innovative Methods and Metrics for Agriculture and Nutrition, London*.
Turner, C., Kadiyala, S., Aggarwal, A., Coates, J., Drewnowski, A., Hawkes, C., Herforth, A., Kalamatianou, S., & Walls, H. (2017b). Concepts and methods for food environment research in low and middle income countries. *Nutrition and Health Academy*.
Ullman, J., & Bentler, P. (2012). Structural equation modeling handbook of psychology. In: John Wiley & Sons, Inc.
USDA. (2015). *Food Environment Atlas*. Retrieved from https://www.ers.usda.gov/data-products/food-environment-atlas/
USDA. (2019). *Annual Retail Trade Survey*. Retrieved from https://www.census.gov/econ/overview/re0200.html

Van der Linden, S. (2014). On the relationship between personal experience, affect and risk perception: The case of climate change. *European Journal of Social Psychology., 44*(5), 430-440.
Ver Ploeg, M. (2010). *Access to affordable and nutritious food: measuring and understanding food deserts and their consequences: report to Congress*: Diane Publishing.
Walker, R., Keane, C., & Burke, J. (2010a). Disparities and access to healthy food in the United States: A review of food deserts literature. *Health and Place., 16*(5), 876-884.
Walker, R. E., Keane, C. R., & Burke, J. G. (2010b). Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & place, 16*(5), 876-884.
Wang, Y., McPherson, K., Marsh, T., Gortmaker, S., & Brown, M. (2011). Health and economic burden of the projected obesity trends in the USA and the UK. *The Lancet, 378*(9793), 815-825.
Weatherspoon, D., Oehmke, J., Dembele, A., & Weatherspoon, L. (2015). Fresh vegetable demand behaviour in an urban food desert. *Journal of Urban Studies, 52*(5), 960-979.
WHO, W. H. O. O. a. o. (2021). Obesity and overweight. *World Health Organization*, 1-2. Retrieved from https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight
Wilde, P. (2018). *Food policy in the United States: An introduction*: Routledge.
Wilde, P. E. (2013). The new normal: The supplemental nutrition assistance program (snap). *American Journal of Agricultural Economics, 95*(2), 325-331.
Willett, W., & Skerrett, P. (2017). *Eat, drink, and be healthy: the Harvard Medical School guide to healthy eating*: Simon and Schuster.
Willis, S., Simonsen, S., Hemmert, R., Baayd, J., Digre, K., & Zick, C. (2020). Food Insecurity and the Risk of Obesity, Depression, and Self-Rated Health in Women. *Women's Health Reports., 1*(1), 308-317.
Woodham, C. (2011). *Food desert or food swamp? An in-depth exploration of neighbourhood food environments in Eastern Porirua and Whitby.* University of Otago,
Wright, S. (1918). On the nature of size factors. *Genetics., 3*(4), 367.
Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics., 5*(3), 161-215.
Xin, J., Zhao, L., Wu, T., Zhang, L., Li, Y., Xue, H., Xiao, Q., Wang, R., Xu, P., & Visscher, T. (2019). Association between access to convenience stores and childhood obesity: a systematic review. *Obesity reviews*.
Yan, R., Bastian, N. D., & Griffin, P. M. (2015). Association of food environment and food retailers with obesity in US adults. *Health & Place, 33*, 19-24.
Zellner, A., & Theil, A. (1962). Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations. *Econometrica, 30*, 54-78. Retrieved from http://www.jstor.org/stable/1911287
Zenk, S., Schulz, A., Izumi, B., Mentz, G., Israel, B., & Lockett, M. (2013). Neighborhood food environment role in modifying psychosocial stress–diet relationships. *Appetite, 65*, 170-177.

Zhang, C., Chen, B., & Pearl, J. (2019). A Simultaneous Discover-Identify Approach to Causal Inference in Linear Models.
Ziliak, J., & Gundersen, C. (2016). Food Security and Geographic Factors in Food Purchase and Acquisition Decisions: A Compilation of Research Conducted under USDA Cooperative Agreements 58-5000-1-0050 and 58-5000-3-0066.

VITA

SULIMAN ALMOJEL

PROFESSIONAL SPECIALTY:

- Research and Teaching
- Mathematical Programming Modeling (AIMMS, GAMS)
- Game Theory in Economics
- Risk Management in Economics
- Econometrics, Statistical Analysis (STATA, R)
- Geographic Information System (ArcGIS)

EDUCATION:

- **Ph.D.** Agricultural Economics, Aug, 2021
  UNIVERSITY OF KENTUCKY, USA
- **M.S.** Agricultural Economics, May, 2016
  UNIVERSITY OF KENTUCKY, USA
- **B.S.** Agricultural Economics, Aug, 2010
  KING SAUD UNIVERSITY, KSA

EXPERIENCE

- **Assistant Professor,** 2021- Present at KING SAUD UNIVERSITY, KSA
- **Reviewer** 2018 - 2021 Saltonstall-Kennedy (S-K) Grant, National Oceanic and Atmospheric Administration (NOAA)
- **Teacher Assistant**, 2012- 2014 at KING SAUD UNIVERSITY, KSA

LANGUAGE

- **Arabic**
- **English**