EXAMINING THE GEOGRAPHY OF FOOD DESERTS AND FOOD SWAMPS IN AUSTIN, TEXAS

by

He Jin, B.A., M.S.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Geographic Information Science
May 2019

Committee Members:

Yongmei Lu, Chair

Benjamin Zhan

Russell Weaver

Tonny Oyana

Lesli Biediger-Friedman

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

DEDICATION

To my father and mother who inspired my pursuit of doctoral degree. This is for both of you.

ACKNOWLEDGEMENTS

This dissertation work was completed under the supervision of my advisor Dr. Yongmei Lu. I want to express my sincere gratitude to her for her guidance, support, and encouragement at Texas State University. I have been deeply influenced by Dr. Lu's outstanding personalities — creativity, earnestness, persistence, and ambitiousness. She brought me to the frontier of food environment and health geography studies. I thank her for teaching me how to conduct scientific research, publish research papers with different journals, and respond to review comments through manuscript revisions. She shared with me without reservation her experiences on the various aspects of academic life, including teaching, scientific writing, and job searching. She also supported me to overcome the difficulty times and helped me fit in the department. Through the many conversations we had, she always encouraged me to stay true to my dream and not give up easily. Dr. Lu funded my dissertation survey; I could not have completed my research without her generous support.

I also would like to thank my dissertation committee members: Dr. F. Ben Zhan, who gave me critical comments and helped me to improve my research; Dr. Russell Weaver, who helped me refine my research questions and recommended alternative statistical analysis methods; Dr. Tony Oyana, for his suggestions on my literature reviews and theoretical framework; Dr. Lesli Biediger-Friedman who assisted me in developing the food survey instrument. I could not have completed my dissertation without their tremendous help and professional support.

My special thanks go to the Department of Geography at Texas State University. I appreciate the finical support from our department for the past four years. I want to thank two faculty members especially Drs. Lawrence E. Estaville and Xingde Jia at Texas State University for their financial support in my fifth year. Many faculty members, including Drs. Nate Currit, Alberto Giordano, Ronald Hagelman, Edwin Chow, and Yihong Yuan, have guided me through the various courses I took. I appreciate the staff members in the department — Allison Glass-Smith, Angelika Wahl, Pat Hell−Jones, and Charles Robinson for their support during my doctoral program.

I want to express my gratitude to my fellow graduate students: Yunuen Reygadas, Shadae Dixon, Zahra Ghaffari, Francesco Zignol, Saeideh Gharehchahi, Nathaniel Dede-Bamfo, Graciela Sandoval, Niaz Morshed, and Brendan Lavy. With my Chinese friends, Xujiao Wang, Guixing Wei, Wen Lu, Xiaowen Cui, Zhaoceng song, and Yahan Teng, we gathered together on weekends and made traditional Chinese cuisines, which warmed up my heart, and kept me from being lonely in the U.S. I thank them for their time, kindness, company, and spiritual support. I highly appreciate the tremendous time and efforts my peer, Dr. Ning Niu, spent to help with some of the technical issues for my dissertation analyses. My appreciation also goes to my best friends Xianhui Li and Xingkong Ma, who comforted me when I encountered difficult times.

I am very grateful to my master thesis advisor, Dr. Shixiong Hu, for his consistent support and care in the past five years. Thanks also go to my master program lecturer Bill Chappel and his wife Kim Chappel for their warm blessing during my pursuit of this doctoral degree.

Finally, I would like to take this opportunity to thank my parents, Tianding Jin and Qunying Zhou, who raised me to be a responsible person. I thank my brother Qiao Jin and sister-in-law Huijie Chen and my lovely niece Yuxi Jin and nephew Baohuan Jin, for bringing so much happiness into my life. Special thanks must go to my loving husband, Shuai Ji, for his love, patience, and encouragement along the way.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

In the past two decades, retail food environment exerts a tremendous fascination with scholars because it can shape individuals' eating behaviors and health outcomes. Food insecurity has been emerging as a priority for many food environment studies. Food deserts (defined as limited access to healthy foods) and food swamps (defined as overexposed to unhealthy foods) are two critical components of food insecurity.

Despite a lot of progress has been made in food environment studies, current retail food environment assessment mainly uses simply descriptive food assess measures, mostly overlooking the role of multiple transportation modes in food access, spatial associations between geographic food access and sociodemographic deprivation, as well as the variations of in-store characteristics across different food stores. This dissertation seeks to fill up the research gaps through pursuing three research objectives.

First, taking advantage of Geographic Information Science, neighborhood community nutrition environment was characterized using a proposed multiple-mode Huff-based 2SFCA method to measure geographic access to food outlets in Austin, Texas. The spatial accessibility score was calculated with a set of impedance coefficients ranging from 1.2 to 2.2. It shows an urban-core and peripheral disparity in terms of spatial accessibility; the spatial access to both healthy and unhealthy food outlets increase as the impedance coefficient increases. The proposed multiple-mode Huff-based 2SFCA was compared with its single-mode counterpart using a paired t-test. The comparison illustrates that the difference between the two methods on healthy food access was

significant at the impedance coefficient 1.4; for the difference in unhealthy food access, it was significant at the impedance coefficient 1.5.

Second, this dissertation research examined the relationship between geographic accessibility to food outlets and sociodemographic marginalization at the block group level. Eight sociodemographic deprivation variables were reduced to two indices: the Economic Deprivation Index (EDI) and Sociocultural Deprivation Index (SDI). Different from the research that uses conventional statistics, this dissertation used spatial statistics to adjust for spatial autocorrelation and spatial heterogeneity problems in the relationships between the two entities. I first employed the spatial autoregressive model to account for the spatial autocorrelation issue. The spatial lag model shows that block groups' EDI was significantly related to the access to healthy food (coefficient = -0.054, $p = 0.037$); spatial error model shows that SDI was significantly associated with the access to unhealthy food outlets ($p = 0.000$). This finding indicates that block groups in low economic deprivation (representing high economic status) enjoyed better spatial access to healthy foods, while those in high sociocultural deprivation (indicating low sociocultural status) were overexposed to unhealthy foods. I then used a semi-parametric Geographic Weighted Regression (GWR) model to explore spatial heterogeneity in the relationship between spatial food access, EDI, and SDI. The semi-parametric GWR allows the flexibility to incorporate both fixed and geographically varying explanatory variables, providing a more satisfactory model fit than the conventional GWR. The result shows that the EDI was a significant global predictor of healthy food access ($p = 0.000$) but was an insignificantly global predictor of unhealthy food access (p = 0.061); the SDI is a varying local variable to predict both healthy and unhealthy food access. Also, the spatial access to food outlets, EDI, and SDI were integrated to identify food deserts and food swamps in Austin. The use of hot spot analysis enables me to account for the spatial dependence issue that was ignored by previous studies. The result shows that food desert neighborhoods were mainly located in the eastern part of IH-35, and food swamp neighborhoods were in the northeast corner of Austin.

Finally, this research fills up a research gap that lacks studies examining the consumer nutrition environment in food-insecure (e.g., food desert and food swamp) and food-secure (e.g., food oasis) neighborhoods. It investigated consumer nutrition environment (i.e., food availability, food price, food quality, and labeling) in three selected neighborhoods (e.g., food desert, food swamp, and food oasis) in Austin, Texas. A food auditing instrument m-TxNEA-S was developed in this dissertation to capture the unique dietary culture and food preferences for Hispanic/Latino groups in Texas. Then this surveying tool was used to survey 14 grocery stores and 32 convenience stores in the three neighborhoods. It shows high inter-rater reliability (mean = 0.96) of the m-TxNEA-S. The result shows that there was a statistically significant interaction (p = 0.019) between the effects of store type (ST) and neighborhood nutrition environment (NNE) on healthy food availability. Food swamp and food oasis neighborhoods' grocery stores offered significantly more healthy foods than convenience stores. Grocery stores in the food swamp (p = 0.018) and food oasis neighborhoods (p = 0.015) had a significantly higher availability of healthy foods than those from the food desert neighborhood; convenience stores in the three neighborhoods did not exhibit any significant difference on healthy food availability ($p = 0.932$). It also shows that the prices for healthy items (lower fat, lower calorie, and whole grain) were not significantly different from their regular food options by ST ($p = 0.374$) and NNE ($p = 0.437$). The analysis of food quality shows that there was not any significant difference by ST ($p = 0.801$) and NNE ($p = 0.272$). It did not exhibit any significant difference in food labelling by ST ($p = 0.897$) and NNE ($p = 0.802$).

1 INTRODUCTION

Background

Food is essential in people's daily life, and it can provide the nutrients (i.e., protein, carbohydrate, fat, vitamins and minerals) and energy to sustain human's growth and development (D'Acosta 2015). Moreover, foods reflect cultural value and identity. Foods to people are not just something that can maintain their physical health status, but also can improve their mental and spiritual health (Behjat 2016). Besides, the World Health Organization (WHO) considers access to healthy and nutritious foods as a fundamental human right. WHO asserts that each individual gets equal access to healthy foods regardless of socio-economic status and that governments and agencies should make policies and implement them to ensure people's equal access to safe, nutritious, and affordable foods (Hodgson 2012).

Overweight and obesity

In Canada and America, the prevalence of obesity or overweight has tripled since 1975 (World Health Organization 2017). In 2000, more than 64 % of the population in the United States were considered either overweight or obese (Flegal, et al. 2002). According to Texas Health and Human Services1, the share of obese adults in Texas has increased dramatically. In 1995, only 15.9 % of Texas adults had BMIs indicating obesity; in 2010, 31.7% of adults fall into this category. The obesity rate of Texan adults was 33.87 % in 2016 had placed Texas 8th nationally, according to a report by Trust for

[1] https://www.dshs.texas.gov/Obesity/Data/

American Health[2]. Approximately 37.1 % of the residents in Austin, TX were overweight, and 27.0 % were obese in 2011[3].

Obesity is known to contribute to a number of health risks and diseases, including hypertension (Brown, Donato, and Obarzanek 1998; Stamler, et al. 1978), cardiovascular disease (Haffner, et al. 1991), type II diabetes (Chan, et al. 1994), stroke (Walker, et al. 1996), and some cancers (Bostick, et al. 1994; Chute, et al. 1991). As J. Michael McGinnis, the Deputy Assistant Secretary for Health concluded that the combination of dietary factors and sedentary activity patterns accounts for at least 300,000 deaths each year; obesity was a key contributor (McGinnis and Foege 1993). Moreover, obesity can lead to mental illness such as depression, anxiety and other mental disorders (Becker et al. 2001). Meanwhile, being obese could be seen as a sign of poor self-control; people who are obese or overweight may make a negative impression on others (Carr and Friedman 2005). The negative attitudes on obesity often lead to discrimination and prejudice in many occasions such as income, employment, college acceptance, and marriage (Judge and Cable 2011).

There are also considerable economic costs on medical care associated with obesity and overweight; as a result, obesity could cause heavy burdens to society and stakeholders. The economic costs of obesity in 1986 for the U.S. was $39.3 billion (Colditz 1992). In 1995 the total cost was $99.2 billion, and more than 50% ($51.6 billion) was spent on direct medical care costs with obesity-related diseases. The indirect

[2] https://stateofobesity.org/adult-obesity/
[3] http://www.governing.com/gov-data/obesity-rates-by-state-metro-area-data.html

cost of obesity-related illness was $47.6 billion, which approximates to the economic costs of smoking (Colditz 1992). In 2008, the cost was increased to $147 billion. It is estimated that the costs of obesity as part of the total health care budget would be approximately $ 344 billion in 2018, and it is likely to be higher than this amount due to the rapid increase in the prevalence of overweight and obesity[4].

Retail food environment matters to overweight and obesity

Individuals' genetics and personal behaviors cannot fully explain the elevating prevalence of overweight and obesity. More and more recent studies have linked obesity with environmental risk factors such as retail food environments (RFE) (Witten 2016). RFE is defined as "a group of factors including the types of retail food outlets and the availability, quality, and price of different kinds of foods, such as prepared foods, fresh produce, and other groceries, in a given geographical area" (Zenk, Schulz, and Odoms-Young 2009, 61). RFE is an indispensable part of the food environment[5]. RFE contains two food environments: food stores (such as grocery stores, convenience stores, and farmers' markets) and restaurants (Glanz, et al. 2005; Luan 2016; Leia Michelle Minaker 2013).

RFE affects people's dietary behaviors and health in two dimensions (Glanz, et al. 2005; Luan 2016). One is geographic/physical access to different kinds of food stores and

[4] https://www.fightchronicdisease.org/latest-news/new-data-shows-obesity-costs-will-grow-344-billion-2018

[5] Note: the food environment is a much broader concept than RFE. Food environment could be categorized as food store environment (e.g., supermarkets, grocery stores, specialty food stores, and farmers' markets), restaurant food environment (fast-food and full-service restaurants), school food environment (e.g., cafeterias, vending machines, and snack shops), and worksite food environment. The retail food environment contains the food store and restaurant food environment. The school and the worksite food environment are out of the scope of this research. The retail food environment (RFE) in this dissertation is synonymous to the food environment.

restaurants in neighborhoods. RFE is associated with dietary behaviors and health outcomes because people tend to consume more fruits and vegetables if they live near a grocery store (Morland, Roux, and Wing 2006). Access to food stores that sell affordable and nutritious food is necessary to maintain a healthful diet (Vallianatos, et al. 2010). Neighborhoods and communities that have inadequate access to food venues that are selling healthy items tend to present its residents with enormous challenges to maintain a nutritious diet and healthy weight status. Unhealthy food environment could result in many issues such as malnutrition and diseases that impair human well-being. For instance, eating at fast food restaurants has been linked to increased caloric intake, lower intake of fruits and vegetables and increased risk of being obese (Anderson, et al. 2011; Wilcox, et al. 2013). The locations of supermarket chains in suburban areas and convenience stores in urban city centers have played an essential role in shaping the unequal access to healthy and nutritious foods (Stein 2011).

The second dimension of RFE shaping individuals' food purchase behaviors and eating patterns is through consumers' experience in food stores and restaurants (Glanz, et al. 2005). It refers to if some kinds of foods (e.g. fruit and vegetables, bread, and milk) are available and affordable to consumers, and if the foods are in good quality. People make food choices based not only on personal preference but also on food availability, price, and quality (Vallianatos, et al. 2010). Food availability and price are influential factors in food purchase and consumption (Glanz, et al. 2005; Glanz, et al. 2007). It is reported that those who lived near grocery stores with higher availability of low fat and high fiber products tend to have healthier diets. Food cost is an important factor in food decisions (Glanz, et al. 2005). Food price affects people's purchase of fruits, vegetables, beef, and pork for low-income households. Beydoun, Powell, and Wang (2008) found that a high price for fast foods is linked to an increase in lower saturated fat intake and a healthier diet. Fruit and vegetable price are positively associated with improved dietary quality, a higher healthy eating index, and lower BMI.

The Issues of Food Desert and Food Swamp

A food desert is an issue that individuals have barriers accessing to nutritional and affordable foods in socially deprived areas (Stein 2011). Supermarkets and grocery stores are the primary food venues to offer a wide range of nutritious foods, which may directly or indirectly influence people's dietary behaviors and health (Stein 2011). The phenomenon food desert was created by the suburbanization of healthy food outlets, in retrospect to the 1950s when the suburbanization began to occur. Then in the 1970s and 1980s American cities were in "urban crisis," and urban areas became more impoverished since manufacturing jobs were lost and the middle class moved away from cities (Wilson 1996). Population and demographic changes in urban neighborhoods during this period also resulted in a significant loss of larger supermarkets and grocery stores. Middle-class households moved to the suburbs and brought food retailers into the suburbs, leaving low-income families near or in urban city core (Stein 2011). The suburbanization also leads to a prevalence of convenience stores and alcoholic beverages in the inner city. These stores may have a poor selection of healthy foods and a wide selection of unhealthy foods, which may contribute to poor diets. Less available of supermarkets, the increasing presence of convenience stores in low-income neighborhoods may create severe barriers to healthy eating.

The term “food desert” originates in the west of Scotland in the early 1990s. In 1995, a policy-working group for Low Income Project Team published a report and coined the term food deserts (Cummins and Macintyre 2002). Beaumont, et al. (1995) defined the food desert as an urban area where residents could not afford to purchase affordable and healthy diets. Since then, the term has been used increasingly by scholars and policymakers. However, the investigation of food deserts was mostly conducted in the United Kingdom between 1995 and 2003. Then more studies were found in the United States after 2003. For the first few years, the exploration of food deserts was mainly in urban areas. Furey, Strugnell, and McIlveen (2001) were the first to investigate the issue of a food desert in rural regions of Northern Ireland. In the United States, Blanchard and Lyson (2006) were the early scholars who have applied the concept of the food desert in less densely populated suburban and rural areas.

Meanwhile, another food access issue is arising with the rapid development of fast-food industry and food consumption in the U.S. This issue is termed “food swamp”, which refers to that people are overexposed to an unhealthy food source (Hager, et al. 2017). The increasing unhealthy food environment could result in many issues such as malnutrition and diseases that impair human well-being. Since the early 1970s, fast food outlets have increased dramatically with the prevalence of childhood obesity increasing rapidly; the number of fast food outlets increased from 30,000 in 1970 to 222,000 in the late 1990s (French, Harnack, and Jeffery 2000). Eating at fast food restaurants has been linked to increased caloric intake, lower intake of fruits and vegetables and increased risk for obesity (Wilcox, et al. 2013; Anderson, et al. 2011).

The term food swamp was introduced in 2009. Rose, et al. (2009) articulated that the limited access to healthy foods is a useful metaphor for under-nutrition. However, the issue in developed countries such as the U.S. is over-consumption of nutrition. The overwhelming unhealthy foods in low-income neighborhoods is a more severe problem than the issue food desert. “Food swamp” was used to describe the areas where high calorie and energy dense food swamps out healthy foods. The authors suggested that food swamp is a more useful metaphor than food desert to depict neighborhood food environment in the U.S. and other developed countries.

Various definitions of food deserts and food swamps

The food deserts literature addresses many ambiguities regarding the actual definition of food deserts, and specific definitions vary in different studies. For instance, it is defined as urban food deserts as "those areas of inner cities where cheap, nutritious food is virtually unobtainable. Car-less residents, unable to reach out-of-town supermarkets, depending on the corner shop where prices are high, products are processed, and fresh fruit and vegetables are poor or non-existent"(Behjat 2016, 5). Donkin, et al. (1999) defined the food desert as areas that are more than a critical distance away from a food store. Wrigley, et al. (2002) elaborated on the definition and included food provisions such as food availability, variety, and price. In the 2008 USDA Farm Bill defined food desert as an “area in the United States with limited access to affordable and nutritious food, particularly such an area composed of predominantly lower income neighborhoods and communities” (Hodgson 2012, 15). Coveney and O’Dwyer (2009) considered areas as food desert when the census collector district had a low percentage of car ownership and the residence was more than 2, 500m from the nearest supermarkets. Morton and Blanchard (2007) identified several characteristics in food desert areas: a high percentage of people without holding a high school degree, high poverty rates, low incomes, and large portions of families/households living in rural areas.

Similarly, there is no formal definition of food swamp. Rose, et al. (2009) defined food swamp as "large relative amount of energy-dense snack foods, inundate healthy food options." Nevertheless, this definition is problematic because it ignores non-spatial factors such as social-deprivation onto food access. Convenience stores and fast food restaurants are two critical resources providing unhealthy food options. It is found that low-income groups are more likely to live close to unhealthy food retailers. For example, Moore and Diez Roux (2006) explored types of food stores in selected census tracts in North Carolina, Maryland, and New York, and the result showed that convenience stores and small liquor stores are located disproportionately in low-income neighborhoods. Fleischhacker, et al. (2011) conducted a systematic review of fast food access research; 21 studies explore the linkages between fast food access and socioeconomic status and 15 studies found low-income areas had more fast food restaurants than middle- and higher-income communities. There is evidence that ethnic minorities groups in comparison with Caucasians were more likely to live in areas with high access to fast food restaurants. Fleischhacker, et al. (2011) asserted that 10 out of 12 studies found that fast food restaurants are more likely to be in areas with higher concentration of minorities, especially for African American and Hispanic people. Other studies also found a similar pattern (Block, Scribner, and DeSalvo 2004; Hager, et al. 2017).

Various measurements of food deserts and food swamps

There are extensive explorations of the food desert in the U.S and other countries during the past 15 years. There are no standard variables and methodologies to measure food desert and food swamp. The variation in approaches creates inconsistency and ambiguity in their results. The lack of consensus on food desert definition has led to that different researchers develop different methodologies and terminologies in their studies.

Some researchers identified food deserts without considering the distribution of food stores; instead, they used the quality and the cost of food as the criteria (Cummins and Macintyre 2002; Larsen and Gilliland 2008). Whereas, other studies considered the type and size of the food stores and the sales volumes to determine the food desert (Hendrickson, Smith, and Eikenberry 2006). Some studies measured food desert considering the access to food stores; of these methods, some only measured spatial accessibility to food stores. The employed measures are often descriptive approaches such as density, proximity, and a variety of food stores. Different buffer distance was to define what is accessible and inaccessible. For example, Coombs, et al. (2010) created a 1-mile buffer around grocery stores. Any neighborhood beyond this buffer were potential food deserts. It assumed that healthy and nutritious foods could only be found in grocery stores and any areas beyond the grocery store buffer zone are food deserts. Schlundt (2014) measured food access in Nashville, Tennessee. He created a 0.5-mile buffer around grocery stores and calculated the distance from each parcel to the nearest grocery store and the nearest bus stop. Then each census block was scored based on the distance. A Food Desert Score was created as an index to identify neighborhoods that may be considered food deserts. He identified four areas as potential food deserts.

There are some studies to combine socio-demographic variables with spatial access to food stores when defining food deserts. They can be divided into two categories. The first category uses individual economic or socio-cultural variables to identify food deserts. For instance, Smoyer - Tomic, Spence, and Amrhein (2006) calculated the spatial accessibility to grocery stores in Edmonton, Alberta. Two methods were employed: proximity (shortest path) and density (number of grocery stores within a 1000-meter network buffer around the centroids of each postal zone). They identified food deserts as the follows: 1) neighborhoods where access to grocery stores falls in the lowest quartile of the study area; 2) residents that are vulnerable groups such as the highest quartile of low income, no car ownership, and old population. A similar method was used by (O'Dwyer and Coveney (2006). The second category is to select a series of socio-demographic variables to construct a composite index (Matheson, et al. 2012). Some researchers developed a deprivation index score based on a set of socioeconomic and demographic variables. A well-cited and classic study is Apparicio, Cloutier, and Shearmur (2007). They developed a social deprivation index based on five variables: low-income population (%), lone-parent families (%), unemployment rate (%), adults with a low level of schooling (%), and recent immigrants (%). Spatial access to supermarkets is based on proximity, density, and competition. Then they combined the spatial access measure and social deprivation index to identify food deserts. The similar approach could be found in (Gustafson, Hankins, and Jilcott (2012) and Larsen and Gilliland (2008).

Compared to the extensive exploration of the food desert, food swamp is relatively rarely studied. The measure of food swamp often uses relative one: unhealthy foods relative to healthy foods. Food swamp is measured by the Retail Food Environment Index (RFEI) (Spence, et al. 2009; Woodham 2011; Luan 2016). It is a ratio of the numbers of unhealthy food stores to the number of retailers. Unhealthy food stores often refer to convenience stores and fast food restaurants, while healthy food stores usually contain supermarkets, grocery stores, specialty stores, and farmers' markets. The RFEI was firstly used by Advocacy California Center for Public Health (2007) at the census tract level. Spence, et al. (2009) calculated RFEI by creating 800 m and 1600 m buffers around survey respondents' homes in Edmonton Canada and found that a higher REFI was associated with a higher risk of being obese. Woodham (2011) calculated REFI for each census area unit and found that Eastern Porirua had higher REFI than Whitby in New Zealand. Gallagher (2006) and Gallagher (2007) studied food swamp issue in Chicago and Detroit by calculating the average distance from unhealthy food venues and the average distance to healthy food outlets to create a Food Swamp Score; this is an alternative way to compute the RFEI. Luan (2016) used a modified REFI (a ratio of the number of healthy food stores to the number of all food stores) to represent food swamp. The result showed that in 2011 - 2014, food swamp became more prevalent than the food desert in the Region of Waterloo, Canada.

Food Deserts, Food Swamps, and Food Insecurity

Food deserts and food swamp are important components of food insecurity

Food insecurity is now a priority for governments and authorities to fight against in both developing and developed countries[6]. Food insecurity is defined as "the social and

[6] http://www.fao.org/docrep/014/i2330e/i2330e.pdf

economic problem of lack of food due to a resource or other constraints, not voluntary fasting or dieting, or because of illness, or for other reasons" (National Research Council 2006, 44). They assert that individuals or households are experiencing food insecurity when: "(1) uncertainty about future food availability and access, (2) insufficiency in the amount and kind of food required for a healthy lifestyle, or (3) the need to use socially unacceptable ways to acquire food" (National Research Council 2006, 44). The USDA divides food insecurity into two levels[7]: low food security (food insecurity without hunger) and very low food security (food insecurity with hunger). Both levels of food insecurity could lead to poor food choices and malnutrition.

With the comparison of definitions of food insecurity and food desert and food swamp, the common word is “access”. The three concepts have some common characteristics. However, they have differences: food desert (or food swamp) focuses more on the physical constraint in obtaining foods, while food insecurity goes beyond the spatial accessibility and emphasizes the economic or financial constraints, such as income, social networks, governmental assistance, housing provisions, and allocation (Behjat 2016). In other words, food insecurity is much more profound in scope than food deserts and food swamps. Behjat (2016) believed that whether the food desert is a strong predictor of food insecurity depends on the resources a community possess. For residents who live in a community where other resources are scarce; they have to entirely rely on the local food retail systems to procure foods. Then food swamp is significantly related to

[7] https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/definitions-of-food-security/

food insecurity (Behjat 2016). In contrast, in communities where other resources are abundant, people do not necessarily have to obtain foods from supermarkets and grocery stores, and they can be food secured even if they have limited access to these stores. In this case, the food desert does not predict food insecurity (Behjat 2016).

Food desert, food swamp, food access, and food insecurity are being studied interchangeably; they are used to characterize the food environment that people reside. Food desert and food swamp are two critical components of food insecurity. Food desert and food swamp are useful to identify areas where residents are potentially suffering from food insecurity when social and economic resources are restricted (Behjat 2016). Therefore, studying food desert (food swamp) could contribute to the understanding of food insecurity in certain areas. It is a useful starting point to develop an appropriate methodology in identifying food desert and food swamp to further understand the potential risks of food insecurity in some geographic regions.

Food insecurity issue in Austin, Texas

Texas is rated as one of the top three states that have food insecurity higher than the national average[8]. The study area of this dissertation, Austin, TX is facing food insecurity issue, especially in the east side of Austin (Pedraza Sanchez 2015). Austin is located in central Texas; the capital of Texas, and it is one of the fastest growing cities in the US (Pedraza Sanchez 2015). Nearly 25% of its residents live in urban food deserts and insecurity. The past two decades Austin has witnessed a rapid population growth. This fast-growing population trend makes properties, and housing values increase rapidly as

[8] https://www.baptiststandard.com/news/texas/food-insecurity-declines-but-1-5-million-texas-households-still-at-risk/

well. The increases in living costs force people to spend more money on rent rather than buying nutritious food. This phenomenon is more severe in the East part of Austin. According to the 2011 Sustainable Food Center (SFC) report, the values of a single-family home in 78721 zip code has increased by 80% from 2000 to 2005 (Sustainable Food Center 2011). Moreover, East Austin is the most impoverished area in Travis County, and 40% of residents live below the federal poverty level (Sustainable Food Center 1996). On the one hand, households in this area are not able to acquire enough food to meet their needs because they do not have sufficient money to afford food. On the other hand, food stores in East Austin are generally limited not only in counts, but also in size, and they do not provide various food options. Large and healthy food outlets are still concentrated in wealthy neighborhoods on the West of Austin (Sustainable Food Center 1996); Not mention that it lacks reliable public transportation system in East Austin, making low-income residents difficult to commute for grocery shopping.

Austin has some initiatives and efforts to deal with the food insecurity, especially for the issue of the food desert. Austin's City Council passed resolution 20160303-020 on March 3, 2016; three months later, the City of Austin's Office of Sustainability has collaborated with other five city departments to develop a program called Healthy Food Access Initiative[9]. Its goals are to enlarge healthy food options in food retailers, increase fresh and organic foods in farms and community gardens, and enrich people's awareness of healthy eating. The initiative also produced six recommendations to cope with the food insecurity in Austin. These include (1) conducting a comprehensive analysis on food

[9] https://www.nycfoodpolicy.org/austin-texas-seeks-make-healthy-food-available-everyone/

environment in Austin and creating maps of food insecurity; (2) establishing grants to provide financial support for various food providers in low-income neighborhoods; (3) simplifying the application of urban farms and community gardens; (4) empowering low-income groups' purchasing power by piloting a new Nutritious Food Incentives program (5) enhancing residents' awareness on healthy eating by developing SNAP outreach pilot program and outreach campaign; and (6) creating safe routes to markets in collaboration with other departments in Austin. An $ 800,000 budget was distributed to implement the initiative in 2017. Moreover, the Sustainable Food Center's Double Dollar Incentive Program was funded by the City of Austin to double the benefits of SNAP and WIC produce. In addition to the government, some no-profit organizations were established to strengthen local access to healthy diets. A prominent one is the Sustainable Food Center (SFC); it hosts four different farmers' markets in underserved communities of Austin and offers incentive programs for low-income residents. It also opens free healthy cooking classes (such as “The Happy Kitchen”) to enhance people's healthy cooking methods. Meanwhile, it is operating a program named Grow Local, aiming at encouraging residents to plant their foods. More importantly, SFC reported the food insecurity issue people in East Austin were facing, especially for the lack of public transportation. In response to the report, the Austin/Travis County Food Policy Council in conjunction with the Austin’s Capital Metro transit system created “The Grocery Bus," which runs in East Austin to give a ride for residents from impoverished communities and neighborhoods to others. There are other initiatives in Austin, and one can refer to these websites for more information.

Identification of food desert and food swamp in Austin, TX

The USDA developed a “food desert locator” tool to map food deserts at census tract level across the United States[10]. It uses two criteria: low-income and low-access. The low-income census tracts are defined as 20% or higher poverty rate or a median family income at (or below) 80% of the area’s median family income. The low-access census tracts are at least 33% of the population who reside more than one mile from a supermarket (or large grocery store) in urban areas and 10 miles in rural areas. The use of "food desert locator" can identify food deserts in Austin, TX. However, this method uses 1 mile or 10 miles as the cutting point to measure the proximity to food stores; the cut-off values are controversial. Besides, it only considers one socio-demographic variable (i.e., income), which cannot fully capture the picture of social and economic constraints on food access.

The Office of Sustainability in the City of Austin has conducted food environment analysis in Austin. They used a scoring tool, called the Healthy Food Availability Index. The survey and research methodology were adapted from the Johns Hopkins Center for a Livable Future[11]. Seven data collectors visited the stores and objectively assessed the types and quantities of food available. This information was used to create a score for each store that is based on the availability of foods in the following categories: protein, dairy, produce, and grains. Healthy Food Availability Index (HFAI) scores were averaged at the block group, and a block group that did not meet a specific average rating was

[10] https://www.ers.usda.gov/data-products/food-access-research-atlas/go-to-the-atlas.aspx
[11] https://www.jhsph.edu/research/centers-and-institutes/johns-hopkins-center-for-a-livable-future/about/index.html

considered not healthy. These areas were overlaid on a map with information on household income, physical proximity to healthy food stores (0.25 mile in urban areas and 1-mile in rural areas from healthy food retail), and vehicle availability. Any place where all four variables overlapped completely was identified as Healthy Food Priority Areas (or food deserts). This method considers the vehicle ownership in addition to the income. Most importantly, it constructs the HFAI score, which measures healthy food status for a census block group. The project sets out to collect data in early 2017 from over 900 food retail establishments to assess the availability of nutritious food throughout the Austin area. Even though it may reveal a more accurate picture of healthy food availability and accessibility, it requires a lot of effort and financial support to recruit surveyors and conduct the survey. The project, which has been ongoing for one and a half year, is far from done. Therefore, there is a need to develop a cost-efficient but a practical methodology to explore the issues of the food desert and food swamp in Austin, TX.

Limitations in Retail Food Environment Research

Recently built environment and its relationship with health outcomes gained a lot of attention to scholars (Booth, Pinkston, and Poston 2005; Feng, et al. 2010; Papas, et al. 2007). Retail food environment is part of the built environment. Therefore, the investigation of this issue fits well in this topic. This dissertation focuses on identifying food deserts and food swamps, as well as examining the consumer nutrition environment in different communities. Even though scholars have made much achievement, there are some limitations in the existing food studies. I summarize them below.

(1) Current food desert and food swamp studies utilize descriptive methods such as density, proximity, and variety to measure physical food accessibility within administrative boundaries or in buffer areas (Blok, Scribner, and DeSalvo 2004; Gordon, et al. 2011; Powell, et al. 2007; Apparicio, Cloutier, and Shearmur 2007). However, these methods do not account for customers' choice of food stores nor the population's share of food store capacity.

(2) Research measuring food accessibility often assume that people are traveling to food outlets by a single transportation mode (usually by automobile) (Mao and Nekorchuk 2013; Kuai and Zhao 2017). This assumption is biased because there are low-income and minority groups that do not have private vehicles (Mao and Nekorchuk 2013). Walking or taking public transportation to food outlets would be the only two options for disadvantaged people (Larsen and Gilliland 2008). We should not neglect these marginalized groups. The mobility for these groups is limited, which may force them to shop in convenience stores or eat in fast food restaurants, which might put them into a long-term exposure of toxic or unsupportive environments. To date, few food access studies have integrated different transportation modes into the measurement.

(3) Previous research used traditional (or classic) statistical techniques such as Pearson correlation or multivariate regression to examine the relationships between physical accessibility to food outlets and socio-demographic characteristics of neighborhoods. These methods neither control for clustering of similar values over space (spatial autocorrelation) nor the heterogeneity of local spatial variations (or spatial non-stationarity) in statistical relationships between variables. Therefore, it is necessary to use

spatial methods to uncover the connections between deprivation and the built environment.

(4) Various studies explore the consumer nutrition environment (food availability, food affordability, and food quality) in different communities across geographic space. These communities are often selected from two socio-economically contrasting ones such as low-income vs. high-income neighborhoods. Few studies have explored the consumer nutrition environment in food deserts and food swamps, as well as how this environment differs in food-insecure neighborhoods (such as food deserts and food swamps) and food-secured neighborhoods (food oases).

Study Objectives and Research Questions

There are two objectives for this dissertation research. Aligning with these research objectives are a series of research questions.

Objective 1: To identify and map out food deserts and food swamps in Austin, TX

The identification of food deserts and food swamps is comprised of three dimensions: physical, economic, and socio-cultural. A novel **multiple-modal Huff-based 2SFCA** method is proposed to measure geographic/physical access to food outlets. Then spatial techniques are used to associate economic (or socio-cultural) factors with physical access for the identification of food desert and food swamp. One general research question and three sub-questions are proposed.

<u>Q1. Where are food deserts and food swamps located in the city of Austin?</u>

<u>*1(a): What are the spatial disparities of geographic food accessibility in Austin, TX?*</u>

This question is about the geographic access to food stores and restaurants. The access to food outlets includes the access to healthy and unhealthy food outlets. A novel **Multi-Mode Huff-based 2SFCA** method is proposed to measure geographic access to both healthy and unhealthy food outlets.

1(b): Is there a significant association between economic (or socio-cultural) deprivation and access to (1) healthy food outlets, (2) unhealthy food outlets, in Austin, TX?

Sociodemographic deprivation could impose adverse effects on the access to food outlets. It would be meaningful to examine whether the access to healthy and unhealthy food outlets is stratified by different levels of sociodemographic deprivation. This question aims to use spatial regression method (i.e., spatial autoregressive method) to explore the relationship between spatial food access and deprivation. The control of spatial autocorrelation in spatial regression could reveal a true relationship between the two entities.

1(c): How does the relationship between economic (and socio-cultural) deprivation and access to (1) healthy food outlets, (2) unhealthy food outlets vary geographically in Austin, TX?

The question in 1(b) serves as a general exploration of spatial food access and sociodemographic deprivation. However, it cannot reveal the varying relationship between the two entities. Thus, this question is to examine the spatial heterogeneity (or non-stationarity) of relationships between food access and marginalization. I aim to undercover their spatial varying relationships in the study area. In addition, I identify food deserts and food swamps by considering both spatial food access and sociodemographic marginalization. The hypotheses are: (1) low access to healthy food outlets is spatially correlated with high social deprivation; (2) high access to unhealthy food outlets is spatially correlated with high sociodemographic deprivation. Based on the hypotheses, spatial food access, economic and socio-cultural deprivation can be integrated to delineate food deserts and food swamps via hot spot analysis.

Objective 2: To investigate the consumer nutrition environment in the selected neighborhoods in Austin, TX

The consumer nutrition environment includes in-store characteristics such as food availability, food price, and food quality. I intend to compare food availability, food price, and food quality in three neighborhoods in Austin, TX. One general research question is:

<u>Q2. How does the consumer nutrition environment vary in three communities that are from food deserts, food swamps, and food oases respectively in Austin, TX?</u>

The three communities should have different characteristics and are selected from the neighborhoods that are classified as food deserts, food swamps, and food oases (defined as high access to healthy food in affluent neighborhoods, opposite to the concept of food deserts). An in-store food audit is conducted to investigate this issue. This part could be broken down into four sub-questions:

<u>*2a: How does the availability of food in food stores vary across food deserts, food swamps, and food oases in Austin, TX?*</u>

<u>*2b: How does the price of the same food products in food stores vary across food deserts, food swamps, and food oases in Austin, TX?*</u>

2c: How does the quality of foods in food stores vary across food deserts, food swamps, and food oases in Austin, TX?

2d: How does the labelling of foods in food stores vary across food deserts, food swamps, and food oases in Austin, TX?

Structure of the Dissertation

This dissertation uses a socio-ecological approach to conceptualize RFE. The remaining chapters are organized as below.

Chapter Two reviews conceptual frameworks directing this dissertation. Three frameworks related to this dissertation are discussed. Empirical studies are reviewed, and research issues are addressed.

Chapter Three provides the theoretical framework of this proposed study. A diagram is present to depict the theories that direct this work.

Chapter Four proposes a novel method by integrating multiple transportation modes into the measurements. Its advantages are emphasized in this chapter.

Chapter Five presents the spatial methods to explore the relationship between spatial access to food outlets and socioeconomic status. Two models are used: the spatial autoregressive model and semi-parametric GWR. Lastly, I present a new method to identify food deserts and food swamps in Austin.

Chapter Six describes a food survey that was conducted in three selected neighborhoods in Austin, Texas to investigate the consumer nutrition environment by store type and neighborhood nutrition environment.

Chapter Seven summarizes the major finding and contribution. It also discusses the limitations of this research and proposes future directions.

2 LITERATURE REVIEW

This chapter reviews the relevant literature for this dissertation. It begins with an overview of theoretical perspectives on how the food environment and health status are related. Next, I review the empirical studies regarding the retail food environment. The review concentrates less on individual epidemiological studies and more on general assumptions and findings.

Food and Health: Shifting to Social Ecological-based Perspective

Although the elevating prevalence of obesity has been well documented, explanations for this emerging epidemic proves to be elusive. From an individual perspective, a gene is a non-modifiable factor that might explain individuals' susceptibility to overweight and obesity. Epidemiological research has attempted to uncover a genetic basis for the population with overweight and obesity; however, to date, a tiny proportion of the population can verify this finding (Han, Lawlor, and Kimm 2010; Strauss 2002). Frayling, et al. (2007) found that people who inherit FTO gene weighed seven pounds more than those without this gene. In spite of this, the authors articulated that inheriting a particular gene will not necessarily make anyone fat; people with this gene may be skinny unless they overeat or do not do any exercise (Piontak 2013). Han, Lawlor, and Kimm (2010) found that some genetic abnormalities affect obesity; however, these do not explain the sharp increases in the global prevalence of obesity in such a short time period; Moreover, Power and Schulkin (2008) showed that less than 5% of obese people have an identified genetic abnormality.

It seems more plausible to blame the modifiable factors (individuals' eating behaviors and physical activities) for obesity. Obesity occurs with an imbalance between the quantity of energy intake and the amount of expenditure. Specifically, a person becomes overweight or obese when he/she intakes more calories than he/she consumes. If individuals want their weights to go back to a healthy state, they have to make caloric expenditure more than caloric intake (lose weight). It is an effective manner to maintain or control weights is reducing caloric intake through eating less energy-dense foods.

Nevertheless, due to the vast technological and economic changes in the contemporary era, individuals' dietary behaviors tend to shift to caloric-dense fast foods, sweeter and fattier foods, sweetened beverages, and larger portion sizes. The dietary changes have caused people fatter than before. Besides, only emphasizing on individual behaviors separates people from their living environment, which ignores the physical and social features in the environment that shape personal eating behaviors (Leia Michelle Minaker 2013). Moreover, from the perspective of intervention, the exclusive focusing on individual behaviors is not beneficial for intervention because an individual has unique eating behaviors. There must be abundant of individual-level interventions, which are difficult to implement for the population as a whole (Rose, et al. 2009; Huang and Glass 2008).

The explanations above mainly focus on biological factor (gene) and individual behaviors, but they tend to overlook the influences of "Environment" (Leia Michelle Minaker 2013). The past two decades have witnessed rapid and widespread changes in the environment; the rising obesity issue reflects these changes. These environmental or contextual modifications have created an environment that is promoting weight gain and is not conducive to weight loss (Witten 2016). This environment is known as an obesogenic environment. The ANGELO is one of the earliest frameworks to conceptualize obesogenic environment. It is an acronym of Analysis Grid for Environments Linked to Obesity and was firstly proposed by Swinburn, Egger, and Raza (1999) in New Zealand. The ANGELO framework has categorized obesogenic (food and physical activity (PA)) environment into two settings: micro and macro environment.

Regarding the microenvironment, it includes multiple local environments such as home, work, schools, and neighborhoods whereas the macro environment defines a broader context, consisting of education and health systems, food industry, all levels of governments, transportation systems, and society's attitude and beliefs. Meanwhile, the ANGELO dissects obesogenic environment into four major types: physical (what is available?), economic (what are the costs?), political (what are the rules?), and sociocultural (what are the attitudes and beliefs?). Table 2.1 depicts the ANGELO framework; examples of different types of food environments are listed. The ANGELO does not seek to propose a mechanism by which the environment influences behavior. Instead, it attempts to identify and categorize all obesogenic features in the environment (Woodham 2011). Therefore, it can be used for stakeholders and practitioners to define a set of prioritized interventions of obesity (Swinburn, Egger, and Raza 1999). For instance, APPLE (A Pilot Programme for Lifestyle and Exercise) program in New Zealand used ANGELO framework to identify the barriers to healthy eating in children aged 5-12 years (Williden, et al. 2006). They found that the significant physical barrier is the lack of availability of convenient, healthy food options, and the economic barrier is the cost of healthy food. The sociocultural barriers include parents' knowledge, children's preference for a less healthy diet, and the lack of fruit and vegetable advertising. The political barriers consist of the absence of parental rules regarding purchasing less healthy food options (Williden, et al. 2006).

Table 2.1 The ANGELO framework.

Environment Size	Environment Type			
	Physical Food and PA	**Economic Food and PA**	**Political Food and PA**	**Sociocultural Food and PA**
Micro settings				
Neighborhoods	Grocery stores and convenience stores	Cost of food in retail stores		Social norms and beliefs on healthy eating
Schools	Cafeteria, vending machines, snack shop	Cost of cafeteria and snack food	Policy on school food programs	Teachers' attitude on healthy eating
Worksite	Vending machines, snack shop	Cost of food in vending machines		
Macro sectors				
Transportation systems	Availability of public transit and stops	Monetary incentives, subsides, tax for food		
Health regulatory systems		Policy on food product and labeling		

Note: It is derived based on Swinburn, Egger, and Raza (1999).

The ANGELO emphasizes the importance of environmental factors or features on health outcomes. However, it neglects the fact that an individual in a given environment has different dietary habitats and patterns. Only focusing on the influences of environment on health outcomes leaves out individuals’ capability of changing their dietary behaviors to resist obesogenic environment or to adapt to leptogenic environment12 (Leia Michelle Minaker 2013). Therefore, a framework that can combine both individual and environmental factors is needed to understand the pathway of foods on health. This kind of framework is known as the Social-Ecological Model. Richard, Gauvin, and Raine (2011) defines the Social-Ecological Model as “a formalized conceptualization of the individual and environmental determinants of health behaviors and public health outcomes." It recognizes that individuals’ behaviors are influenced by multilevel contexts (Story, et al. 2008; Robinson 2008). The Social-Ecological Model, firstly developed by American physiologist Urie Brofenbrenner, explains how the inherent qualities of a child and his/her environments influence his/her behaviors (Greenfield 2012; Grzywacz and Marks 2001; Lounsbury and Mitchell 2009). NAPO (Nutrition and Physical Activity Program to Prevent Obesity and Other Chronic Diseases) adopted this framework and tailored it to address the multiple levels of influences on overweight and obesity (Brown 2011; Hamre, et al. 2014).

The Social-Ecological Model used by the NAPO has five levels: individual, interpersonal, organizational, community, and society (Brown 2011; Hamre, et al. 2014). The first one is an individual level, and it refers to a person’s knowledge, attitudes, and beliefs and behaviors towards diet and exercises, which may increase the likelihood of being overweight and obesity (Sarrafzadegan, et al. 2013). The interaction between children and their physician or dietician is considered at this level. The interpersonal level includes any group of people who share a relationship. Family members and friends

[12] Leptogenic environment, opposite to Obesogenic environment, refers to an environment that is conducive to healthy weights through the promotion of healthy food choices and participation in physical activities. In short, Obesogenic factors could be reviewed as barriers to healthy weights, whereas leptogenic factors are the enhancers to maintain health status.

influence individual behaviors can be classified into this level (Brown 2011). The organization level includes organizational policies (e.g., school policy) and membership rules that can reinforce positive or negative behaviors on nutrition, diet, and physical activity (Control and Prevention 2009). The community or neighborhood level corresponds to any feature in obesogenic environments that can affect people's dietary and physical activity behaviors. It not only contains built (physical) environment such as neighborhood design, proximity to schools, parks, walking paths, traffic, the location of fast food and recreation centers, but also includes social environment such as social coherence, social support, and crime and neighborhood safety. The societal level refers to any policy and operation that is implemented as a nation or state level (Brown 2011; Leia Michelle Minaker 2013). These policies and campaigns may create healthier environments that promote energy balance for people. The examples include national media campaigns, wellness legislation, and federally and state-mandated school wellness policies. The five levels of SEM are inter-correlated and thus are often depicted as nesting dolls (Story, et al. 2008; Leia Michelle Minaker 2013; Brown 2011).

Later researchers are motivated by the Social-Ecological Model and have conceptualized better frameworks that incorporate personal behaviors and environmental factors. One of them is ENRG (Environmental Research framework for weight Gain prevention), which were proposed by (Kremers, et al. 2006). This framework is built upon ANGELO and attempts to outline the causal mechanisms between environmental influence and energy balance-related behaviors. Cognitive mediators and moderators between ecological factors and behaviors are added in the framework. Cognitive mediators include attitudes, subjective norms, perceived behavioral control, and intention. Moderators contain person moderator (i.e., demographic, personality, awareness and involvement) and behavior moderator (such as habit strength and clustering). The influences of environment on energy balance-related behaviors are moderated and mediated by these factors. The ENRG framework may guide causal mechanisms to link environmental features with personal behaviors in different populations.

In all, the Social-Ecological Model serves as a starting point for scholars to understand different levels of exposures and all levels of risk factors that influence overweight and obesity. It suggests that health is the result of the interactions between people and their environment, and people's behaviors cannot be separated from their environment. In addition, the Social-Ecological Model is beneficial for the creation of intervention and prevention programs to tackle the issue of overweight and obesity since they can be operated at multiple levels (Glanz, et al. 2005; Leia Michelle Minaker 2013; Luan 2016; L Minaker 2013).

Conceptual Frameworks of Retail Food Environment

The Social-Ecological Model has gained great popularities over the past two decades (Robinson 2008; Greenfield 2012; Leia Michelle Minaker 2013; Luan 2016; Ball, Crawford, and Mishra 2006). Researchers have acknowledged that dietary behavior and health outcome should be understood through a socio-ecological perspective at five different levels. Among these levels, the community (micro-environment) level such as neighborhood recently has received more attention as more and more evidence reveal that neighborhood-level variables could be potential determinants of health behaviors, especially in disadvantaged areas (Luan 2016; Leia Michelle Minaker 2013; L Minaker 2013; Stein 2011). Meanwhile, within-store factors such as food availability, food price, food quality, and promotions may create a micro-environment to affect people's dietary behaviors. This dissertation aims to explore the characteristics of the retail food environment, and its association with health outcomes. One of the conceptual frameworks --- Glanz and colleagues' nutrition environment framework stands out since it has been widely adopted from many researchers in multiple fields (Glanz, et al. 2005; Luan 2016; Leia Michelle Minaker 2013; L Minaker 2013).

As shown in Figure 2.1, three sets of environmental variables are proposed to affect dietary behaviors or health outcomes potentially; they are the community, consumer, and organizational nutrition environment (Glanz, et al. 2005; Witten 2016). The community nutrition environment contains neighborhood characteristics such as type, number, location, and accessibility of food stores and restaurants (Glanz, et al. 2005). It is commonly evaluated by geographic access measures. The consumer nutrition environment refers to in-store characteristics that consumers encounter when they reach a food retailer (Glanz, et al. 2005; Leia Michelle Minaker 2013); these characteristics typically include food availability, affordability, and quality. Portion sizes, promotions, snack foods near counters might be included as well (Glanz, et al. 2005; Saelens, et al. 2007; Ball, Timperio, and Crawford 2006). The organizational nutrition environment involves multiple contexts including home, work, school, and neighborhoods (Glanz, et al. 2005). Multiple-Contexts environmental exposure occurs in daily life. For example, people may travel far away from their residential locations (home) to seek special ethnic foods (Witten 2016). Other factors that affect dietary behaviors in the framework contain government and industry policies, the information environment (i.e., media and advertising), and individual variables (i.e., socio-demographics, psychosocial factors, and perceived nutrition environment) (Glanz, et al. 2005; Witten 2016).

This framework is quite useful for food environment assessment and the evaluation of the relationships between RFE and health outcome such as obesity (Glanz, et al. 2005; Leia Michelle Minaker 2013; Luan 2016). The authors highlight the importance of studying the community nutrition environment and the consumer nutrition environment because a large number of variables could be measured in these two environments (Glanz, et al. 2005). Moreover, the framework proposes two pathways that food environment affects dietary behaviors. One is a direct effect. The other one is an indirect effect; environmental food variables on eating behaviors or health outcomes are moderated or mediated by individual variables (Glanz, et al. 2005; Witten 2016). These two pathways could serve as theoretical foundations of food environment as a determinant of diet and health (Leia Michelle Minaker 2013; L Minaker 2013; Luan 2016).

From the food equity and social justice perspective, only consideration of geographic/physical access to food outlets might create bias in evaluating retail food environment (Luan 2016). Certain groups who are in disadvantaged socioeconomic status (i.e., low-income people and racial/ethnic minorities) have reduced healthy food access or elevated unhealthy food access (Apparicio, Cloutier, and Shearmur 2007; Blok, Scribner, and DeSalvo 2004). It can be explained by Macintyre's deprivation amplification hypothesis (Macintyre 2007). The core concept of this hypothesis is "the disadvantageous environments magnify individual vulnerability, resulting in (built-) environmental characteristics more detrimental to health in deprived areas" (Luan 2016, 20). This hypothesis emphasizes the amplification of disadvantaged social environment or context on individuals or household deprivation. That is to say, viduals who are socially disadvantaged individuals are further exposed to the contextual disadvantage in terms of access to affordable, nutritious food or physical facilities (Macintyre 2007; Macintyre, Macdonald, and Ellaway 2008).

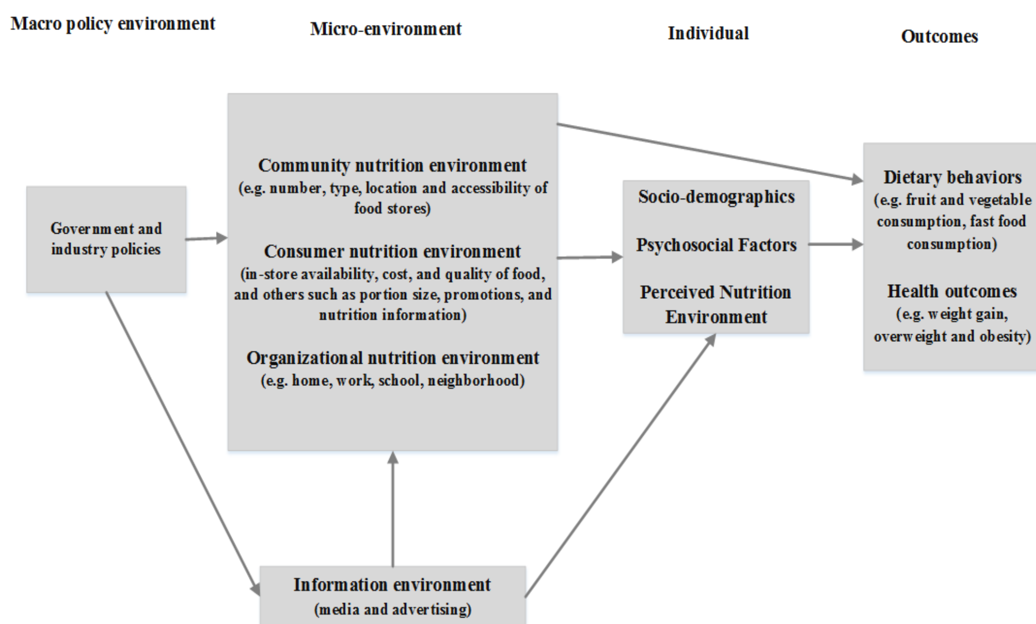


Figure 2.1 Glanz and colleagues' conceptual framework of nutrition environment.
Note: It is adopted from Glanz, et al. (2005, 331).

On top of the deprivation amplification hypothesis, Lytle's ecological model of individuals' eating behaviors provides another theoretical basis of social marginalization and built environment (Luan 2016; Leia Michelle Minaker 2013). This model considers how individual, environmental, and social factors explain differences in eating behaviors. Lytle (2009) asserts that the environmental factors explain more variance of eating behaviors with individual and social factors becoming more restricted (e.g., low social coherence and support). As social factors become less restricted, the individual factors explain more on dietary behaviors. In other words, the environment constrains individuals' eating behaviors more severely in more disadvantaged neighborhoods. Therefore, the understanding of how retail food environment affects diet and eating behaviors may be particularly important in populations for whom social factors are very restricted (Lytle 2009). Lytle's conception may have many merits on empirical studies. For instance, Chung and Myers (1999) found that the poor in the Minneapolis –Saint Paul area faced great difficulties purchasing foods due to large grocery stores and chains were not located in that area. Morland, et al. (2002) have shown that there were four times as many supermarkets in the wealthiest areas compared to the least wealthy areas. Beaulac, Kristjansson, and Cummins (2009) concluded that low-income neighborhoods often lack supermarkets and chain stores, which may create barriers to accessing healthy food, especially for those who lack access to transportation. There were 26.5% Americans with low-income (below $20,000) did not have their vehicles in 2001 (Wallace, et al. 2005). Their food purchases were largely constrained by walking and public transportations. Even if they can shop with their family members and friends who have personal vehicles, their shopping frequencies might be limited. They are often constrained on a bi-weekly or monthly basis, and the purchased foods may become stale after several weeks' stock (Wiig and Smith 2009; Walker, Block, and Kawachi 2012).

Besides, Dr. Hillary Shaw's work "Food Deserts: Towards the Development of a classification" classified three broad themes relating to food access (Shaw 2006). The three barriers to food accessibility are 1) ability, 2) asset, and 3) attitude. The ability refers to the ability of physical access to healthy food whereas assets and attitudes reflect economic and sociocultural factors to prohibit healthy eating.

In conclusion, the influences of the food environment on diet and diet-related health outcomes are multi-facet and multi-scales. The effect of the built food environment on health outcomes is an essential focus for public health policy. Multiple SEMs have been developed to facilitate the understanding of how individuals and environments interact with each. I only discussed the most relevant theoretical frameworks that guide this dissertation. Regarding other socio-ecological models that link the food environment and eating behaviors and diet-related health outcomes, (Leia Michelle Minaker 2013) conducted a meta-analysis on conceptual models and summarized them systematically. One can refer to her work for more information.

Community Nutrition Environment

The community nutrition environment, according to Glanz and colleagues' framework (Glanz, et al. 2005), focuses on the geographic distribution of food outlets such as the type, number, and location. It refers to geographic accessibility to different types of food outlets. This framework emphasizes spatial elements in access, which complies with transportation specialists and experts' definition of accessibility as "spatial distribution of potential destinations, the ease of reaching each destination, and the magnitude, quality, and character of the activities found there" (Handy and Niemeier 1997, 1175). Based on this broad definition, we may declare that food accessibility is a measure of the ease or difficulty to obtain food in a given neighborhood. But food accessibility is a complex concept to define and quantify (Neff, et al. 2009; Ver Ploeg

2010; Walker, Keane, and Burke 2010) because food accessibility is multi-facet including the number and variety of food stores, opening hours, distance from residential locations, travel modes (Donkin, et al. 1999). To date, no formal definition is assigned to food accessibility. In spite of that, as Kimberley Hodgson pointed out, scholars articulate that there are three elements in food accessibility: “(1) nutritionally adequate, culturally appropriate, and affordable food; (2) income sufficient to purchase healthy food; and (3) proximity and ability to travel to a food source that offers such food” (Hodgson 2012, 15). Food accessibility is multifaceted or multi-dimensional per this definition. It includes not only the physical factor (i.e., proximity) but also contains economic and cultural elements. Food and Agriculture Organization of the United Nations (2006,1) included food accessibility as one of the four components of food security. It defines food access as below:

> Access by individuals to adequate resources for acquiring appropriate foods for a nutritious diet. Entitlements are defined as the set of all commodity bundles over which a person can establish command given the legal, political, economic and social arrangements of the community in which they live (including traditional rights such as access to shared resources).

This definition, indeed, is consistent with the rationales of Lytle's (2009) ecological model and Shaw's (2006) three broad classifications of barriers to having access to healthy foods.

Geographic/physical access to food outlets

Geographic accessibility has been extensively studied and explored, mainly because Geographic Information Science (GIS), a robust system to capture, display and analyze spatial information, has been used increasingly in public health research (O'Dwyer and

Burton 1998). The review of geographic/physical food accessibility mainly focuses on GIS methods.

The ease or difficulty of food access has been measured in different ways depending on the definition employed (Charreire, et al. 2010; Forsyth, Lytle, and Van Riper 2010; Hilmers, Hilmers, and Dave 2012). The methods are generally classified into two categories: descriptive approach and modeling approach (Luan 2016). The descriptive one includes density, proximity, variety, competition, while the modeling one contains spatiotemporal access, raster cost analysis, and gravity-based modeling (Luan 2016). The two approaches are reviewed below. The gravity-based model is explicitly explained since this dissertation employs this method.

Descriptive approach. Density: This approach computes the access or availability of food stores within a food environment (Charreire, et al. 2010). Density measures the number of stores in a certain geographical area. For instance, the count of food outlets (e.g., grocery stores, supermarkets, and convenience stores) per resident or areal unit within administrative boundaries or researcher-defined zones has been used to describe the retail food environment (Gallagher 2007). Density measures in administrative boundaries are often seen in Census Blocks (Smoyer-Tomic, et al. 2008), Block groups (Sharkey, Horel, and Dean 2010), Census tracts (Apparicio, Cloutier, and Shearmur 2007; Austin, et al. 2005; Baker, et al. 2006; Powell, et al. 2007; Moore, et al. 2008; Wang, et al. 2007; Zenk, et al. 2005b), and Postal Code (Clarke, Eyre, and Guy 2002; Donkin, et al. 1999). A problem using these units is that they do not necessarily correspond to where individuals may travel, since such administrative areas may have unequal sizes and little relevance to residents' shopping habits. A more common density approach is to use researcher-defined zone or said buffer. It draws a zone around a location within a specific distance to quantify accessibility of food stores. The location could be a respondent's home (Bodor, et al. 2008; Laraia, et al. 2006), school (Austin, et al. 2005; Zenk, et al. 2005b), food store (Clarke, Eyre, and Guy 2002; Donkin, et al. 1999), centroid of neighborhood (Block, Scribner, and DeSalvo 2004; Winkler, Turrell, and Patterson 2006), and even around bus route (Larsen and Gilliland 2008). Note that the buffer shape could be either circular or network. Regarding the difference between the two types of buffers, one can refer to Frank, Andresen, and Schmid (2004). The threshold of buffer distance is important to measure accessibility. For a circular buffer, its radius is between 100m and 2500m, which depends on the study settings. For example, 400m and 800m were used in Austin, et al. (2005); 2, 500m was employed in Winkler, Turrell, and Patterson (2006). For a network buffer, the size varies as well. It could be 400m (Russell and Heidkamp 2011), 500m (Furey, Strugnell, and McIlveen 2001), 1000m (Apparicio, Cloutier, and Shearmur 2007), and 2,500m (O'Dwyer and Coveney 2006).

Proximity: this approach uses either travel distance or travel time to represent the proximity to food outlets and residential locations (Ver Ploeg 2010). Regarding the distance, three distance formats are often used in research. They are Euclidean distance, Manhattan distance, and network distance (Charreire, et al. 2010; Wang 2014). Euclidean distance is a straight-line distance (also known "as the crow flies") between points on a flat surface. Though many studies have used this distance measure (D'Acosta 2015), its accuracy is limited because it cannot represent the actual travel routes in the real world. Manhattan distance, as the name suggests, describes a restrictive movement in rectangular blocks such as Manhattan in New York City. Manhattan distance can be used as an approximation for network distance if the street network is close to a grid pattern (Wang 2014). In most cases, roads, sidewalks, and driveways, however, are not in the regular or grid patterns. Therefore, Manhattan distance is not common to use in studies (Wang 2014). Instead, network distance has been widely utilized; it is the shortest distance between two locations along a transportation network. Studies often evaluate the shortest distance from consumers' residential locations to the nearest food stores (D'Acosta 2015; Larsen and Gilliland 2008). However, travel distances in urban or suburban areas are not comparable to the ones in rural areas. As a result, many studies consider access in rural areas and urban areas separately (Opfer 2010; Pearce, et al. 2007; Sharkey, Horel, and Dean 2010). Walkable distance has often been used to characterize access in urban and rural areas. The walkable distances are 1000 meters or 800 miles for urban areas. Some studies use travel time to measure the proximity to food outlets (Pearce, Witten, and Bartie 2006; Pearce, et al. 2007; Pearce, Day, and Witten 2008). OD matrix is often used to measure the shortest travel time from consumers' residential locations to food stores.

Variety: As Charreire, et al. (2010) pointed out, food studies often combine density and proximity measures to measure food accessibility because any single measure cannot adequately characterize the food environment. Measuring distance to the nearest food stores does not consider whether the consumer has other choices that may offer better products or lower prices. Thus, in addition to measuring density and proximity, Apparicio, Cloutier, and Shearmur (2007) and Sparks, Bania, and Leete (2011) calculated the mean distance to three different stores to measure the variety of food outlets in the study areas.

Competition: Gallagher (2007) and Gallagher (2006) included a competition of healthy and unhealthy foods in the evaluation of the food environment. He and his team argued that it is more important to measure how easy or difficult to choose between a mainstream food retailer (supermarket and grocery stores) and fringe food retailers (convenience stores and fast food restaurants) when relating food access to diet-related health outcomes such as overweight and obesity. The authors created the Food Balance Score to measure the relative balance of healthy foods to unhealthy foods. This score was calculated using the average distance to any mainstream food store divided by the average distance to a fringe food store.

Modeling approach. A modeling approach is different from the descriptive ones. It accounts for realistic constraints in assessing geographic accessibility. These constraints include time schedules and temporal variation (Spatiotemporal approach), travel cost (least-cost path), and distance decay (gravity-based modeling) (Luan 2016).

Spatiotemporal approach: Time plays a vital role in shaping the availability of food outlets thus influencing food access (Chen and Clark 2013; Chen and Clark 2016). However, traditional accessibility overlooks the constraint of time. Not all food outlets open 24 hours a day and seven days a week, and different store hours create an hourly and daily variation of availability to purchasing food. Therefore, a spatiotemporal approach can take daily maximum opening hours of food stores into account. Chen and Clark (2016) found that neighborhoods with low-SES in Columbus were not at a disadvantage of spatial access, but the limited temporal access is an issue. Some studies trace people's movements over one or more days with GPS units to understand how these movements influence food-shopping behaviors of specific population groups (Zenk, et al. 2011; Shannon 2016). Other scholars such as Michael Widener published a series of papers relating to the spatiotemporal access to healthy food stores (Widener, et al. 2013; Widener, et al. 2015). Widener and his colleagues integrated temporal patterns such as public transit schedules and commuting flows into accessibility to study the dynamics, as well as how these dynamics affect shopping behaviors at daily scale. On the seasonal level, temporal examples of food availability include seasonal opening and closing of farmers markets and produce stands in urban spaces. These studies measured a seasonal variation of food access in different locations (Lucan, et al. 2015; Widener, Metcalf, and Bar-Yam 2011).

<u>Raster cost surface analysis:</u> it is a cost-based surface in which each cell represents the accumulated cost between a cell and the nearest source. Then a point-in-raster function extracts the value of cost surface at a particular location to calculate the least cost path from an origin to destination. The cost could either be travel time or money. Some research developed a walk-time cost surface based on natural features such as slope, land cover, streams and trails (Balstrøm 2002; Sherrill, Frakes, and Schupbach 2010), and these features could create impedance for people to traverse. Hallett IV and McDermott (2011) derived a monetary surface in dollars (i.e., money spent on vehicles as a cost) and measured monetary cost to operate a motor to full-service food outlets. Specifically, they used $0.505 per mile of driving as a cost to construct the cost surface. Cost surface analysis is a raster-based method, while the descriptive and spatiotemporal approaches are vector-based methods. The use of raster cost has many merits: (1) vector-

based studies tend to use single points centered within census unit polygons, which may break up clusters, obscure underlying patterns, and add uncertainty to the result. However, raster cost does not have such problems; (2) Raster cost is more appropriate in rural food studies. In these remote areas, descriptive approaches such as food proximity or density measure are likely low because of vast distances and low population density (Mulrooney, et al. 2017). (3) Raster analysis has much more computation capacity than the vector one, and it works better for larger regions.

Gravity-based modelling: Gravity-based accessibility is another modeling approach (Luan 2016). It takes the distance-decay effect into account. The assumption is that people are more likely to procure foods in their immediate neighborhoods (Luan 2016). The probability of people visiting a food venue is decreasing as their residential locations are getting far away from this venue. This approach is often used in healthcare accessibility measures (Luo and Wang 2003; McGrail and Humphreys 2014; McGrail and Humphreys 2009; Ngui and Apparicio 2011), but is rarely utilized in food access studies. Up to date, only two studies are found to measure food access (Dai and Wang 2011; Kuai and Zhao 2017). A detailed review of this approach is depicted below.

The methods mentioned above (e.g., density, proximity, variety, and competition), however, are subject to two problems. First, they do not account for the spatial variation of food stores accessibility within the analysis unit. For example, three food stores are within one- mile buffer of a residential location; these methods assume that people in this buffer zone have equal access to the three stores. Second, they do not consider the difference in stores size or business capacity, assuming each food store has the same production capacity. Hansen's gravity model could address these issues. It calculates the accessibility at a particular location using capacity at supply sites divided by the travel impedance from that specific location to supply sites.

Hansen's simple gravity model, however, only considers the supply side (consumer's decision-making), but does not account for demand side (consumer's competition for the service of food provider). To account for both sides many existing studies calculate accessibility using the ratio of the number of food stores to the population within an administrative unit (Michimi and Wimberly 2010; Moore and Diez Roux 2006; Raja, Ma, and Yadav 2008).

Supplies and demands are not always in the same analysis unit. Peng (1997) and Wang (2000) used a floating catchment area (FCA) as a dynamic unit to solve the spatial mismatch problem. The FCA draws a circle or square around a location to define a filtering window and uses the average (the ratio of supply to demand) value within the window to represent the value at the site. The window moves across the study area until averages at all locations are obtained. However, FCA (one-step FCA) ignores the interactions between supply and demand across boundaries of residential areas: competition for supply from adjacent neighborhoods and competition for demand in nearby residential units.

<u>Two-step Floating Catchment Area (2SFCA) family:</u> Luo and Wang (2003) proposed a well-known two-step floating catchment area (2SFCA) method to measure spatial accessibility to healthcare. This method not only uses the ratio of supply versus demand as a measure of availability, but also considers the interaction of supply and demand across the unit boundaries. It consists of two steps to measure healthcare access. The first step focuses on the supply side and the second step focuses on the demand side.

The 2SFCA method can be interpreted as follows: 1) from a particular healthcare provider searches a 30-minute driving catchment, sum up the total population that the supplier can reach within that catchment, then calculate a supplier-to-population ratio. 2) from a particular population-weighted centroid searches a 30-minute driving catchment; obtain the computed provider-to-population ratio in step 1 and sum up all provider-to-population ratios to calculate the accessibility.

2SFCA is intuitive and easy to interpret, and it considers the interaction between supply and demand at the boundary of analytic units. However, it has limitations. 2SFCA assumes that people within the catchment area (i.e., 30-minutes driving zone) have equal access to the supply site regardless of the actual travel distance or time. That is to say; it does not consider distance decay within the catchment area.

Luo and Qi (2009) developed an Enhanced Two-Step Floating Catchment Area (E2SFCA) method to overcome the limitation of 2SFCA. Travel preference is decreasing while moving away from the origin; this is known as the term "distance decay." To account for the distance decay in the catchment, it divides the catchment area into three drive time zones and applies different weights for these zones in both steps. It also consists of two steps: 1) from a particular healthcare provider searches three drive time zones (catchments): 0-10, 10-20, and 20-30 minutes, sum up the total population that the supplier can reach within each zone, then compute a provider-to-population ratio. 2) From a particular population-weighted centroid searches the three drive time zones: 0-10, 10-20, and 20-30 minutes, obtain the calculated provider-to-population ratio in step 1 in each drive time zone and sum up all provider-to-population ratios to compute the accessibility.

E2SFCA is a special case of gravity-based accessibility index model; the different weights for three sub-zones in E2SFCA reflect the impedance function in a discrete manner. E2SFCA has drawn criticism because of three problems. First, it still does not consider the distance decay within each subzone. Second, there is an abrupt weight change at the boundary of different subzones. Third, it does not explain why the weights were chosen in each subzone, and the choice of weights seems arbitrary (Vo, Plachkinova, and Bhaskar 2015).

To compensate for the deficiency of E2SFCA, many researchers have attempted to improve it using gradual-decay 2SFCA methods. Like the E2SFCA, these gradual –decay 2SFCA methods have two steps as well. The most notable difference is the distance decay function. They used a continuous gradual-decay function to replace the discrete weights in E2SFCA. There are various ways of conceptualizing the distance decay function, which includes power function (Wang 2014), exponential function (Wang 2014), Gaussian function (Dai 2010; Shi, et al. 2012), and Kernel density function (Guagliardo 2004). Their methods transform the discrete and stepwise distance decay in E2SFCA into a continuous distance decay. The only difference between E2SFCA and Gradual-decay 2SFCA is the way to assign weight values (discrete vs. continuous). In essence, they share the same principle. Thus, they could be classified into the same class of metrics.

Wan, Zou, and Sternberg (2012) argued that the E2SFCA does not consider the potential competition among multiple supply sites available for a population site. That is to say, the probability of people visiting a supply site would decrease if there are other supply sites nearby. Therefore, Wan, Zou, and Sternberg (2012) proposed a three-step

Floating Catchment Area (3SFCA) method to adjust this problem. Firstly, a selection weight (G) is computed for all population sites and supply sites parings. The selection weight (G) is then integrated into the two subsequent catchment calculations (step two and three).

The 3SFCA method introduces a selection weight to adjust the population demand (or account for the competition among multiple supply sites); the selection weight is a probability for people to visit a supply site in the catchment area. However, as Luo (2014) argued, the calculation of the selection weight in 3SFCA method is only based on travel minutes. The supply capacity of a facility also affects people's selection. Moreover, the use of four subzones in 3SFCA method is problematic because it assumes that the weight in each zone is fixed. To fix these problems, Luo (2014) incorporated the Huff model into Floating Catchment Area to propose a Huff-based 2SFCA (single mode) model. Similar to the 3SFCA, the first step of Huff-based 2SFCA is the calculation of selection weight (G). The difference between the two methods is the way of the computation of selection weight. The Huff-based 2SFCA computes the selection weight (G) on a supply site based on the "Huff model", which quantifies the probability of people's selection on a supply site with considerations both travel cost and capacity of a supply site. The Huff-based selection weight (G) is then integrated into the supply and demand catchments.

The Huff-based 2SFCA essentially is comparable to the 3SFCA. The Huff model-based selection probability further adjusts the population demand considering both travel cost and the attractiveness of the supply site. Luo (2014) articulated that the Huff-based 2SFCA method could add more variabilities on the measure of spatial accessibility due to its more delicate adjustment of population demand. It is informative to identify under-served areas to allocate resources to where they are needed most.

Neighborhood social deprivation and distribution of food outlets

The distribution of food outlets varies spatially between neighborhoods, which may force people to reside in their locations to develop a healthy or unhealthy eating habit. It is more likely for racial/ethnic and socioeconomically disadvantaged residents to be located in food deserts and food swamps. The below part reviews studies that examine these disparities.

Racial/ ethnic disparities in access to food outlets. Morland, et al. (2002) found that neighborhoods dominated by African American people had fewer supermarkets compared to White-dominated neighborhoods in Maryland, Minnesota, Mississippi, and North Carolina. Zenk, et al. (2005b) found that the most impoverished African Americans neighborhoods were 1.1 miles farther from the closest supermarket compared to the most impoverished White neighborhoods. When drawing a three-mile radius to count the number of supermarkets, the White neighborhoods had 2.7 more supermarkets than the most impoverished African neighborhoods. Powell, et al. (2007) found that even if accounting for neighborhood income, the availability of chain supermarkets in African American neighborhoods was only 52% of the White communities.

Racial disparities are also found in access to fast-food restaurants, which has been linked to the increases in obesity. Pearce, et al. (2007) found significantly negative statistical associations between access to the nearest fast-food outlets and social deprivation in neighborhoods and schools in New Zealand. It is found that predominantly Black neighborhoods in New Orleans were more likely to contain a higher density of fast-food restaurants. Gordon, et al. (2011) found that African American block groups had a significantly lower proportion of healthy bodegas and greater accessibility of fast-food outlets. Kwate, et al. (2009) found that the percentage of African Americans in block groups was positively associated with fast-food outlets density in New York City. James (2004) argued that dietary habits among African Americans are poor; their diets often involve high calorie and high amount of sodium intake, low intake of fruits and vegetables, and grains.

Hispanic- dominated neighborhoods are found to have high access to convenience stores and fast food restaurants. Galvez, et al. (2008) analyzed food outlets in 165 census blocks in East Harlem, NY and found that Hispanic census blocks had a significantly higher density of convenience stores and fast-food outlets than racially mixed census blocks. Another research in Nieces County in TX found that it is more likely for Hispanic neighborhoods to have convenience stores than in-Hispanic White neighborhoods (Lisabeth, et al. 2010). A more comprehensive review of ethnic disparities in access to food stores can be seen in Hilmers, Hilmers, and Dave (2012).

Economic and socio-demographic disparities in access to food outlets. Research has found that in economically deprived areas people's access to healthy food is more reduced, but access to unhealthy food is more abundant (Dowler, et al. 2001). Broda, Leibtag, and Weinstein (2009) compared the difference in food prices across household income levels. They found that lower-income consumers often shop at stores offering lower rates. Moreover, high-income households spent the most considerable amount of money on groceries, which is 2-3% higher than low-income ones. This study suggests that food price is not the single indicator of economic aspect to determine people's food shopping choices; income should also be considered as an essential component of food affordability.

It has been noticed that low-income people often purchase cheap, and energy-dense food. Compared with higher-income households, they are buying more discounted items and private-label brand products; take advantage of volume discounts, and purchase less monetary foods (Leibtag and Kaufman 2003). Food, to them, is a way to survive. Therefore, they try to maximize calories percent to avoid being hunger (Drewnowski 2009). In many cases, they do not care whether the food they consume is healthy or not. Instead, they are more concerned with whether the purchased food can stuff their stomach. By contrast, high-income groups have more privileges to choose what they eat. It is more likely for them to put food nutrition rather than food price on their priority. They generally purchase more on fruit and vegetables (Cassady, Jetter, and Culp 2007).

Chung and Myers (1999) found that the poor in the Minneapolis –Saint Paul area faced disadvantaged difficulties in purchasing foods due to large grocery stores and chains were not located in that area. Morland, et al. (2002) showed that there were three times as many supermarkets in the wealthiest areas compared to the least affluent areas. Beaulac, Kristjansson, and Cummins (2009) concluded that low-income neighborhoods often lack supermarkets and chain stores, which may create barriers to accessing healthy food, especially for those who require access to transportation. There were 26.5% Americans with low-income (below $20,000) did not have their vehicles in 2001. Their food purchases were constrained mainly by walking and public transportations because foods could be heavy to carry. Even if they can shop with their family members and friends who have personal vehicles, their shopping frequencies might be limited. They are often constrained on a bi-weekly or monthly basis, and the purchased foods may become stale after several weeks' stock (Wiig and Smith 2009; Walker, Block, and Kawachi 2012).

Larsen and Gilliland (2008) found that low-income people in the inner city have the poorest access to supermarket, and this situation was getting worse over the years. It is noticed that fast food restaurants are more available in low-income neighborhoods (Fleischhacker, et al. 2011; Hilmers, Hilmers, and Dave 2012). These restaurants serve many energy-dense, nutrient-poor foods at relatively low prices (Kestens and Daniel 2010). The frequent consumption of fast food may lead to weight problems in the long run (Powell and Nguyen 2013).

Canto, Brown, and Deller (2014) have shown a strong relationship between rural poverty and healthy food access where higher poverty is associated with poor healthy food access and adverse health. Bhattacharya, Currie, and Haider (2004) explored the relationship between nutritional status, poverty, and food insecurity. The result shows that poverty is predictive of poor nutrition among all age groups except school-age children. Drewnowski and Specter (2004) reviewed the relation between obesity and diet quality, energy density and costs. They concluded that poverty is associated with the lower expenditure of food, moderate consumption of fruit and vegetable, and lower-quality diets. A study investigated the relationship between neighborhood characteristics, grocery store accessibility, fruit and vegetable intake, and weight status (Mushi-Brunt, et al. 2007). It is found that 50% of high-poverty neighborhoods did not have grocery stores. Moreover, children in high poverty communities consumed significantly fewer servings of fruit and vegetables than those in low poverty one ($p< 0.001$).

Milicic and DeCicca (2017) examined the impact of unemployment on fruit and vegetable consumption in Candia. Their finding showed that there was a robust negative association between the unemployment rate and the consumption of fruits and vegetables regardless of gender and education levels. In other words, a high unemployment rate reduced people’s fruit and vegetable consumption. Smed, et al. (2018) explored the consequences of unemployment on diet and purchase behavior in Denmark using longitudinal data from 2008 to 2012. The authors summarized the influences of unemployment on a diet in three periods. The short-term impact is that unemployment led to people increasing in food expenditure and consumption of saturated fat, protein, and animal-based foods. The medium-term effect is the declining consumption of these nutrients whereas carbohydrates and sugar will replace these nutrients in the long term. Therefore, they concluded that unemployment could substantially influence people's diet and consumer behaviors.

Barker, et al. (2008) examined the relationship between education attainments, food involvement, and fruit and vegetable consumption. They found that food involvement for women decreased with decreasing educational attainment, and women with low educational attainment ate fewer fruits and vegetables. Lawrence, et al. (2009) organized a semi-structured interview to examine why women with lower educational attainment struggled to make healthier food choices. They identified following factors: (1) less control on families’ food choices; (2) less support for attempts to eat healthily; (3) fewer opportunities to learn good food-related practices; (4) more ambiguous beliefs about the negative consequences of eating nutritiously.

Research gaps in measuring food deserts and food swamps

Based on the above reviews, I could identify three gaps in the measurement of food deserts and food swamps.

(1) The spatial access to food outlets is often analyzed by employing descriptive approaches including density, proximity, variety, and competition. Few food environment studies utilized modeling approaches such as 2-Step Flowing Catchment Area (2SFCA) family for the measurement. The use of this family has two advantages. First, it considers both a customer's choice of food retail and population's share of food store capacity. It is important for people in deprived areas who experience more competition for food as the population density is high and supermarkets and quality grocery stores with healthy food are few. Second, it accounts for competition for food stores from adjacent neighborhoods and competition for customers from food stores in nearby residential units. In other words, it considers the supply and demand interaction beyond boundaries. The most recent model in the 2SFCA family is Huff-based 2SFCA. This model has the third merit. For example, if there are other supply sites are near a food store, the probability of people visiting this store will decrease. The Huff-based 2SFCA method further adjusts this problem and solve the over-demands issue by considering both travel cost and the attractiveness of the supply site.

(2) When measuring food accessibility, most of the literature assumes that people travel by automotive and vehicles. This assumption does not separate the population by mobility or transportation modes. People who are in disadvantaged status often end up walking or taking the public transit to shop. In food desert studies, researchers are mostly concerned about disadvantaged groups. These marginalized groups often cannot afford

personal vehicles. Therefore, it is necessary to consider alternative transportation modes such as public transportation and walking. Widener, et al. (2015) measured food accessibility in Cincinnati, Ohio based on transit network focusing on the disadvantaged population. Larsen and Gilliland (2008) measured healthy food accessibility in London, Ontario for 1961 and 2005 based on two transportation modes: walking and public transportation. They argued that most of the low-income people could not afford personal vehicles. Network Analyst in ArcMap was used to calculate proximity to the nearest grocery stores and also the density, or number, of grocery stores within 1000-meters of each block centroid. They identified one food desert in an east London neighborhood. However, their research only focuses on the disadvantaged population and neglects the comparison between the groups with and without private vehicles. It is also necessary to integrate multiple transportation modes into the measurement to obtain an integral view of the overall picture of accessing to retail food environment.

(3) Most studies used simple or traditional statistics to associate physical access with social-demographic variables. For example, Gordon, et al. (2011) developed a Food Desert Index based on the competition of healthy and unhealthy foods in New York City. These food access index components were measured, ranked and scored as low, medium, and high to create a scale range of 3 (poor) to 9 (high). The relationship between socio-demographic variables (i.e. race and ethnicity and median income) were also analyzed. The Food Desert Index and the demographic variables were combined to create a total food desert score. The areas with the highest food desert scores were identified as food deserts. This is a typical approach being taken for the combination of physical access and sociodemographic variables, which is also used in the classic work by Apparicio,

Cloutier, and Shearmur (2007). There are two problems for this approach. First, the classification of low, medium, and high is arbitrary. There are no standard thresholds in the literature. Second, these measures do not consider spatial associations between physical access and socio-demographic variables. Spatial data frequently conform to Tobler's First Law of Geography "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). It refers to those observations in nearby locations are more similar than would be expected on a random basis (Stein 2011). This phenomenon is called spatial dependence (e.g., positive spatial autocorrelation), which increases the likelihood of similar values in neighboring units. If any significant spatial dependence is present in the dataset, the analysis violates the assumption of independence of observations in conventional statistics.

For food swamp studies, they are also subject to the same issues with food deserts. Besides, most studies only focused on spatial dimension but neglected economic and sociocultural aspects. The present research tackles these issues and constructs a more robust method to identify food deserts and food swamps in Austin, TX.

Consumer Nutrition Environment

Food availability

Ver Ploeg (2010) argued that food accessibility only could be measured once the availability of food has been measured. Therefore, it is beneficial to understand the concept of food availability. Similar to the concept of food accessibility, there is no consensus on this concept. For instance, Ver Ploeg (2010) articulated that food accessibility is a measure of availability of an adequate number of foods in quality regardless of domestic or imported products. While Donkin, et al. (1999) viewed food availability to measure the presence or absence of types of food stores and food sold in the stores. The former definition is generic, but the latter is intuitive and specific.

Food availability measures whether certain food items are available or not. This measure overcomes the limitation of the assumption that supermarkets always have healthy foods while convenience stores always have less healthy food options (L Minaker 2013; Leia Michelle Minaker 2013). The food availability is variant in different types of food outlets. Food availability can be measured by checklists (e.g., yes/no questions on the availability of specific foods) and shelf-space (e.g., linear length of shelf-space, yes/no questions on whether healthier food items account for 50% of the total shelf-space) (L Minaker 2013).

Farah and Buzby (2005) examined food availability with dietary outcomes. It was conducted among children to find whether fruit and vegetable availability predict people's intake of these products. It is found that the association between restaurant fruit and vegetable availability and fruit and vegetable intake was significant and positive, although the association between grocery store fruit and vegetable availability was not correlated with fruit and vegetable consumption (Farah and Buzby 2005). Slater, et al. (2009) suggested that local availability of fruits and vegetables are significantly associated with intake. Weiss, et al. (2007) found that neighborhood availability of dark green and orange vegetables was significantly associated with residents' consumption. While Frank, et al. (2009) had some unusual findings. They found that higher healthy food availability was associated with higher BMI among urban residents where are predominantly by White, and lower BMI was found among urban residents where are predominantly African American and low-SES people.

Food affordability

Food affordability is conventionally used to depict “the cost of food relative to an individual’s or household’s income or purchasing power” (L Minaker 2013). The food affordability has been adapted to understand food costs within a neighborhood, and it can be absolute (e.g., the cost of a food basket) or relative (e.g., the cost of healthy foods relative to their unhealthy counterparts) (Leia Michelle Minaker 2013).

The cost of food has been identified as an essential barrier to healthy eating in low-income communities (Chung and Myers 1999) and households (Cassady, Jetter, and Culp 2007). It also considers food price as an indispensable measure of food accessibility (Chung and Myers 1999; Ver Ploeg 2010). Imagine that stores sell food items at a high price, and it is not likely for low-income groups to purchase them even if they are physically close to that store.

Food price not only affects people’s food choice and their purchasing behaviors, but also impacts health outcomes such as obesity. Bowman (2006) used USDA's survey data to investigate women's attitudes about food purchasing. He found that 46.8% of women considered food price very important, especially for African-American and Hispanic women food shoppers. Energy-dense foods such as fats, sweetened beverages, and chips are usually inexpensive and cost lower; people from low-income households tend to buy more of these innutritious foods thus affect their diets in the long run (Briefel, Wilson, and Gleason 2009). The relatively low cost of energy-dense foods could contribute to the prevalence of obesity among low-income groups. By contrast, nutritious food such as fruits and vegetables could increase diet costs, and food price of these fresh and healthy items could create a barrier to their eating behaviors and health status.

It has been found that food price tends to be lower in deprived areas (Block and Kouba 2006; Horowitz, et al. 2004). For example, Block and Kouba (2006) compared the availability and affordability of food market basket in two communities (Austin and Oak Park) in Chicago; Austin is a deprived area with many low-middle-class African Americans, whereas Oak Park is an affluent area with a large number of upper-middle households. They used the US Department of Agriculture's standard market basket to survey 134 food stores. Meanwhile, they articulated that USDA's standard market basket does not include many items that are culturally important for African American people. They thus added some more items such as greens, sweet potatoes and baby formula into the list, and ultimately 102 food items were surveyed. It is found that priced averaged lower in Austin than in Oak Park. A similar study was conducted by Horowitz, et al. (2004). The survey places were East Harlem (a racial/ethnic minority neighborhood) and Upper East Side (White and affluent community) in New York. It is found that the median prices of all surveyed food were significantly higher in the Upper East Side than in East Harlem.

The cost of healthy food items such as fruit and vegetables is found to be more expensive in deprived and rural areas (Harrison, et al. 2007; Bovell-Benjamin, et al. 2009). A report by Kaufman (1997) compared food costs on a per pound basis in both all-income and low-income households that low-income in the U.S. in the year 1977 and 1987. The results showed that low-income families paid higher per-unit food costs for fruit juices and eggs, but they paid lower for other food categories. Bovell-Benjamin, et al. (2009) found that the price of fruit and vegetables was generally higher in Tuskegee (95.1 % were African American) than in Auburn (affluent, predominantly White).

Harrison, et al. (2007) surveyed 97 stores in 2004 and 81 stores in 2000; the cost of healthy food in remote areas was 29.6% ($p < 0.001$) higher than cities. The price of healthy foods had increased more than that of less healthy foods in these two years. However, there was no consensus about this viewpoint. A food market study showed that higher and middle- income neighborhoods have significantly higher priced fruits, orange, and other vegetables than very low-income neighborhoods (Cassady, Jetter, and Culp 2007).

Food price for similar items varies in different types of food stores. The most consistent finding is that on average food price is higher in convenience stores than that in grocery stores (Broda, Leibtag, and Weinstein 2009). For example, Kaufman (1997) found that small stores charged 10% more than supermarkets on average. Convenience stores usually are smaller than grocery stores in size. It is likely to have limited food items and located in low-income and rural neighborhoods than in the suburbs. Consumers face a relatively constricted price range so that they cannot take advantage of "economies of size". Whereas grocery stores are more likely to locate in the suburbs, and they provide an extensive range of brand, size, quality, and products, which allows consumers to choose items with prices falling within their food budgets. Grocery stores and supermarkets can charge lower prices due to their "economies of size" by offering store labels and more generic food items (Kaufman 1997).

Food quality

Food quality measures the quality characteristics of foods in food retailers. This measure is subjective, and the assessment is usually based on the direct observation of the appearance of fruits and vegetables. For instance, whether fruits and vegetables appear withered or bruised. Food quality measurement can reflect residents' satisfaction with the foods in their neighborhoods. Poor food quality (e.g., withered or bruised fresh produce, rotting meat, and expired canned foods) could deter food purchasing behaviors (Brown 2014) and therein impose an adverse effect on diet quality and health outcomes. Food quality, like availability and affordability, is often regarded as a component in food access measures because store type is used as a proxy for food quality. For instance, grocery stores are often regarded to have good food quality; while convenience stores generally offer a lower quality of food items than grocery stores (Sallis and Glanz 2009). Only a few studies have assessed neighborhood food quality (Thow, et al. 2011). Some of them evaluated food quality in an objective manner (Coulter 2009; Cummins et al. 2009; Glanz et al. 2005) by directly observing the appearance of specific food items in food stores. Others took quantitively method such as interviews to ask participants' perception of food quality, especially for low-income populations and minority groups. For example, research focused on the perceived quality of produce among African-American women in Detroit (Zenk, et al. 2005a). The authors found that positive perceptions of product quality were positively associated with increased intake regardless of store type and location, residential' age, income, and education.

Food labels

Food labels can provide important nutrition information for consumers. The information such as calories, total fat, saturated fat, cholesterol, protein, serving size, etc., can help consumers understand the ingredients they take, which may assist them in tracking their daily consumption of certain healthy nutrition and avoiding unhealthy ingredients. Also, it is easy to compare the nutrient content of different food options via reading food labels. Through the label comparison, people can make a better choice to select the one with lower fat and thus improve their overall diet quality. Food labels also enable consumers to quickly find food items high in fiber and protein if they intend to increase the intake of these nutrients. Some research found that using food labels was significantly positively associated with lower fat intake (Neuhouser, Kristal, and Patterson 1999; Ollberding, Wolf, and Contento 2011) and higher consumption of fruit and vegetables (Satia, Galanko, and Neuhouser 2005). It is also found that food labeling was related to a declined trend of body weight and obesity. For example, Variyam and Cawley (2006) tested whether the release of the Nutrition Labelling and Education Act (NLEA) influenced American adults' body weight. They found that the use of food labels was negatively associated with body weight for non-Hispanic white women. More importantly, the use of NLEA regulation was estimated to save $ 63 to $ 166 billion due to the decrease in body weight and obesity prevalence. Therefore, an exploration of in-store characteristics such as food labeling is crucial to improve deity quality and fight against food-related diseases.

Review of food survey instrument

To evaluate the consumer nutrition environment the most important part is to decide food items to be surveyed since more than 20,000 food items are sold in U.S. grocery stores (Kaufman 1997). At the minimum, the most typical foods that represent what the population consumes in specific areas should be selected. In other words, this food basket typically should meet the minimum nutrition requirements of a defined population group or individuals (Anderson, et al. 2011; Gans, et al. 2010). Nevertheless, constructing such a food basket could be a complex issue, since very similar items could be sold in different brand type, flavors, and package size.

A common way to measure consumer nutrition environment is to use a standard food basket survey. A couple of standard food baskets have been developed to satisfy the needs of different purposes (Block and Kouba 2006; Anderson, et al. 2011; Ling 2005; Cassady, Jetter, and Culp 2007; Harrison, et al. 2007; Palermo, et al. 2008; Bovell-Benjamin, et al. 2009). For example, the Diet and Health Knowledge Survey (DHKS), Victorian Healthy Food Basket in Australia, and Healthy Eating indicator shopping basket tool (HEISB) in U.K are commonly used in the survey

USDA's Thrifty Food Plan (TFP) is a standard food survey instrument used in the U.S (Block and Kouba 2006; Bovell-Benjamin, et al. 2009; Cassady, Jetter, and Culp 2007; Jetter and Cassady 2006). It is a meal plan that demonstrates how a diet that meets the minimum recommendations of the Dietary Guidelines for Americans may be achieved by a family of four on a modest budget or by food stamp recipients. The TFP consists of 87 food items and could be divided into seven food categories: (1) Fruits and Vegetables, (2) Bread, Cereals and Grain, (3) Milk and Cheese (4) Meat and Meat Alternatives (5) Fat and Oils (6) Sugars and Sweet (7) Condiments and Species. This survey tool includes all the necessary food categories for healthy eating. However, it does not specify healthier foods and their regular counterparts, so it is not entirely useful for comparison.

NEMS (Nutrition Environment Measures Survey) tools are appropriate survey instruments to measure consumer nutrition environment (Glanz, et al. 2007). The NEMS Tools are observational measures to assess the consumer nutrition environments in food outlets, specifically stores, corner stores and restaurants (Glanz, et al. 2007). They are used to measure the availability of healthful choices, prices and quality. NEMS-S (for stores) and NEMS-R (for restaurants) are two earliest tools developed by (Glanz, et al. 2007) and (Saelens, et al. 2007), respectively. The successive NEMS tools include NEMS-CS (for corner stores), NEMS-V (for vending machines) and NEMS-P (for perceived nutrition environment). The present study focuses on the nutrition environment in food stores, and thus NEMS-S can satisfy the research needs. For NEMS-S instrument, 11 measures are included: milk, fresh fruits and vegetables, ground beef, hot dogs, frozen dinners, baked goods, beverages (soda/juice), whole grain bread, baked chips, and cereal. Except for the fruits and vegetable group, other measures include both healthier and less-healthy options (such as regular ones). The advantage of NEMS-S is its distinction between healthy and less healthy foods, which may facilitate food price comparison of both healthy and less-healthy foods in different geographic locations. For example, semi-skimmed milk versus whole milk, whole wheat bread versus white bread. Design as such could lead to a direct availability and food price comparison in different types of food stores over geographic locations.

NEMS-S has been used in many studies and projects (Coulter 2009). For example, Coulter (2009) used NEMS-S to conduct the availability, price, and quality of healthy food options in a low-income racially diverse neighborhood in Seattle. The author found that less healthy food options were much more available than healthy food alternatives. There is no significant difference in availability, price, and quality of healthy food options by either race or income. Leone, et al. (2011) evaluated the availability and affordability of healthy food items in Leon County, Florida. The results show that the availability of all healthy food items was significantly different by store types. The availability and affordability of healthy options are different by income level, but not by racial composition.

NEMS-S is a reliable tool designed by Glanz, et al. (2007) for the target population in the Atlanta metropolitan area. It might not be appropriate to other states or cities because it may miss some culturally appreciated foods that are important to residents. Take Texas as an example: 35 % of the population is Hispanic/Latino. An ethnic group such as Hispanic/Latino people have their own cultural identities. Hispanics reveal differences in beliefs, preferences, and culture on diets. The cultural values make them more prefer on high-fat foods. Therefore, there is a need to tailor the NEMS-S tool to make it more adaptive to local communities. Physical Activity, and Obesity Prevention (NPAOP) developed such a tool, and named it as "Texas Nutrition Environment Assessment in Stores (TxNEA-S)" (Gloria and Steinhardt 2010). It is an adaptation of the NEMS tool that included additional foods that are culture-important to Hispanics and other minorities in Texas. This tool has been adopted by the Texas Department of State Health Services as a reliable tool to assess the availability, price, and quality of healthy foods in retail stores in Texas[13]. There are 134 food items on the list. It includes some culturally important foods for Texan such as tortillas and tropical fruits. It is comprised of 14 food categories[14]: Fresh Fruits (16 items), Fresh Vegetables (13 items), Convenience-added Produce (4 items), Dairy Milk (4 items), Dairy Yogurt Cottage Cheese (4 items),

[13] http://www.dshs.texas.gov/Obesity/TXNEAS/

[14] http://www.dshs.texas.gov/uploadedFiles/Content/Prevention_and_Preparedness/obesity/TxNEATool%20.pdf

Dairy other cheese (8 items), Canned Fruits (10 items), Canned Beans and Legumes (9 items), Grains Cereal (8 items), Bread and Commercial baked goods (9 items), Grains (7 items), Bulk Dry Grains/Beans (8 items), and Frozen Fruit and Vegetables (13 items). Gloria and Steinhardt (2010) used TxNEA-S instrument to survey food stores in Austin, Texas and confirmed its reliability and validity in measuring consumer nutrient environment. Despite these advantages, TxNEA-S is still subject to several problems. For instance, it does not have meat on the list; beef, chicken, and fish are essential sources for protein. Moreover, it does not contain beverages such as coke, which is consumed heavily by American people and is believed to have great potential to make individuals obese. Therefore, I would customize the TxNEA-S in this research to make it adaptive to my research purposes.

3 THEORITICAL FRAMEWORK

In Chapter two, research gaps have been identified. Retail food environment is complicated since it involves multifaceted entities such as food access, economic and sociocultural barriers, and consumer nutrition environment. I thus must consider how to conceptualize these connections between them. In other words, this chapter attempts to explore the theories beneath the scene in this study.

Theoretical Framework

A diagram is present to show the theoretical framework of my dissertation (Figure 3.1). The well-cited Glanz and colleagues' nutrition framework guides this dissertation (Glanz, et al. 2005). The assessment of RFE in their framework consists of two dimensions: the community nutrition environment and consumer nutrition environment.

First, this dissertation proposes a new method to measure geographic food accessibility (Aim 1). Geographic accessibility to food stores and restaurants is one of the essential components of the food environment and reflects the community nutrition environment. The methodology of geographic access is predominantly descriptive. It is often operationalized as proximity, density, and variety. Modeling approach such as the Two-Step Floating Catchment Area (2SFCA) family, however, has rarely been used in food accessibility studies.

Moreover, accessibility varies by transportation modes, and it is necessary to integrate transportation modes into the measurement. A novel multi-modal Huff-based 2SFCA is proposed to measure geographic/physical accessibility to food stores and restaurants. Three transportation modes (i.e., walking, driving, and taking public transit) are taken into consideration. The physical accessibility indices for both healthy and unhealthy food venues are computed.

Second, I explore the relationship between spatial food access and sociodemographic marginalization (Aim 2). As discussed in Chapter two, social deprivation may create barriers to access to foods. Three theories guide this dissertation: Macintyre's deprivation amplification hypothesis (Macintyre 2007), Lytle's ecological model (Lytle 2009), and Shaw's food access classification (Shaw 2006). More importantly, Shaw's three broad groupings had identified three barriers to access healthy foods: ability, asset, and attitude. Based on this classification, I further break down sociodemographic marginalization into two broad categories: economic factors (i.e., asset) and sociocultural factors (i.e., attitude). The economic factors are measured by four variables: household income, unemployment, income below the poverty line, and household lacking kitchen facilities; these four variables are combined to create a composite index —Economic Deprivation Index. The sociocultural factors consist of four variables: house ownership, education attainment, Hispanic population, and linguistic isolation; these four variables are combined to create a composite index — Sociocultural Deprivation Index. Then the relationships between the three indices (physical accessibility index, economic deprivation index, and sociocultural deprivation index) are explored by spatial regression model (e.g., spatial lag model and spatial error model) and semi-parametric GWR model to account for the spatial autocorrelation and spatial non-stationarity problems. The ultimate goal is to identify food deserts and food swamps in the study area. The three indices are combined to delineate food deserts and food swamps by considering their spatial relationships (i.e., Low-High and High-High relationships). The consideration of spatial relationships between the three indices (or physical, economic, and sociocultural factors) ensures more accurate identification of problematic areas.

Third, this dissertation examines the consumer nutrition environment (Aim 3). The identification of food desert and food swamps in Aim 2 reflects the food access and sociodemographic deprivation issues at neighborhood (or census block group) level, which cannot reveal consumers' food shopping experiences (i.e., food quality and price) at an individual store. Therefore, an exploration of consumer nutrition environment is complementary to the Aim 1 and 2 and is integral to this dissertation. The consumer environment includes food availability, food price, and food quality in stores as proposed by Glanz and colleagues’ framework (Glanz, et al. 2005; Glanz, et al. 2007). I explore how the consumer nutrition environment vary in different community nutrition environments such as food deserts, food swamps, and food oases (i.e., the areas are food secured neighborhoods). In this dissertation, an in-store food audit is conducted in food stores in three communities in Austin. The three neighborhoods are selected from the areas that are classified as food deserts, food swamps, and food oases through the Aim 2 of this dissertation. Four types of food stores are surveyed: supermarket, grocery stores, supercenter, and convenience store.

This dissertation explores both community and consumer nutrition environment, which is complete for an objective assessment of the retail food environment. The community nutrition environment mainly focuses on geographic access to food outlets. It emphasizes the "spatial" factors in access to food outlets (Aim 1). However, "non-spatial" factors such as sociodemographic marginalization could create barriers to access to healthy foods or promote access to unhealthy foods (i.e., deprivation amplification hypothesis). Therefore, food access issue should consider the spatial and non-spatial entities together. The Aim 2 identify the marginalized factors and explores the relationships between the two entities. Aim 1 and 2 are connected by the deprivation amplification hypothesis.

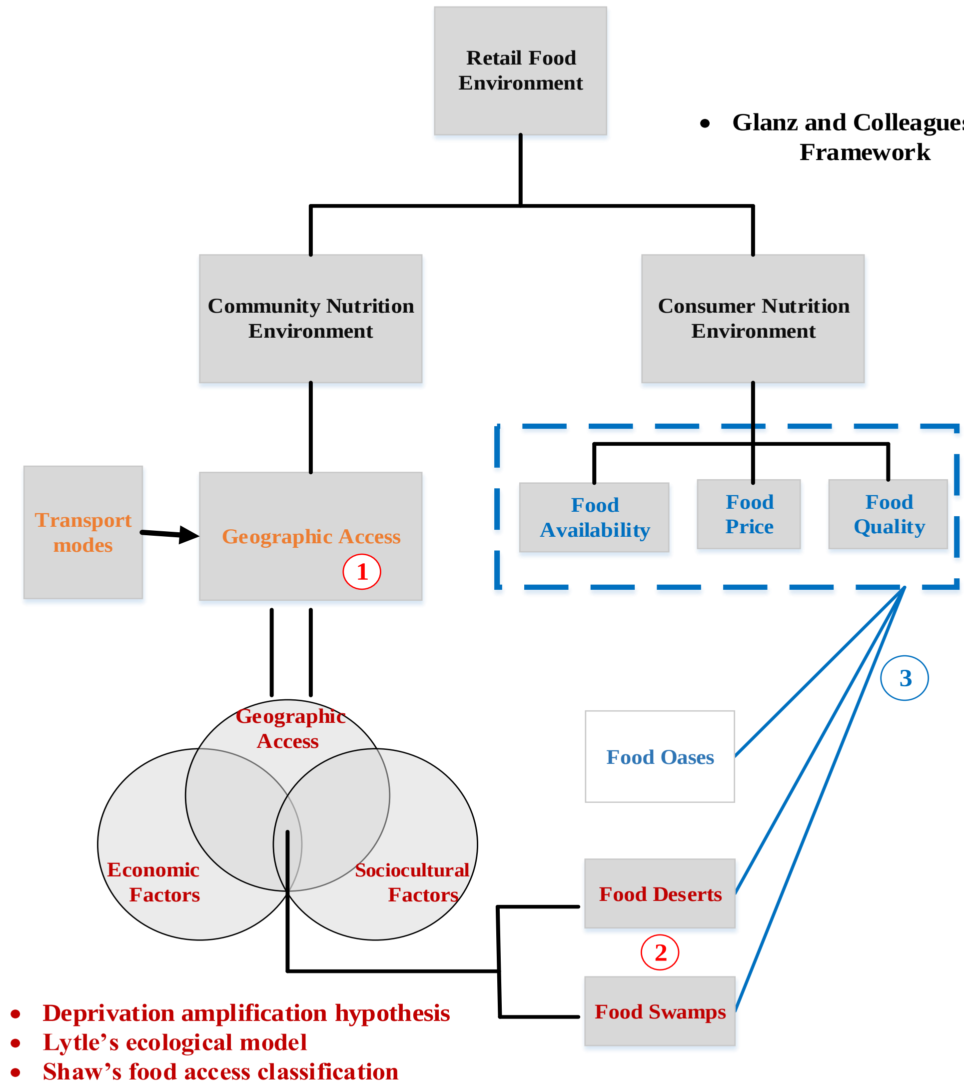


Figure 3.1 Theoretical framework of this research.

Meanwhile, the Aim 3 explores the consumer nutrition environment in food outlets. It is a "non-spatial" dimension of the food environment and can reveal distinct characteristics from the community nutrition environment. Hazardous neighborhood conditions (such as food deserts and food swamps) may impose an adverse effect on consumer nutrition environment (i.e., low healthy food availability, high food price, and poor food quality), which is also motivated by the deprivation amplification hypothesis. Aim 2 and 3 are thus connected by this hypothesis. Therefore, community level (Aim 1) and consumer level (Aim 3) of nutrition environment are indirectly related by the deprivation amplification hypothesis as well. Community nutrition environment serves as a basis to identify food deserts and food swamps. Consumer nutrition environment enriches our understanding of food deserts and food swamps in another dimension; this dimension (such as food price and quality) might be much more important than the geographic distance to food stores for people who live in the marginalized neighborhoods when they make decisions on food purchasing and food consumption. In summary, through the exploration of both community and consumer nutrition environment, this dissertation provides a thorough understanding of food access issues in Austin, Texas. It is anticipated that the findings from this dissertation would offer useful suggestions for the intervention of food deserts and food swamps in the study area.

4 EXAMINING GEOGRAPHICAL ACCESSIBILITY TO FOOD OUTLETS IN AUSTIN, TEXAS

Introduction

Retail food environment is critical to individuals' dietary behaviors and health outcomes (Glanz, et al. 2005; Luan 2016; Kuai and Zhao 2017). Food access is one of the critical components of the retail food environment (Glanz, et al. 2005). It is reported that residents with higher access to nutritious and affordable food sources have a high consumption of fruits and vegetables and low consumption of energy-dense foods (Morland and Evenson 2009; Brown and Miller 2008; Rundle, et al. 2009). On the contrary, lower access to healthy foods (i.e., grocery stores and supermarkets) and higher access to unhealthy foods (e.g., convenience stores and fast food restaurants) was linked to higher risk of overweight/obesity (Morland, Roux, and Wing 2006; Wang, et al. 2007; Powell, Chaloupka, and Bao 2007), which is believed to contribute to cardiovascular disease, stroke, diabetes, and some cancers (Bostick, et al. 1994; Chan, et al. 1994; Stamler, et al. 1978).

Spatial food accessibility measures the ease or difficulty of procuring foods for individuals or population groups in specific geographic units (Glanz, et al. 2005; Luan 2016; Kuai and Zhao 2017). The distribution of food providers (i.e., grocery stores) and consumers is usually not evenly distributed, which leads to disparities in spatial food accessibility (Wang and Luo 2005). Practices and programs have been launched to eliminate the disparities and inequities of food access (Thornton, et al. 2016). However, it remains a challenging task to equalize food access in some geographic regions, leading to that the elimination of food access disparities is still a significant public health issue (Algert, Agrawal, and Lewis 2006; Dai and Wang 2011).

A good measure of accessibility is the foundation to evaluate food accessibility disparities. In the past two decades, with the success of Geographic Information Science (GIS) — a robust system to capture, store, manipulate and manage spatial information, GIS-based spatial accessibility has been extensively explored (O'Dwyer and Burton 1998; Langford and Higgs 2006; Luo 2014). Various methods have been developed to measure GIS-based spatial accessibility (Charreire, et al. 2010; Forsyth, Lytle, and Van Riper 2010; Hilmers, Hilmers, and Dave 2012). They can be grouped into two categories: descriptive approach and modeling approach (Luan 2016). The descriptive one is simple and straightforward, and it is used widely in most research. This approach consists of density (Block, Scribner, and DeSalvo 2004; Winkler, Turrell, and Patterson 2006; Apparicio, Cloutier, and Shearmur 2007; Austin, et al. 2005; Baker, et al. 2006; Powell, Chaloupka, and Bao 2007; Moore, et al. 2008; Wang, et al. 2007; Zenk, et al. 2005b), proximity (Charreire, et al. 2010; Wang 2014; D'Acosta 2015; Larsen and Gilliland 2008; Opfer 2010; Pearce, et al. 2007; Sharkey, Horel, and Dean 2010), variety (Apparicio, Cloutier, and Shearmur 2007; Sparks, Bania, and Leete 2011), and competition (Gallagher 2007; Gallagher 2006). Despite its simplicity, the descriptive approach is subject to two problems (Luan 2016; Wan, Zou, and Sternberg 2012): (1) it assumes that individuals in the unit are equally accessible to a service site no matter how far away they are from it; (2) it also implies that people always do food shopping in their neighborhoods rather than beyond their neighborhood boundaries. The descriptive approach does not consider realistic constraints or impedance in the assessment of spatial accessibility.

By contrast, the modeling approach is more advanced and sophisticated. This approach contains time, schedules and temporal variation (spatiotemporal method (Chen and Clark 2016; Zenk, et al. 2011; Shannon 2016)), travel cost (least-cost path method (Balstrøm 2002; Sherrill, Frakes, and Schupbach 2010)), kernel density method (Guagliardo 2004), and gravity-based model (Joseph and Bantock 1982). Among these methods, the gravity-based model assumes that people's access to a service site decreases as they are far away from this site. In other words, it considers the distance-decay effect in the modeling, which is a relatively complete concept in measuring distance-based accessibility (Wan, et al. 2012).

Two Step Floating Catchment Area (2SFCA) method is a special case of the gravity-based model (Luo and Wang 2003; Luo 2014). It has been used in medical healthcare studies (Luo and Wang 2003) since it not only accounts for the distance-decay effect but also emphasizes the interactions between health services and population demands in neighboring analytical units, which formulates the competition among the population for limited resources. Luo and Wang (2003) were the first to propose the 2SFCA. However, the 2SFCA has limitation since it still assumes that all individuals within the catchment area (i.e., 30-minutes driving zone) have equal access to a service site. Many subsequent studies have proposed various methods to improve 2SFCA (Luo and Qi 2009). (Single-mode) Huff-based 2SFCA method is one of the successful modifications to the original 2SFCA. It can reveal more variability of accessibility score due to that it accounts for more realistic constraints (i.e., quantifying the probability of people's selection of a supply site with consideration of both travel cost and capacity of a supply site) in the measurement. The (single-mode) Huff-based 2SFCA method and other variants are under the 2SFCA framework and calculate a supply-to-population ratio to measure accessibility, identifying underserved areas and provide reliable evidence on interventions and resource allocation (Vo, Plachkinova, and Bhaskar 2015).

Transportation modes are important factors influencing an individual's travel capacity. For instance, a 30-minute driving distance is substantially different from a 30-minute walking distance. In the United States, 90% of households drive for food shopping. However, people who drive could be as low as 46% in some urban areas (i.e., New York City) because of the well-developed public transit systems, as well as the traffic and parking issues in cities. In addition, some minority and low-income groups cannot afford personal vehicles; they must rely on walking or public transportation. Therefore, incorporating multiple transportation modes in the measurement of accessibility is necessary; it would not overestimate the accessibility in urban areas and thus would produce a more accurate estimation. However, little efforts have been put into combining multiple transportation modes in the 2SFCA method. One exception is Mao and Nekorchuk (2013) who proposed a multi-mode 2SFCA method to measure healthcare accessibility in Florida. This method was then adopted by Kuai and Zhao (2017) to measure healthy food accessibility in Baton Rouge, Louisiana. However, it still has some drawbacks. First, it applies multi-modes to the original 2SFCA, but it does not solve the inherent problem of the 2SFCA, that is --- it assumes the equal access to a supply site for all individuals in the catchment area. Second, their methods only considered two transportation modes: driving and taking a bus. However, walking can be an alternative transportation mode if people live close to the service sites. Moreover, they assume that personal vehicles and bus travel on the same routes with the same speed, which is unrealistic in life. It could introduce errors in the measurement if driving and taking the bus are not differentiated.

I propose a multi-mode Huff-based 2SFCA method in this chapter to fill up the research gaps. On the one hand, it incorporates transportation modes into the (single-mode) Huff-based 2SFCA, which could overcome the overestimation of population demand of the Huff-based 2SFCA. On the other hand, it incorporates Huff-based model into the multi-mode 2SFCA method and adjusts the equal access issue in the catchment area of the multi-mode 2SFCA method. Meanwhile, I differentiate the routes and speeds for personal vehicles and buses in this research. Then, the proposed multi-mode Huff-based 2SFCA method is applied in the city of Austin, Texas to estimate geographic accessibility to both healthy and unhealthy food outlets. Furthermore, I compare our results between multi-mode and single-mode Huff-based 2SFCA.

The remaining of this chapter was organized in the following manner. Section 4.2 presents study area and analytic unit, followed by data source and data preparation. In section 4.3 methodology part, we briefly overview the equations for the traditional 2SFCA method and its variants. Then we propose a multi-mode Huff-based 2SFCA method. The result is present in section 4.4. The last part is the discussion and conclusion.

Study Area and Unit of Analysis

The study area is in Austin, the 4th-most populous city in Texas and the capital city of State of Texas. It seats across three counties: Travis, Hays, and Williamson Counties (see Figure 4.1). Austin has the second largest population as a state capital in the U.S, and it is the fastest growing city in the nation. Austin represents 271.8 square miles. According to the 2016 American Community Survey, approximately 947,890 people are living in Austin (American Community Survey 2016). Regarding the racial and ethnic composition, White is the predominant one (68.3%), and Hispanic or Latino (35.1%) and African American (8.1%) are minorities (American Community Survey 2016). The median household income is $ 42,689, and per capita income is $ 24,163 per year. There are approximately 9.1% of families and 14.4% of the population below the poverty line. Between 2000 and 2016, the metropolitan Austin area grew by 20 %, and approximately 157,000 new residents made it the fastest growing metro area in the nation.

I used block group as the unit of analysis for this study. Using census block can better represent a finer scale for analysis. However, demographic and socioeconomic data are not available at this level for privacy and confidentiality concern. Block group is the smallest unit that the U.S. Census Bureau tabulates sociodemographic data. Census block groups are more similar in regards to population characteristics than census tracts and usually contain between 600 and 3,000 people or 240 to 1,200 housing units, depending on population density (Iceland and Steinmetz 2003). Most of the literature reviewed in Chapter Two were at the census tract level, and the use of the smaller block groups instead of census tracts in this study may increase the precision of the results and mitigate the potential modifiable areal unit problem (MAUP). The MAUP is a common problem in geographic analyses (Wong 2004). It refers to a statistical bias that leads to different results during the aggregation of data from one unit to another unit. For instance, a method using data aggregated by the county will produce results that differ from the same process using data aggregated at the state level. Therefore, it is vital to examine food access at an analytic scale as finer as possible to mitigate the potential MAUP. There are 555 census block groups in the study area.

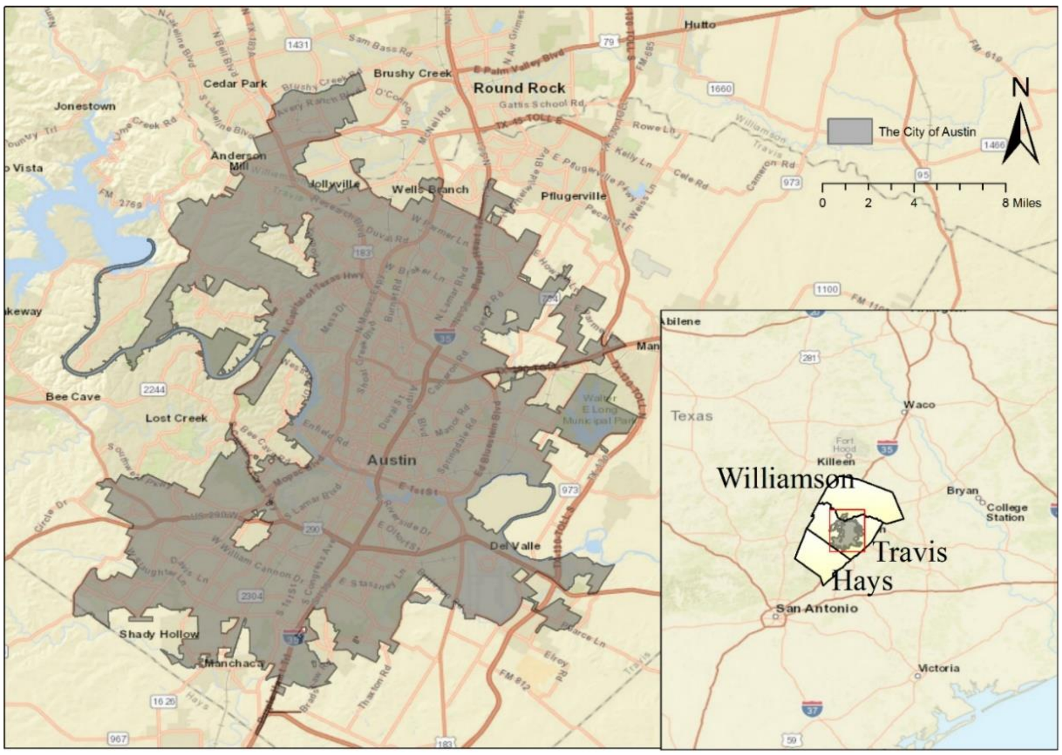


Figure 4.1 Study area — 555 census block groups in Austin, Texas.

Data Source and Data Preparation

Food store and restaurant data

Food store and restaurant data were collected from ReferenceUSA, which is an online database of business that offers the complete coverage of business establishments in the US and Canada[15]. This website is one of the most comprehensive data sources and provides up-to-date data for 50 million businesses in the United States. The ReferenceUSA has been subscribed by Texas State University Alkek Library and is ready to use. The retail food stores include healthy food sources — supermarkets, grocery stores, supercenters, and specialty food stores (meat and fresh fruit vegetable markets), and unhealthy sources such as convenience stores and fast food restaurants. Supermarkets are defined as large, corporate-owned “chain” stores, while grocery stores are defined as small non- corporate-owned food outlets. Supermarkets and grocery stores are the reliable providers of healthy food because they consistently have greater variety and availability of healthy food options than other stores (Glanz, et al. 2007). Supercenters such as Walmart are also considered as reliable, healthy food sources (Krukowski, et al. 2010). Some other studies included specialty stores as sources providing healthy food options (Moore and Diez Roux 2006; Walker, Keane, and Burke 2010). In contrast, convenience stores and limited-service fast food restaurants are classified as unhealthy food outlets since they primarily carry processed foods and high caloric food that do not meet people’s nutrition needs (Glanz, et al. 2005; Saelens, et al. 2007). The search for food stores in ReferenceUSA using North American Industry Classification System (NAICS) codes and the NAICS indices are seen in Table 4.1.

[15] See http://resource.referenceusa.com/

Table 4.1 2017 North American Industry Classification System (NAICS) codes of food stores.

Industry Group	**NAICS Definition**	**NAICS Index**	**Examples**
Convenience Stores	445120 Convenience Stores	445120	7-Eleven, Circle K, Corner Store, Murphy USA
Fast Food Outlets	722513 Limited-Service Restaurants	722513	Burge King, Domino's, Mc Donald's, Chick-Fil-A, KFC
Supermarkets & Grocery stores	445110 Supermarkets and Other grocery stores (except Convenience stores)	445110	H-E-B Foods, Whole Foods Market, Randall's, Trader Joe's, Man Pasand Grocery, Natural Grocery
Supercenters	452311 Warehouse Clubs and Supercenters	452311	Sam's Club, Costco Wholesale, Walmart
Specialty Food Stores	4452 Specialty Food Stores:		
	445210 Meat Markets	445210	Discount Meats, University Meat Market
	445220 Fish and Seafood Markets	445220	Cawoods Produce
	445230 Fruit and Vegetable Markets	445230	LA fruta Feliz, Brothers produce of Austin
	445291 Baked Goods Stores	445291	
	445299 All others	445299	Yogurt land, Pinkberry

The stores requested from ReferenceUSA were geocoded to obtain their X and Y coordinates. As Zhan et al. claimed, two different datasets are required for the geocoding; one is an address record table, and the other one is a reference street network database. Each store address has four required attributes: street address, city name, state name, and ZIPcode. USA Geocoding Service was used as the reference street network database because it contains dual range street address and location information, which can be useful to determine the side of the street segment where the fast food chains are located. The geocoding was performed in ArcGIS 10.5. The mean geocoding matching scores are all high, ranging from 95.809% to 97.50% (see Table 4.2). The geocoded stores were initially in a geographic coordinate system (WGS 1983), which is not suitable for precise distance and size measurements. I then projected them to a coordinate project system — NAD 1983 UTM 14N, precise preservation of distance and shape for the study area in the North-South orientation such as Austin.

Table 4.2 The number of each type of food stores in the study area.

Types of stores	**Number of stores**	**Geocoding accuracy (Mean ± SD)**	**Total number of stores**
Heathy Food Source			156
Supermarkets & Grocery stores	101	96.128% ± 3.097%	
Supercenters	14	97.500% ± 5.408%	
Specialty Food Stores		97.021% ± 3.514%	
Meat Markets	14		
Fish and Seafood Markets	1		
Fruit and Vegetable Markets	16		
Baked Goods Stores	0		
All Others	10		
Unhealthy Food Source			811
Convenience Stores	245	96.057% ± 3.851%	
Fast Food Outlets	566	95.809% ± 4.073%	
Total			967

Van Meter, et al. (2010) found considerable bias at the edge of study areas for accessibility measures, and they thus recommended that any study involving accessibility measures should correct for the edge effects. To do this, I created a 2000-meters buffer around the city of Austin boundary. Any stores within this buffer zone are considered in my study. Then I clipped each type of stores into Austin buffer zone. It eliminates any store that does not fall within the buffer zone. Based on this criterion, 245 convenience stores and 566 fast food outlets were identified and successfully geocoded (Table 4.2). Meanwhile, 156 healthy food stores were in the Austin buffer zone. Of these 156 stores, there are 101 supermarkets and grocery stores, 14 supercenters, 14 meat markets, one fish and seafood markets, 16 fruit and vegetable markets, and ten other specialty food stores (Table 4.2).

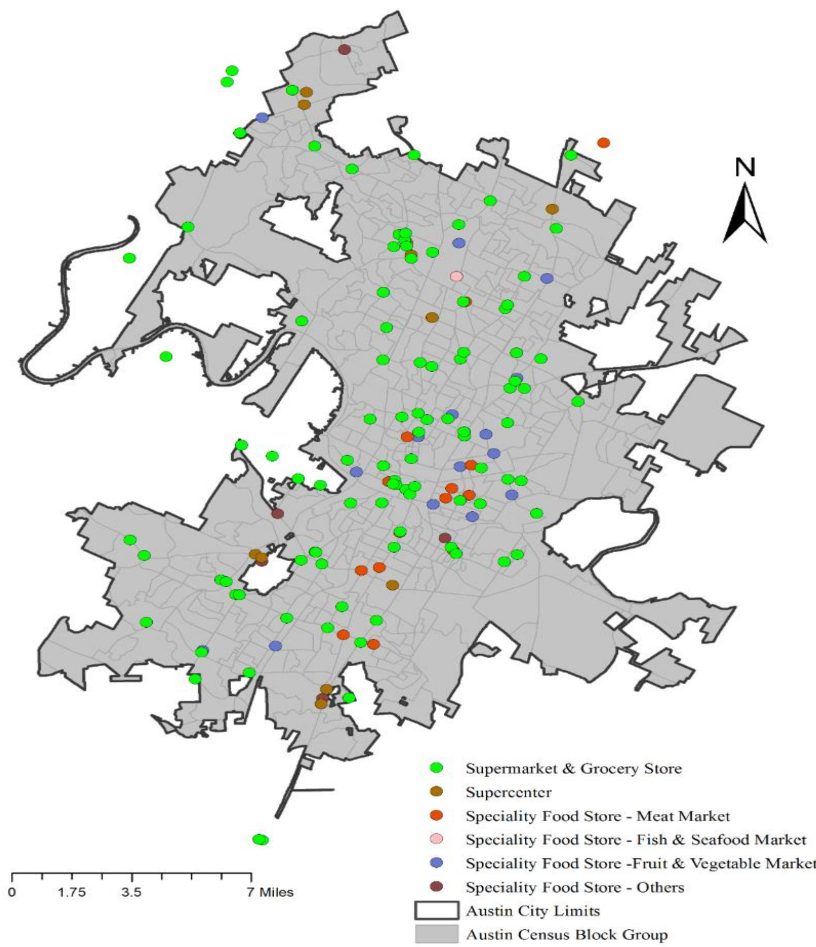


Figure 4.2 Spatial distribution of healthful food stores in the study area.

Healthy food stores are mainly distributed in the urban center along the highway IH-35 (Figure 4.2). Note that some streets have more than one store in competition and these locations may be stacked on the map. Figure 4.3 is the spatial distribution of unhealthy food stores in the 2000- meters buffer zone of Austin. Convenience stores and fast food restaurants are also located along the IH-35.

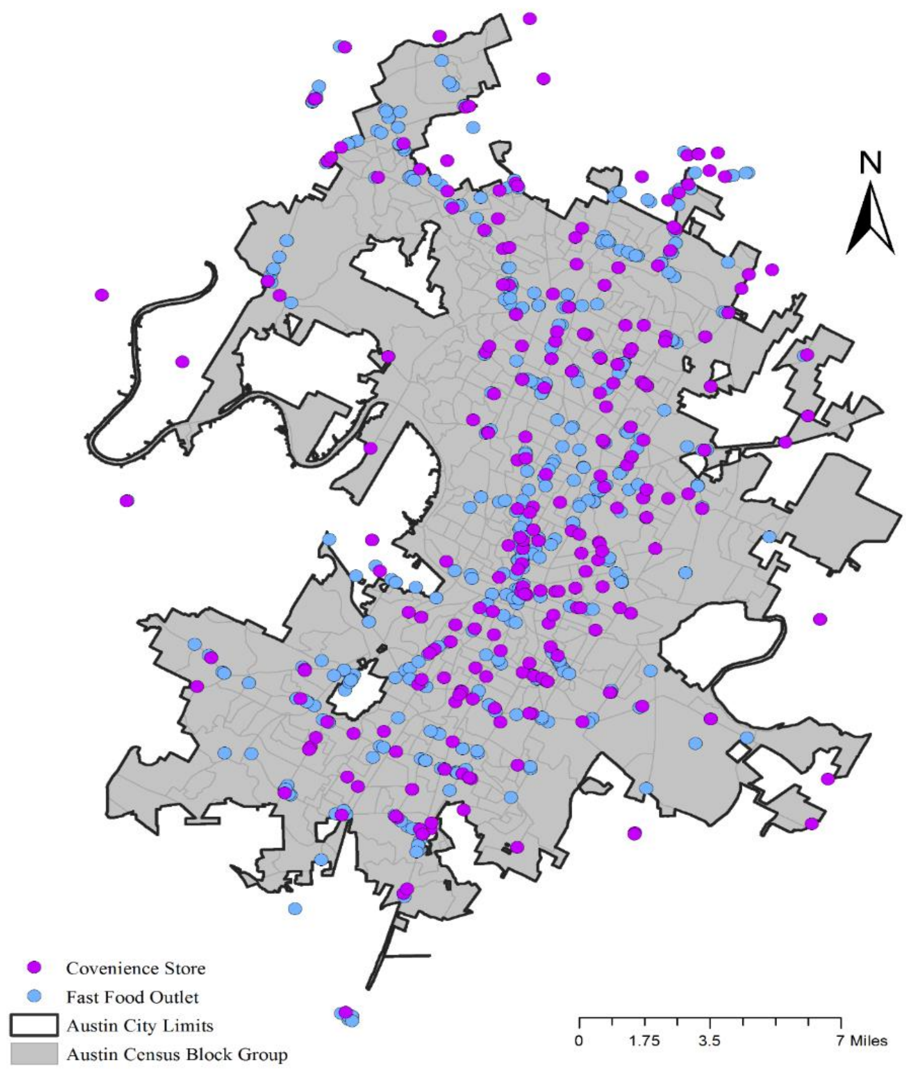


Figure 4.3 Spatial distribution of unhealthful food stores in the study area.

The food store dataset obtained from the ReferenceUSA has a “sales volume range” column attached to each record. According to Kuai and Zhao (2017), food stores' business capacity could be estimated by the logarithm of food stores' sales volume. Even though the sales volume range is not a specific number, I could assign a number to it. For instance, the sales volume range “Less than 0.5 million” is assigned “500,000”. Other assignments are shown in Table 4.3. The capacity (logarithm of sales volume) for healthy and unhealthy food stores is depicted in Table 4.3.

Table 4.3 Business capacity of food stores in the study area.

Sales volume range	**Sales volume**	**Store capacity**	**# healthy food stores**	**# unhealthy food stores**
Less than 0.5 million	500,000	5.69	15	128
0.5 ~ 1.0 million	1,000,000	6.00	9	234
1.0 ~ 2.5 million	2,500,000	6.39	36	385
2.5 ~ 5 million	5,000,000	6.69	13	61
5 ~ 10 million	10,000,000	7.00	11	1
10 ~ 20 million	20,000,000	7.30	8	2
20 ~ 50 million	50,000,000	7.69	35	NA
50 ~ 100 million	100,000,000	8.00	24	NA
100 ~ 500 million	500,000,000	8.69	5	NA

Public transit and street centerline data

Public transit and street centerline data are needed to create a network database. General Transit Feed Specification (GTFS) data is used to develop public transit routes and calculate travel time between transit stops. Street centerline is used to create a road network for people who drive or walk.

GTFS, the acronym of General Transit Feed Specification (GTFS), is formerly known as Google Transit Feed Specification. The original purpose of GTFS is to create a transit trip planner. Google collaborated with TriMet to launch the first Google Transit

Trip Planner in Portland, Oregon (Ma and Jan-Knaap 2014). Since then, more and more cities have participated in Google Transit Trip Planner. Up to December 2017, there are about 400 GTFS available in the United States. Before the advent of GTFS, there was no standard transit format (i.e., transit schedule, directory, and timetable) available to use, making it difficult for developers to develop transit-based applications. Currently, hundreds of public transit agencies worldwide voluntarily publish their up-to-date GTFS data and share it with the public.

GTFS is a standardized data format to store public transit routes, stops, and schedules (Antrim and Barbeau 2013). GTFS contains a set of feed files that consists of transit information: (1) administration information, such as operator agency and service calendar; (2) spatial information, including the location of stops, timing of stops, and routes; (3) schedules such as trips and stop times; and (4) optional information, including fares, calendar dates, shapes, frequencies, and transfers. The specific GTFS definitions can be found in Table 4.4. Austin GTFS was obtained from the City of Austin website[16].

The potential use of GTFS in transit accessibility is well recognized in food studies. For instance, Widener and colleagues studied the spatiotemporal access to supermarkets using GTFS public transit data in various places at the daily scale (Widener, Metcalf, and Bar-Yam 2011; Widener, et al. 2013; Widener, et al. 2015). The GTFS data can be added to ArcGIS to form a transit-travel-time matrix, which allows researchers to investigate the food environment dynamically based on the time of day. The integration

[16] https://data.austintexas.gov/Transportation/GTFS-June-2016/hmh7-7zmg

of time and schedules into food studies can provide a complete and realistic picture of food environment rather than solely using spatial information.

Table 4.4 The definitions of GTFS feed files, adapted from Google Developers, 2015[17].

File Name	Required	Definition
agency.txt	Required	Transit agencies that provide the GTFS data
stops.txt	Required	Individual locations where vehicles pick up or drop off passengers.
routes.txt	Required	A route is a group of trips that are displayed to riders as a single service.
trips.txt	Required	A trip is a sequence of two or more stops that occurs at specific time.
stop_times.txt	Required	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	Required	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.
calendar_dates.txt	Optional	Exceptions for the service IDs defined in the calendar.txt file.
fare_attributes.txt	Optional	Fare information for a transit organization's routes.
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes.
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes.
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service.
transfers.txt	Optional	Rules for making connections at transfer points between routes.

Street centerline shapefile. Austin street centerline shapefile was obtained from the City of Austin as well. This shapefile has some critical fields such as Road Type (Code) and Speed Limit (Table 4.5). I could add a field called "minutes" to calculate travel time

[17] https://developers.google.com/transit/gtfs/reference/

on each road segment using Shape Length divided by Speed Limit. This "minutes" filed will be used as travel cost in the creation of a road network. The "elevation" filed has binary integers: 1 and 0. If two coincident endpoints have elevation values of 1, the edges connect. However, if one endpoint has a value of 1, and the other coincident endpoint has a value of 0, the edges don't connect. The filed "one-way" has three values: "B", "FT", and "TF". "B" indicates that the road is allowed to travel on both directions, whereas “FT” and “TF” mean that that road is restricted on one way, either “From-To” or “To-From” direction. These two fields will be utilized as a “One Way” restriction in the network dataset.

Table 4.5 Road type in Austin Shapefile and speed limit for each road type.

Road Type Code	Road Type	Speed Limit (Miles/Minute)
0	Category Unknown	25
1	Interstate, Fwy, Expy, Toll	70 , 65
2	US and State Highways	65, 60, 50
4	Major Arterials and County Roads	45, 40
5	Minor Arterials	40
6	Local City/County Street	30 or 25
8	City Collector	35
10	Ramps and Turnarounds	60, 50, 30
12	Driveway	25
14	Unimproved Public Road	25
15	Private Road	30, 25
16	Routing Driveway/Service Road	5
17	Platted Row/Unbuilt	1

Creating multiple-modal transportation network. This study used network distance rather than Euclidean distance. The proposed model needs to incorporate multimode transportation. Thus, it is necessary to create a multiple-mode network database in ArcGIS Network Analyst extension. I consider three travel modes: driving, public transit, and walking. Other methods such as biking, motorcycling, and taking a taxi or cab are out of our consideration.

The obtained GTFS text file contains the transit service on June 1st-30th, 2016. Melinda Morang and her team at ESRI developed a toolkit — Add GTFS to a Network Dataset, which can convert GTFS text file to transit routes and combine GTFS data with street centerline data into an ArcGIS network dataset. One can refer to Add GTFS Data to a Network Dataset Users' Guide[18] for more information. Figure 4.4 illustrates the process of creating a multi-modal network in ArcGIS 10.5. It contains three steps, and one can refer to Appendix A for more details about the three steps. The generated transit stops and transit route segments are seen in Figure 4.5.

Population with transportation modes

Transportation modes are incorporated into physical accessibility measurement. The accessibility measure differs between the population groups with different transportation means. The distribution of transportation means among the population for grocery shopping is difficult to measure, but I could use the number of people who take different transportation forms to work as an estimation. The data of transportation means to work consists of different modes of commuting forms such as driving, public transit, and walking. I obtained transportation modes data from the 2016 American Community Survey (ACS) 5-year estimations. Three modes were included: public transit, walk and drive. The number of the population aged 25 to 64 years using the three transportation

---

17 https://esri.github.io/public-transit-tools/AddGTFStoaNetworkDataset.html

modes was extracted for the analyses. As illustrated in Figure 4.6, most people in the suburban areas and the periphery of Austin use driving as their transportation means, but the neighborhoods in the urban center are most likely to commute by walking and taking public transit.

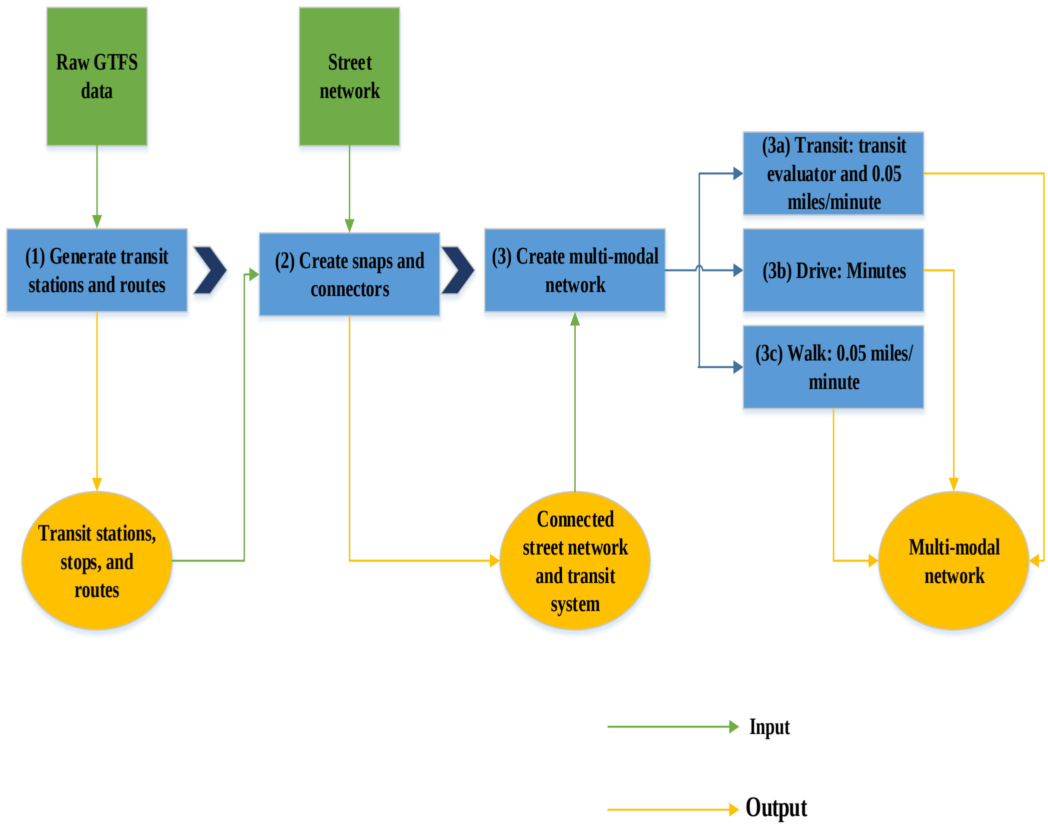


Figure 4.4 Steps to create a multi-modal network using GTFS text file and street shapefile.

Note: This figure is modified based on Ma and Jan-Knaap (2014, 8)).

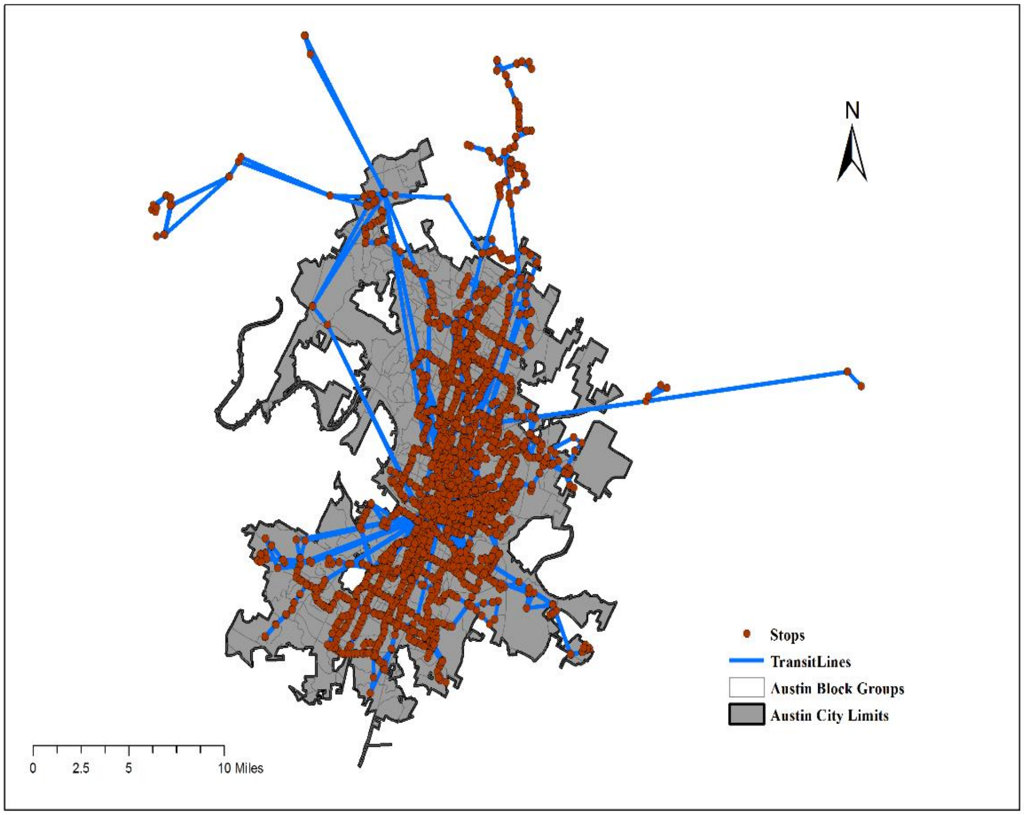


Figure 4.5 Transit stops and route segments were generated by GTFS data in Austin, Texas.

Shapefile data

Shapefiles data come from different sources (Table 4.6). The last row of Table 4.6 is the data source for population in each Census Block. I will use this data to generate the population-weighted centroid for each census block group. The ACS, unfortunately, does not provide population data for census blocks. I have to download it from the US Census Bureau TIGER/Line. However, this data is available at the census block level for the year 2010. The data in the year 2016 is not available. Despite this mismatch, I determined to utilize the data in the year 2010 because it is the best data that could be obtained.

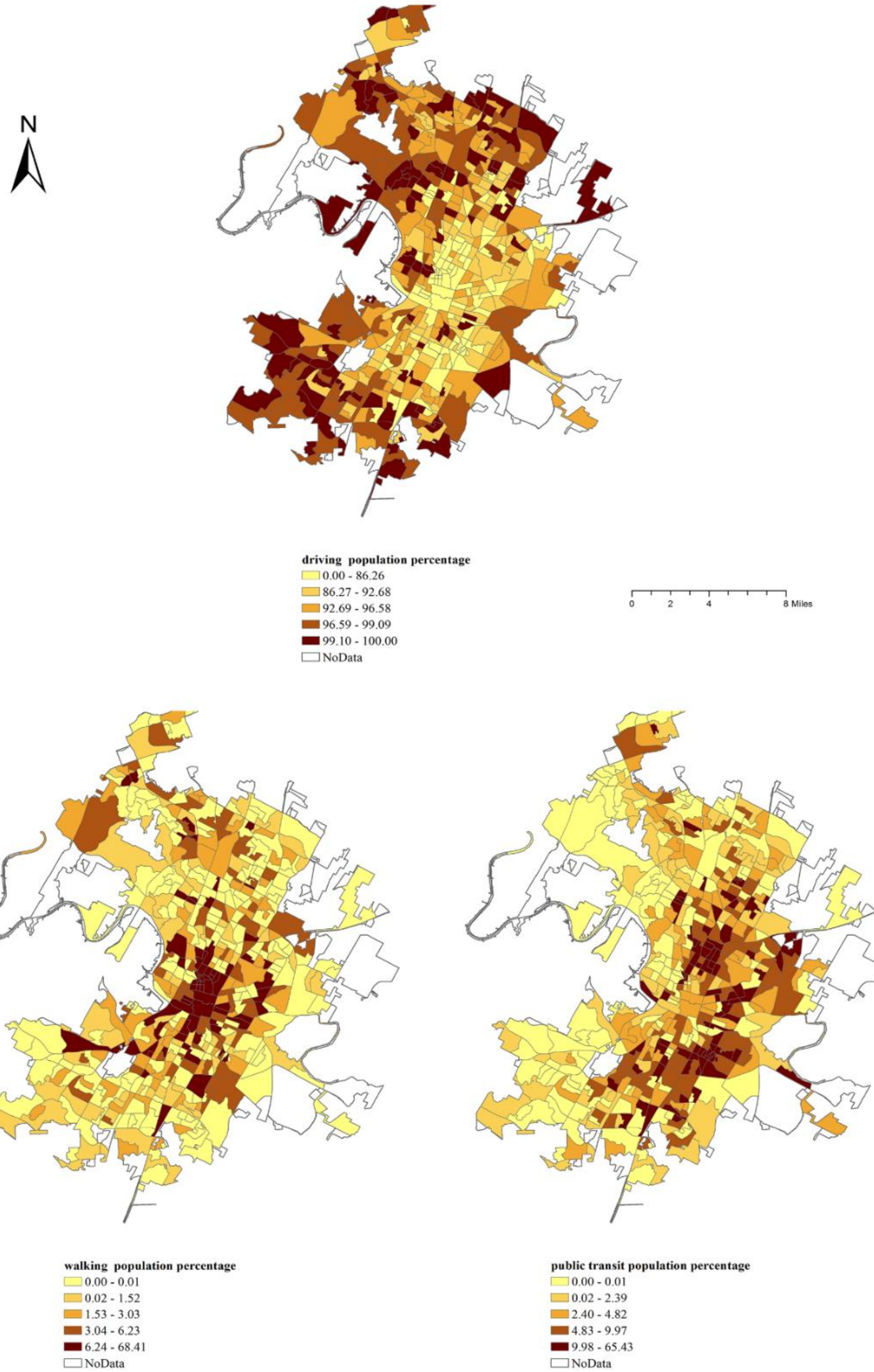


Figure 4.6 Spatial distribution of population percentage with transportation modes.

Table 4.6 Summary of shapefile data in this research.

Dataset	File type	Data type	Source	Year of data
Austin Census Block Group	Shapefile	Polygon	US Census Bureau	2016
Austin City Limits	Shapefile	Polygon	Texas Department of Transportation	2016
Austin Census Block with Population	Shapefile	Polygon	US Census Bureau	2010

Methodology

The unit of the analysis is at the block group level. Since the geographic centroid of each block group may not represent where the majority of the population resides, population-weighted centroids are used as the reference point of each block group. I generated populated-weighed centroids of each block group using (Luo and Wang 2003) method. The equations are shown below.

$$X = \frac{\sum_{i=1}^{n} P_i X_i}{\sum_{i=1}^{n} P_i} \quad \text{4-1}$$

$$Y = \frac{\sum_{i=1}^{n} P_i Y_i}{\sum_{i=1}^{n} P_i} \quad \text{4-2}$$

where X and Y denote the coordinates of population-weighted centroid, X_i and Y_i represent the x, y coordinates of each block's geographic mean center, and P_i is the census block population.

The equations above could be implemented in ArcGIS operations. Simply perform the following techniques to calculate the population-weighted centroid of each block group. Open the “Mean Center” tool, put the 2010 census block shapefile as the input dataset, set the population of each block as weight field, and block group ID as case field.

Measuring spatial accessibility to food outlets

Traditional 2SFCA method and family. This study proposes a multi-modal Huff-based 2SFCA method. It incorporates multiple transport modes into Huff-based 2SFCA method. Nevertheless, 2SFCA, *E2SFCA, 3SFCA, and Huff-based 2SFCA* methods are essential to understanding the proposed model. Their equations and formulas are discussed below.

<u>2SFCA:</u> It is the foundation of two-step floating catchment area family method. It has two critical steps. First, each supply site j searches all demand sites (k) in a catchment area d_0, and summarize the population with in the catchment area d_0 for store j. Then calculates the supply-to-demand ratio R_j within the catchment area (d_0):

$$R_j = \frac{S_j}{\sum_{k \in \{d_{kj} \leq d_0\}} P_k} \qquad 4\text{-}3$$

where R_j is the supply-to-demand ratio at supply site j that falls within the predefined catchment area d_0; S_j is the capacity of supply at site j, d_{kj} is the travel time between site k and j, and P_k is the population demand at site k that falls within the catchment ($d_{kj} \leq d_0$).

Second, each demand site i searches all supply sites (j) that are within the catchment area d_0 and sum up the supply-to-demand ratio R_j.

$$A_i^F = \sum_{j \in \{d_{ij} \leq d_0\}} R_j = \sum_{j \in \{d_{ij} \leq d_0\}} \frac{S_j}{\sum_{k \in \{d_{kj} \leq d_0\}} P_k} \qquad 4\text{-}4$$

where A_i^F is accessibility at location i; d_{kj} or d_{ij} is the travel time between location i (or k) and j.

E2SFCA: The 2SFCA method does not account for distance decay in the catchment d_0. To overcome this deficiency, E2SFCA divides the catchment area into three drive time zones (0-10, 10-20, and 20-30 minutes) and applies different weights for these zones in both steps. The supply-to-demand ratio R_j can be rewritten as:

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\leq d_r\}} P_k W_r} \quad \text{4-5}$$

where S_j is the capacity of supply at site j, d_r is the rth drive time zone; P_k is the population demand at site k that falls within the rth catchment ($d_{kj} \leq d_r$); W_r is the impedance weight of the rth time zone, which is a Gaussian function of travel time.

The accessibility at the second step is the summation of weighted supply-to-demand ratio R_j within the catchment area. Its equation is:

$$A_i^F = \sum_{j\in\{d_{ij}\leq d_r\}} R_j W_r == \sum_{j\in\{d_{ij}\leq d_r\}} \frac{S_j W_r}{\sum_{k\in\{d_{kj}\leq d_r\}} P_k W_r} \quad \text{4-6}$$

where A_i^F is accessibility at location *i;* other notations remain the same as for Eq. 4-5.

3SFCA: The E2SFCA method does not consider the potential competition among multiple supply sites available for a population site. A *3SFCA* method can adjust this problem. At the first step, a selection weight (G) is computed for all population sites *i* and supply sites *j* parings. The G is calculated by the weight between site i and j (W_{ij}) divided by the summation of all W values for supply sites within population i's catchment(d_0). The equation is:

$$G_{ij} = \frac{W_{ij}}{\sum_{s\in\{d_{is}\leq d_0\}} W_{is}} \quad \text{4-7}$$

where G_{ij} is the selection weight of population i on supply j; W_{ij} and W_{is} represent the assigned weight for supply sites j and s; d_{is} is the travel time from population site i to any supply site s, and d_0 is the catchment of population i.

The second step of 3SFCA incorporates the selection weight G into the Eq. 4-5 to calculate the supply-to-demand ratio of a supply site:

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\leq d_r\}} G_{kj} P_k W_r} \qquad 4\text{-}8$$

where S_j is the capacity of supply at site j, d_r is the rth drive time zone; G_{kj} is the selection weight of population k on supply j; P_k is the population demand at site k that falls within the rth catchment ($d_{kj} \leq d_r$); W_r is the distance impedance weight of the rth time zone, and it is a Gaussian function of travel time.

The third step of the 3SFCA also incorporates the selection weight G into the Eq. 4-6 to calculate the spatial access to supply sites for a population site i.

$$A_i^F = \sum_{j\in\{d_{ij}\leq d_r\}} G_{ij} R_j W_r \qquad 4\text{-}9$$

where A_i^F is accessibility at location *i;* other notations remain the same as for Eq. 4-8.

<u>(Single-mode) Huff-based 2SFCA</u>: The 3SFCA method calculates the probability of people's selection (the selection weight G) only considering the travel time (or cost). Huff model quantifies the probability of people's selection on a supply site with considerations both travel cost and capacity of the supply site. The equation of Huff model is:

$$\mathrm{Prob}_{ij}^{H} = \frac{S_j d_{ij}^{-\beta}}{\sum_{s\in\{d_{is}\leq d_0\}} S_s d_{is}^{-\beta}} \qquad 4\text{-}10$$

where $Prob_{ij}^{H}$ is the probability of population location i visiting supply site j based on Huff model; s is any supply site within the catchment d_0; β is the travel time impedance coefficient.

The first step of Huff-based 2SFCA method is to utilize $Prob_{kj}^{H}$ and a continuous negative power distance weight W_{kj} to replace G_{kj} and W_r in the Eq. 4-8. The equation can be rewritten as:

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\leq d_0\}} Prob_{kj}^{H} P_k W_{kj}} \qquad \text{4-11}$$

where R_j is the supply-to-demand ratio of supply site j within the catchmentd_0; W_{kj} is the inverse power impedance weight between k and j.

The second step is to summarize R_j at all supply sites within the catchment area d_0. Similar to the second step, it replaces G_{ij} and W_r in the Eq. 4-9 by $Prob_{ij}^{H}$ and W_{ij}, respectively. The equation is:

$$A_i^F = \sum_{j\in\{d_{ij}\leq d_0\}} Prob_{ij}^{H} R_j W_{ij} \qquad \text{4-12}$$

where A_i^F is accessibility at location *i;* W_{ij} is the inverse power impedance weight between demand site i and supply site j; other notations remain the same as for Eq. 4-11.

Proposal of a novel method — Multi-mode Huff-based 2SFCA. Inspired by Mao and Nekorchuk (2013), I seek to improve the (single- mode) Huff-based 2SFCA model by incorporating multiple transport modes into it. This new method is named as multi-mode Huff-based 2SFCA, which still complies with the framework of 2SFCA methods. It utilizes different means of transportation as weights and then assigns the weight for each mode of transportation in the calculation of the supply-to-demand ratio and the spatial accessibility index. The proposed method was implemented in the following three steps.

First, the multi-mode Huff-based 2SFCA calculates the probability of people's selection on a supply site for different transportation modes. It considers both travel cost and capacity of the supply site simultaneously. The calculation resembles the one calculated in the Huff-based 2SFCA method. The difference is that the proposed method incorporates n ($n \geq 1$) transportation modes $\{M_1, M_2, \ldots M_n\}$ into the Equation 4-10. As a result, the equation is updated as:

$$\mathrm{Prob}_{\mathrm{ij,M_n}}^{\mathrm{H}} = \frac{\mathrm{S_j} * (\mathrm{d_{ij,\,M_n}})^{-\beta}}{\sum_{\mathrm{r} \in \{\mathrm{d_{ir,M_n}} \leq \mathrm{d_{0,\,M_n}}\}} \mathrm{S_r} * (\mathrm{d_{ir,\,M_n}})^{-\beta}} \qquad 4\text{-}13$$

where $Prob_{ij,M_n}^{H}$ is the probability of population location i visiting supply site j based on Huff model by transportation mode M_n; d_{ij,M_n} or d_{ir,M_n} is the travel time between i and j(or r) by transportation mode M_n; d_{0,M_n} is the predefined travel catchment defined by transportation mode M_n ; r is any supply site within the catchment d_{0,M_n} ; and β is the travel time impedance coefficient.

Second, the supply-to-demand ratio $of\ R_j$ is calculated. At this step, n transportation modes $\{M_1, M_2, \ldots M_n\}$ is incorporated into the Eq. 4-11. Correspondingly, the population at location k is divided into n subpopulations by transportation modes $\left\{P_{k,M_1}, P_{k,M_1}, \ldots P_{k,\,M_n}\right\}$ (Mao and Nekorchuk 2013); the probability of people at population k selecting a supply site j is updated by transportation modes $\left\{Prob_{kj,M_1}^{H}, Prob_{kj,M_2}^{H}, \ldots Prob_{kj,M_n}^{H}\right\}$. Therefore, Eq. 4-11 is rewritten as:

$$R_j = \frac{S_j}{\sum_1^n \sum_{k \in \{d_{kj,M_n} \leq d_{0,M_n}\}} Prob_{kj,M_n}^{H} P_{k,M_n} W_{kj,M_n}} \qquad 4\text{-}14$$

where d_{kj,M_n} is the travel time by the transportation mode M_n between location k and j; d_{0,M_n} is a predefined threshold travel time from j by mode M_n; $Prob^{H}_{kj,M_n}$ is the Huff-model based selection probability for population at k who visit the supply site j by mode M_n; W_{kj,M_n} is an inverse power impedance weight between k and j by mode M_n.

Lastly, the overall accessibility A_i at a population site is computed. The R_j calculated in the second step at all supply sites by different transportation modes within the catchment area $d_{0,\,M_n}$ is summarized. Instead of directly adding all R_j within a catchment area, the multi-mode method weights R_j of each facility by the size of subpopulation as per the catchment area(s) it falls within. Then, it sums the weighted values to calculate the overall accessibility (Ai) of the population i. The spatial accessibility A_i should be the weighted average of accessibility of n subpopulation groups. The equation is:

$$A_i = \frac{\sum_1^n P_{i,M_n} \sum_{j \in \{d_{ij,M_n} \leq d_{0,M_n}\}} Prob^{H}_{ij,M_n} R_j W_{ij,M_n}}{\sum_{v=1}^{n} P_{i,M_v}} \qquad 4\text{-}15$$

where A_i denotes spatial accessibility to supply sites for each block group *i*; P_{i,M_n} is the population at location *i* by transportation mode M_n; other notations remain the same as for Eq. 4-14.

Defining travel time thresholds

In the spatial accessibility equations, it is critical to define thresholds of travel time t_0 for each transport mode. For walking mode, the 2009 National Household Travel Surveys revealed that the median walking time duration in the U.S. population was 10 minutes (Yang and Diez-Roux 2013). Therefore, this study uses 10 minutes as the threshold for walking. The average value for urban American's driving time to the grocery store is approximately 15 minutes. This study thus set 15 minutes as the threshold for driving. For the public transit mode, Kuai and Zhao (2017) used 15 minutes as the threshold to examine healthy food accessibility in Baton Rouge, Louisiana. However, they assumed that the travel time by transits is same as by vehicles. In this study, I used the TransitEvaluator to calculate transit travel time; it recalls the schedules and trips, waiting time, and ride time between stops in GTFS. For this reason, the travel time by transits should be longer than by vehicles. Dai and Wang (2011) used Kernel Density 2SFCA (KD2FCA) to measure food store accessibility in southwest Mississippi using 30 minutes threshold. Therefore, I used 30 minutes as a cut-off value for public transit mode. Note that the transit time of 30 minutes could indicate 30 minutes of walking, 30 minutes of riding on transit, or any combination of walking, waiting, and riding that adds up to 30 minutes.

Implementation of multi-mode Huff-based 2SFCA method

The proposed method was implemented in ArcGIS 10.5. Network Analyst OD Cost Matrix solver was used to calculate the travel time of each mode for each census block/ food store pair. The OD matrix uses Dijkstra's algorithm to find the shortest path through the network. The OD Time tables of the three modes were generated, respectively. The predefined travel time thresholds for the three modes (i.e., 15-min driving, 10-min walking, and 30-min public transit) were used. For each population site, all the supply locations within its catchment area by transportation mode were identified and joined to the population site catchment layer. The one-to-many relationship (i.e., one population site to many supply locations) was then established, and drive-time properties were joined to calculate the Huff-based selection probability of a population location on supply sites within its catchment by different transportation modes. The calculation involves two factors: (1) the business capacity of a supply site; (2) the inverse power drive time weight ((travel time) ^(-β)) by transportation modes.

The second step computes the supply-to-demand ratio for each of the supply sites in the study area using Eq. 4-14. For each supply site, all the population locations within its catchment area by transportation mode were identified and joined to the supply site catchment layer. The one-to-many relationship (i.e., one supply site to many population locations) was then established, and drive-time properties were joined to calculate the supply-to-demand ratio. The population demand was further adjusted by three factors: (1) the Huff-based selection weight; (2) the subpopulation groups by different modes; and (3) the inverse power drive time weight ((travel time) ^ (-β)) by transportation modes.

The last step of the analysis sums the supply-to-demand ratios of each population location to calculate the accessibility using Eq. 4-15. The overall accessibility was also adjusted by three factors: (1) the Huff-based selection probability of a population location on a supply site; (2) the subpopulation groups by different modes; and (3) the inverse power drive time weight ((travel time) ^(-β)) by transportation modes.

It can be seen that the three steps all contain the impedance coefficient β. That is to say, the choice of impedance coefficient β is vital to the calculation of spatial accessibility. Mao and Nekorchuk (2013) incorporated transportation modes into the 2SFCA method and used six coefficients suggested by ESRI ranging from 1.5 to 2.0. A range of coefficients (i.e., from 1.2 to 2.2 with an increment of 0.1) was employed to conduct comparative analysis for the results from these 11 coefficients.

Comparison analysis between multi-mode and single-mode Huff-based 2SFCA

Multi-mode and single-mode Huff-based 2SFCA methods were implemented in 476 Austin block groups. I compared the results of the two methods. First, a paired t-test was utilized to assess whether there is a significant difference between the two methods. Second, a scatterplot was drawn to illustrate the trend of underestimation or overestimation by the results of these two methods. Third, the relative difference between the two accessibility measures was computed to examine the magnitude and direction of the difference, as suggested by Mao and Nekorchuk (2013). The equation of the relative difference for each block group is shown in the below.

$$\mathrm{RD} = \frac{\mathrm{A_i^F(m)} - \mathrm{A_i^F(s)}}{\mathrm{A_i^F(s)}} * 100 \qquad \text{4-16}$$

where $A_i^F(m)$ and $A_i^F(s)$ are the spatial accessibility score for multi-mode and single-mode Huff-based 2SFCA methods, respectively.

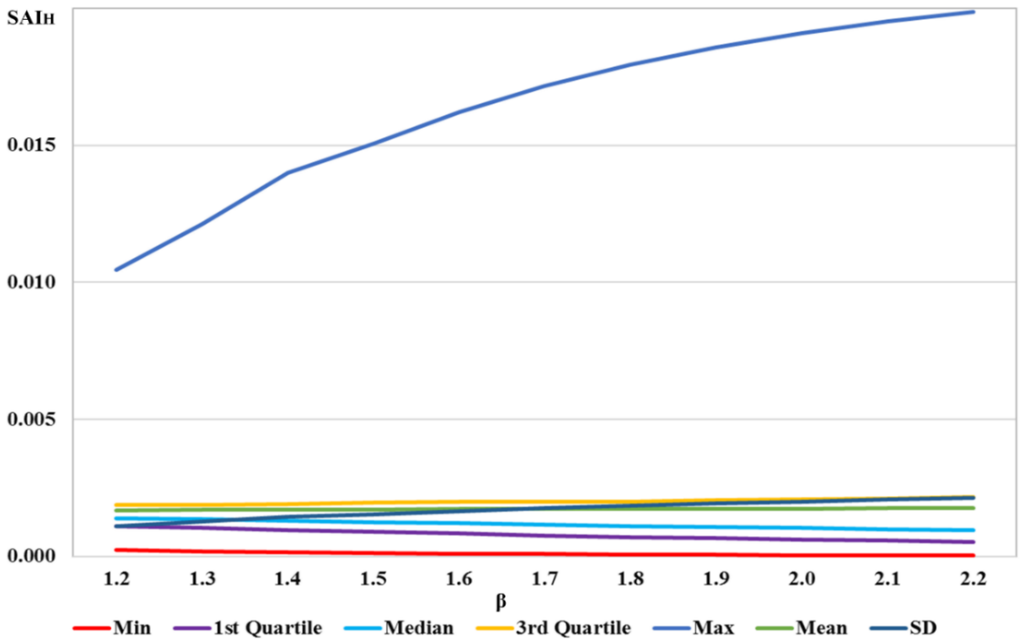
Result

Spatial access to healthy food outlets using multi-mode Huff-based 2SFCA

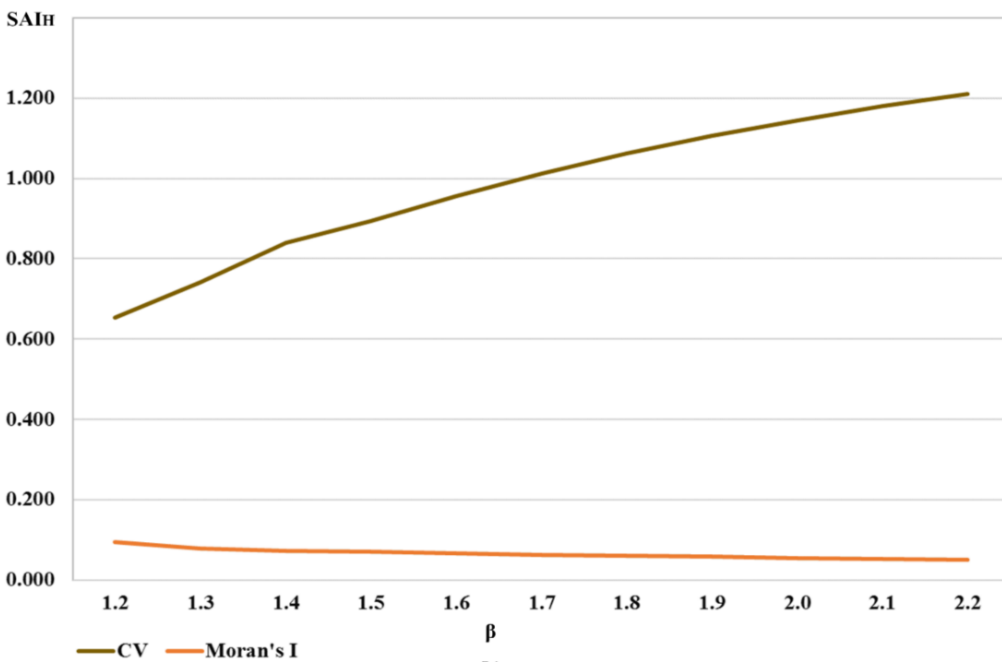
Table 4.7 summarizes the statistics of the spatial accessibility index to healthy food outlets for the 11 impedance coefficients. It can be observed that the maximum, mean, and standard deviation (SD) values increase as the impedance coefficient increases (Figure 4.7(a)). The coefficient of variation (CV) is the ratio of the standard deviation to the mean, and it shows the extent of variation from one data series to another. It indicates that, as the impedance coefficient increases (Figure 4.7(b)), the level of dispersion for the spatial accessibility increases as well. Moran's *I is* reported to show the spatial dependence of the measures with different impedance coefficients. All the Moran's *I* value are positive and significant (p = 0.000), which indicates that the accessibility measures in block groups have similar values close to each other. The Moran's I value decreases as the impedance coefficient increases (Figure 4.7(b)), meaning that the spatial interaction between neighboring units become lower with higher accessibility indices.

Table 4.7 Descriptive statistics of the SAI_H.

β	Min	1st Quartile	Median	3rd Quartile	Max	Mean	SD	CV	Moran's I
1.2	0.00023	0.00111	0.00139	0.00188	0.01047	0.00168	0.00110	0.65451	0.09493
1.3	0.00019	0.00104	0.00134	0.00189	0.01213	0.00170	0.00126	0.74160	0.07861
1.4	0.00015	0.00096	0.00129	0.00191	0.01400	0.00171	0.00143	0.83982	0.07153
1.5	0.00012	0.00089	0.00124	0.00195	0.01506	0.00171	0.00153	0.89320	0.06990
1.6	0.00010	0.00083	0.00121	0.00198	0.01622	0.00172	0.00165	0.95570	0.06644
1.7	0.00008	0.00077	0.00116	0.00198	0.01718	0.00173	0.00175	1.01175	0.06312
1.8	0.00006	0.00071	0.00111	0.00200	0.01796	0.00174	0.00184	1.06173	0.06005
1.9	0.00005	0.00066	0.00106	0.00204	0.01859	0.00174	0.00192	1.10559	0.05734
2.0	0.00004	0.00061	0.00103	0.00207	0.01910	0.00175	0.00200	1.14446	0.05492
2.1	0.00003	0.00057	0.00099	0.00212	0.01952	0.00175	0.00206	1.17961	0.05268
2.2	0.00002	0.00053	0.00096	0.00216	0.01989	0.00175	0.00212	1.21151	0.05061



(a)



(b)

Figure 4.7 Line graph of the SAIH at the block group level with a range of impedance coefficients.

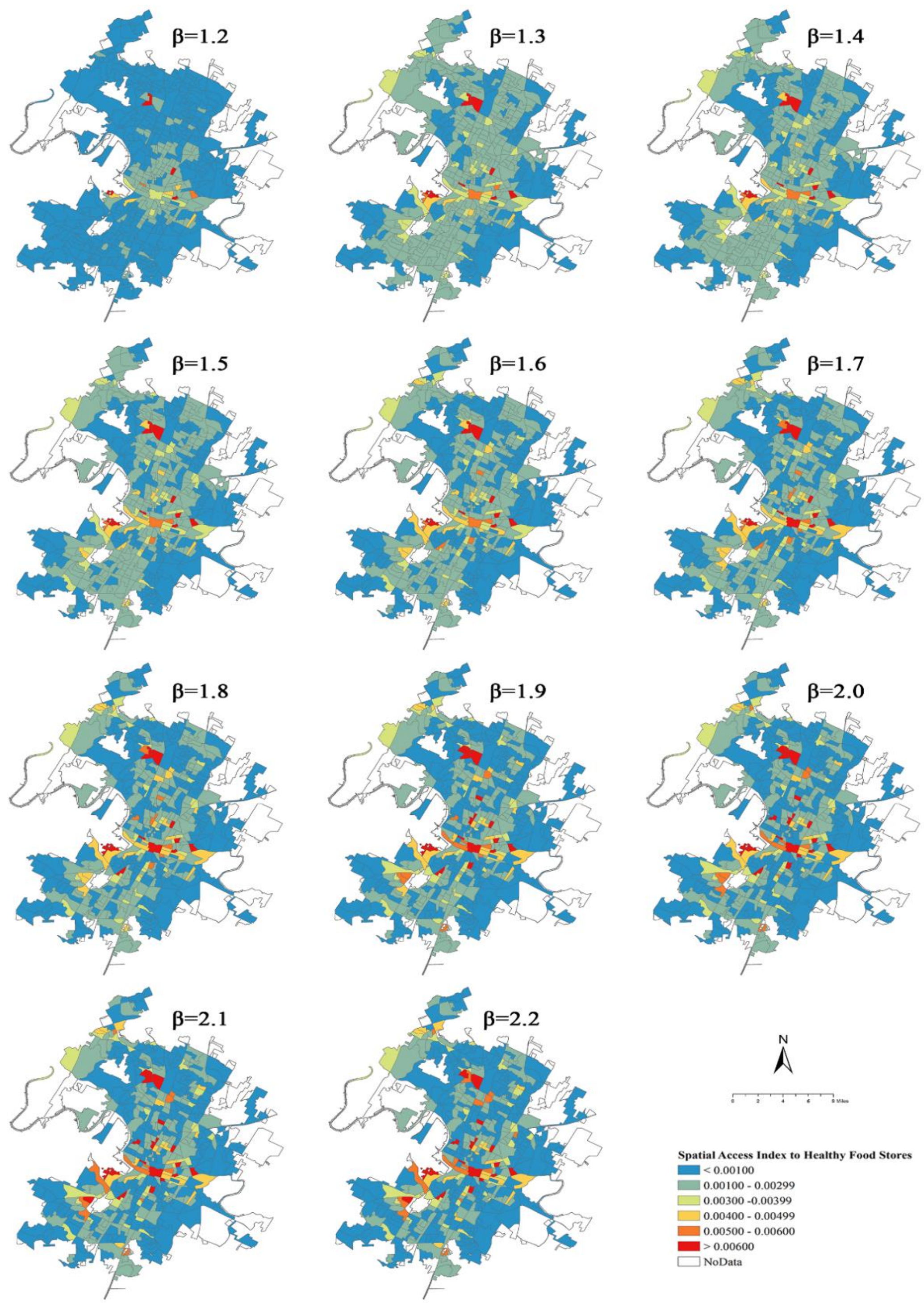


Figure 4.8 Spatial distribution of the SAI_H at the census block group level for a range of impedance coefficients

Figure 4.8 shows the spatial distribution of spatial accessibly to healthy food outlets in Austin with a set of distance impedance coefficients from 1.2 to 2.2. There is a general trend for all the impedance coefficients — accessibility index to healthy food outlets is high in the urban core and low in the peripheral areas of Austin. In other words, spatial access to healthy foods decreases when the distance is far away from the urban center. It can be seen that when the impedance coefficient is low (β = 1.2-1.4), and spatial autocorrelation is high, block groups with high spatial access to healthy foods are in the urban core and its surroundings, whereas the low spatial accessibility is found in the periphery of Austin. When impedance coefficient is high (β = 1.8-2.2) and spatial autocorrelation is low, the dark blue colors expend towards the surroundings on the maps, indicating more block groups in the periphery fall into the low accessibility interval. Meanwhile, more red colors appear on the maps and imply that more block groups in the inner urban have higher accessibility values than those with low impedance coefficients. This intriguing pattern shows that larger impedance coefficients make high accessibility higher and low accessibility lower, and thus it increases the variability of the measurements. This can be justified by the CV values in Table 4-8 that shows an increasing trend with larger impedance coefficients.

Spatial access to unhealthy food outlets using multi-mode Huff-based 2SFCA

Table 4.8 summarizes the statistics of the spatial accessibility index to unhealthy food outlets for the 11 impedance coefficients. The maximum, mean, and standard deviation (SD) values increase as the impedance coefficient increases (Figure 4.9(a)). The CV values indicate that, as the impedance coefficient increases, the level of dispersion for the spatial accessibility increases (Figure 4.9(b)). Moran's I values are positive and significant (p = 0.000), which indicate that the accessibility measures in block groups tend to have clusters or similar values close to each other. The Moran's *I* value decreases as the impedance coefficient increases (Figure 4.7(b)), indicating that the spatial autocorrelation of the accessibility to unhealthy foods become low with high distance impedance measures.

Table 4.8 Descriptive statistics of the SAI_U.

β	**Min**	**1st Quartile**	**Median**	**3rd Quartile**	**Max**	**Mean**	**SD**	**CV**	**Moran's I**
1.2	0.00157	0.00439	0.00571	0.00769	0.03897	0.00667	0.00372	0.55729	0.11667
1.3	0.00131	0.00411	0.00558	0.00779	0.03938	0.00667	0.00410	0.61398	0.10644
1.4	0.00109	0.00390	0.00553	0.00794	0.04006	0.00671	0.00446	0.66479	0.09795
1.5	0.00091	0.00367	0.00538	0.00800	0.04051	0.00674	0.00481	0.71360	0.09170
1.6	0.00075	0.00345	0.00534	0.00809	0.04093	0.00676	0.00514	0.76022	0.08692
1.7	0.00063	0.00327	0.00527	0.00801	0.04133	0.00677	0.00544	0.80234	0.08311
1.8	0.00052	0.00310	0.00528	0.00814	0.04167	0.00679	0.00574	0.84475	0.07899
1.9	0.00043	0.00296	0.00520	0.00815	0.04205	0.00680	0.00596	0.87567	0.07742
2.0	0.00036	0.00284	0.00510	0.00831	0.04309	0.00682	0.00619	0.90792	0.07515
2.1	0.00031	0.00274	0.00502	0.00841	0.04391	0.00683	0.00640	0.93718	0.07311
2.2	0.00026	0.00259	0.00491	0.00844	0.04457	0.00684	0.00660	0.96456	0.07121

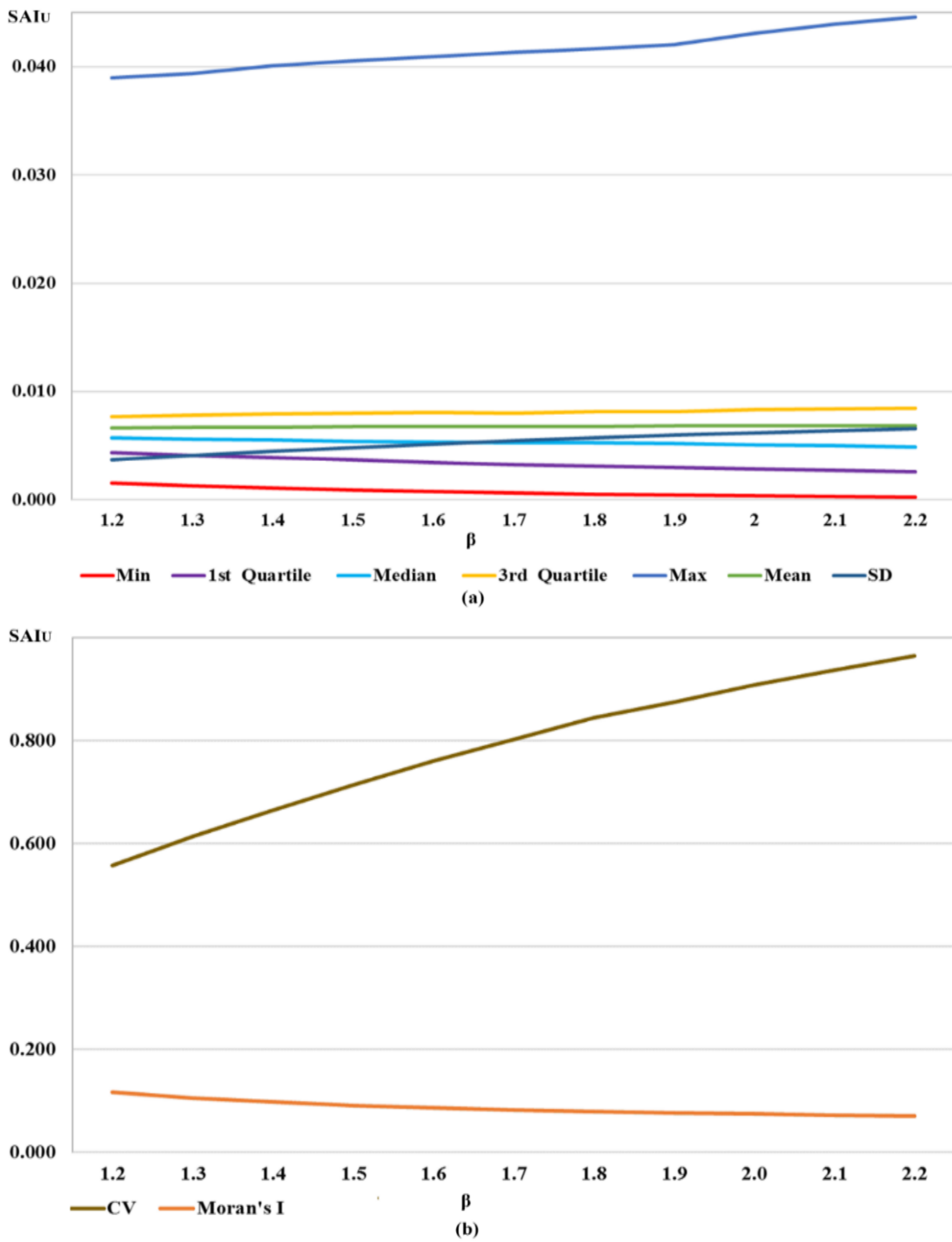


Figure 4.9 Line graph of the SAI_U at the block group level with a range of impedance coefficients

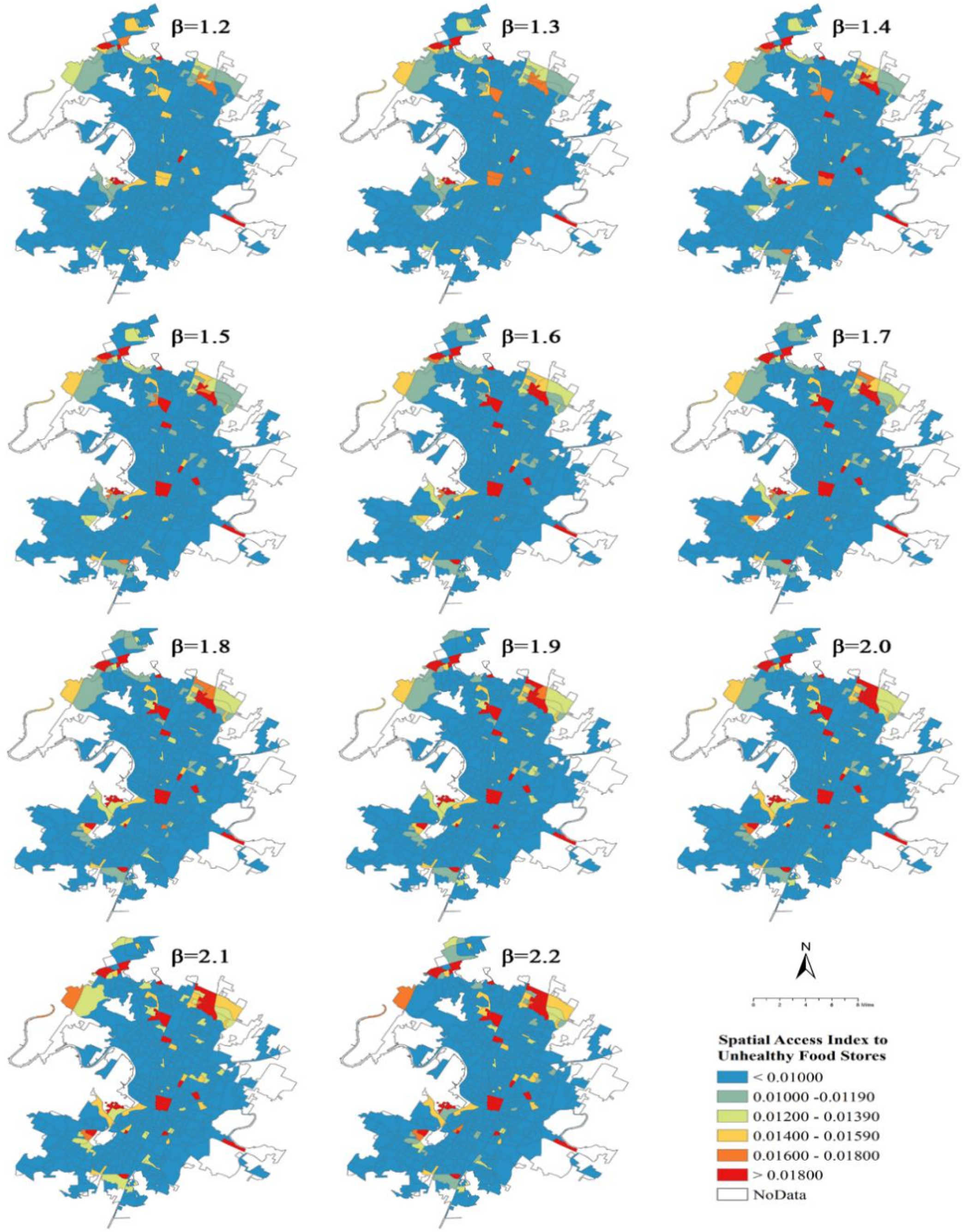


Figure 4.10 Spatial distribution of the SAI_U at the census block group level for a range of impedance coefficients

I also mapped out the spatial distribution of spatial access to unhealthy food outlets (SAI_U) in Austin with impedance coefficients ranging from 1.2 to 2.2 (Figure 4.10). It is found that accessibility index to unhealthy food outlets is high in few block groups of the urban core and the northwest and northeast of Austin; low accessibility is in the peripheral areas of Austin, except the northeastern, northwestern, and southernmost corners. When the impedance coefficient is low (β = 1.2-1.5) and spatial autocorrelation is high, as the impedance coefficient increases, the values in block groups with high accessibility become much higher. The accessibility values in their surrounding block groups become higher as well. Different from the distinct pattern in the spatial accessibility to healthy food outlets, it appears that there is a trivial change in terms of the spatial accessibility to unhealthy food outlets when β increases from 1.8 to 2.2.

Results of the comparison analysis

Comparison of spatial accessibility to healthy food outlets. The paired t-test was utilized to assess the differences between the multi-mode and single-model (i.e., automobile) methods. The result is shown in Table 4.9. It can be seen that in most of the cases the two methods with different impedance coefficients do not exhibit significant differences since their p values are larger than 0.05. There is only one exception when β equals to 1.4. Its t value is 1.981 with a p-value 0.048, which rejects the null hypothesis that there is no significant difference between the two methods. The mean difference between the two methods is largest at β = 1.4, while the smallest mean difference 0.0007 can be observed with β values ranging from 1.9 to 2.2. It seems that the mean difference between the two methods does not change much when the impedance coefficient β becomes larger.

Table 4.9 Paired t test between the multi-mode and single-mode methods regarding the SAI_H.

	Paired Difference [a]					
β	**Mean**	**Stand Deviation**	**Standard Error**	**95% Confidence Interval of the Difference**	**t-value**	**P value**
1.2	0.000013	0.000189	0.000009	(-0.000004, 0.000030)	1.526	0.128
1.3	0.000012	0.000186	0.000009	(-0.000004, 0.000029)	1.452	0.147
1.4	0.000013	0.000125	0.000006	(0.000002, 0.000024)	1.981	0.048*
1.5	0.000009	0.000175	0.000008	(-0.000007, 0.000025)	1.151	0.25
1.6	0.000010	0.000166	0.000008	(-0.000005, 0.000025)	1.303	0.193
1.7	0.000009	0.000159	0.000007	(-0.000005, 0.000023)	1.260	0.208
1.8	0.000008	0.000152	0.000007	(-0.000006, 0.000022)	1.124	0.262
1.9	0.000007	0.000146	0.000007	(-0.000006, 0.000021)	1.111	0.267
2.0	0.000007	0.00014	0.000006	(-0.000005, 0.000020)	1.155	0.249
2.1	0.000007	0.000136	0.000006	(-0.000005, 0.000019)	1.132	0.258
2.2	0.000007	0.000131	0.000006	(-0.000005, 0.000018)	1.081	0.280

Note: *significant at 0.05 level; a: statistics of paired difference are multi-mode Huff-based 2SFCA minus the single-mode Huff-based 2SFCA.

Since the two methods exhibit a significant difference with β value 1.4, I used the accessibility measures at this coefficient for the following comparison. For a better illustration purpose, I multiplied both multi-mode and single-mode accessibility values by 10,000 and then performed a logarithm transformation on them. The mean percentage for people who are driving in Austin is 90.367. Therefore, I selected the block groups below 90% of people who are driving (i.e.,147 block groups) and sketched the two methods on a scatterplot (Figure 4.11(a)). The differences are mixed. For the log-transformed accessibility index less than 3.5, multi-mode method tends to result in higher estimation than single-mode one. The difference is remarkable when the accessibility rate is low. As the log-transformed accessibility index increases, the differences become minimal. For the log-transformed accessibility value larger than 3.5 (only few block groups), multi-mode method tends to have a higher estimation than single mode method. Figure 4-11(b) shows the comparison when block groups' driving percentage above 90% (e.g., 329 block groups). The multi-mode method has a higher estimation than a single method. For the log-transformed values larger than 3.6, block groups with a multi-mode method all have a higher estimate, leading to that no estimation fall below the reference line 1:1.

Figure 4.12 shows the spatial accessibility to healthy food outlets for the two methods. To better compare the two methods, I multiplied the value by 1,000 for each block group and categorized their values into six intervals (i.e., < 1.5, 1.5-2.5, 2.5-3.5, 3.5-4.5, and > 4.5). As illustrated in Figure 4.12 (a) and 4.12 (b), multi-mode and single-mode methods exhibit similar spatial pattern of accessibility values. Urban centers and its immediate surroundings have high accessibility index (> 3.5), whereas peripheric areas have low accessibility index (< 1.5). However, they have some dissimilarities. As shown in the two inset maps, in the urban core the multi-mode method generated more medium accessibility values (2.5-4.5), while the single-mode method produced more high accessibility values (> 4.5).

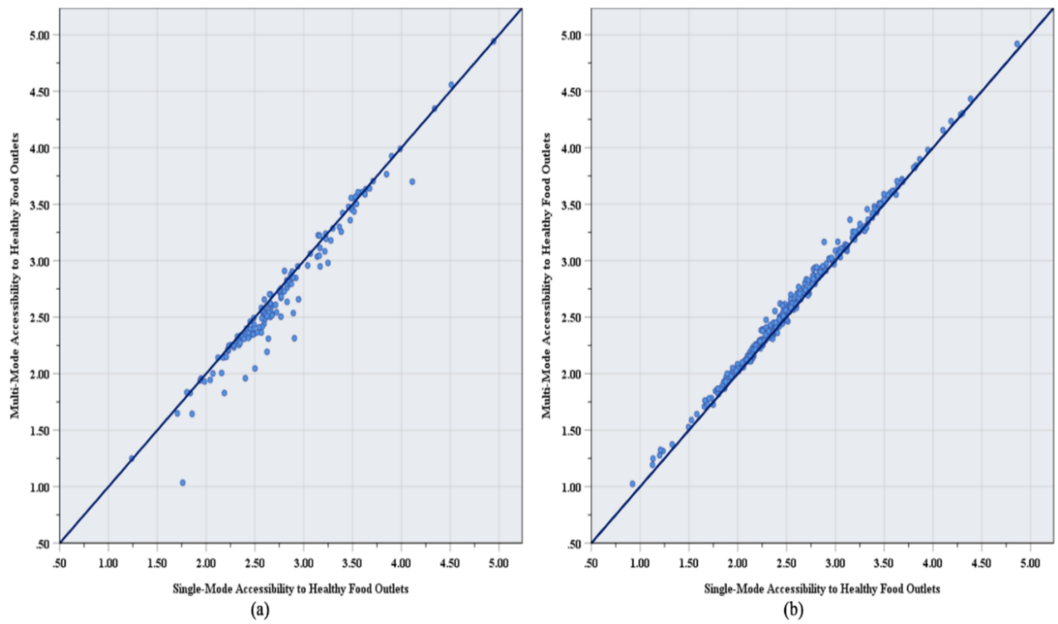


Figure 4.11 Comparison of the multi-mode and single-mode Huff-based 2SFCA on the LnSAIH in (a) urban core block groups (b) peripheric block groups.

Figure 4.12 (c) shows the magnitude and direction of the percent difference between the two methods with β = 1.4. In the urban center (i.e., downtown of Austin and the University of Texas at Austin), the percentage difference is negative (< -10 %), indicating that the multi-mode method generates more than 10% lower accessibility index than the single-mode one. Whereas in most of the peripheric areas, the percentage difference is positive (0 % - 10 %). It suggests that the multi-mode method produced 0 to 10 percent higher accessibility index than the single-mode one. I also observe that the block groups on the immediate west side of the downtown area in bright red color (i.e., > 15 %, positive percent difference). These areas are geographically adjacent to the urban center; I expect that these locations could exhibit negative difference since I assumed a low percentage of residents who drive in urban centers (i.e., less than 80%). However, I examined the percentage of people who drive in these block groups and found that their driving percentages are somehow high (e.g., more than 96 %). Therefore, it makes sense that a positive difference is observed in these areas.

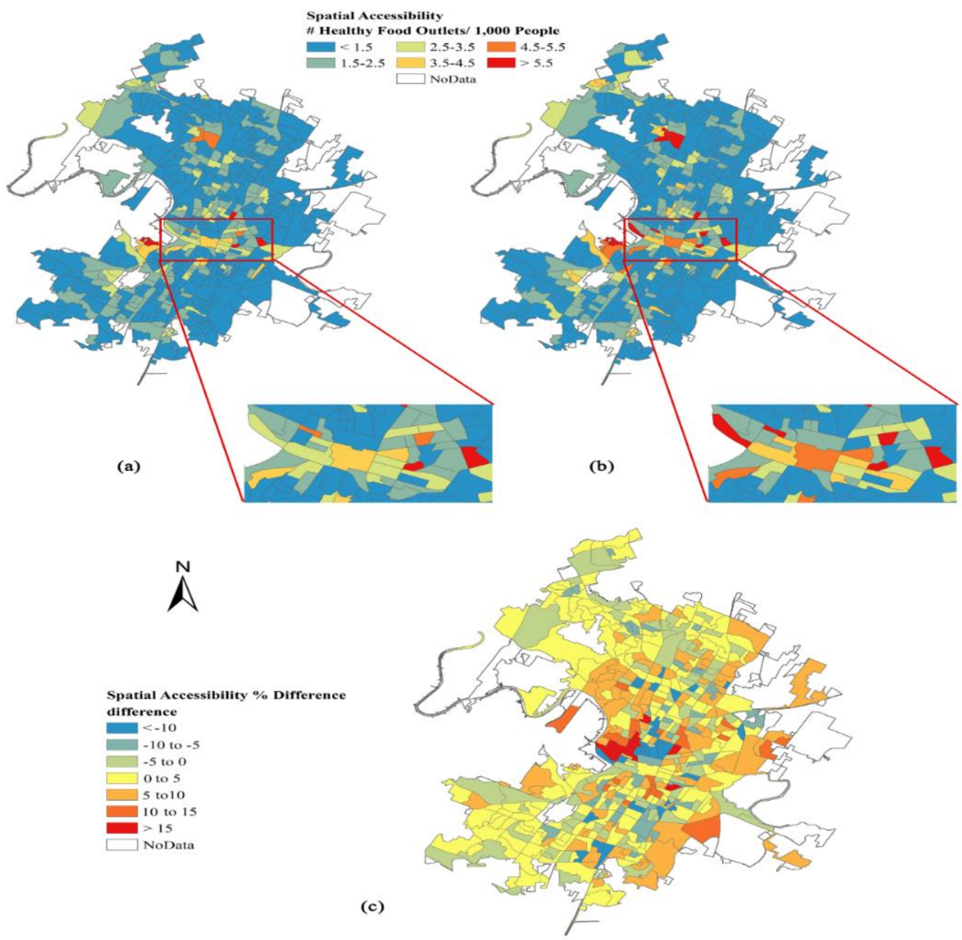


Figure 4.12 Spatial distribution of the SAI_H with (a) multi-mode and (b) single-mode; (c) the percent difference between the multi-mode and single-mode Huff-based 2SFCA.

Comparison of spatial accessibility index to unhealthy food outlets. Similar to the analysis of spatial access to healthy food outlets, I also employed the paired t-test to examine whether there is a significant difference between multi-mode and single-mode methods. I found that only when β equals to 1.5, the two methods had substantial mean differences (t value = 2.091, p value = 0.037) (Table 4.10). The mean difference between the two methods is positive with each impedance coefficient, indicating that the mean values for multi-mode method are larger than the single-model method in the study area. In addition, the most substantial difference occurred at β = 1.5 (0.000048), while the smallest mean difference can be observed with β = 1.3 (0.000042).

Table 4.10 Paired t test between the multi-mode and single-mode methods regarding the SAI_U.

Paired Difference [b]						
β	**Mean**	**Stand Deviation**	**Standard Error**	**95% Confidence Interval of the Difference**	**t- value**	**p -value**
1.2	0.000043	0.000524	0.000024	(-0.000004, 0.000090)	1.780	0.076
1.3	0.000022	0.000546	0.000025	(-0.000027, 0.000071)	0.869	0.385
1.4	0.000042	0.000508	0.000023	(-0.000004, 0.000088)	1.803	0.072
1.5	0.000048	0.000499	0.000023	(0.000003, 0.000093)	2.091	0.037*
1.6	0.000041	0.000498	0.000023	(-0.000003, 0.000086)	1.813	0.070
1.7	0.000041	0.000500	0.000023	(-0.000004, 0.000086)	1.787	0.075
1.8	0.000041	0.000516	0.000024	(-0.000005, 0.000087)	1.735	0.083
1.9	0.000041	0.000506	0.000023	(-0.000005, 0.000086)	1.752	0.080
2.0	0.000039	0.000511	0.000023	(-0.000007, 0.000085	1.680	0.094
2.1	0.000041	0.000513	0.000023	(-0.000005, 0.000087)	1.738	0.083
2.2	0.000041	0.000515	0.000024	(-0.000005, 0.000087)	1.737	0.083

Note: *significant at 0.05 level; b: statistics of paired difference are multi-mode Huff-based 2SFCA minus the single-mode Huff-based 2SFCA.

I used the β value of 1.5 to compare the two methods. The result is shown in Figure 4.13. It can be seen that the single-mode method tends to generate a higher estimation than the multi-mode method in urban core block groups (Figure 4.13(a)). The difference was remarkable when the log-transformed accessibility index is medium (3.7- 4.5). The difference declined when the value is more than 0.010. Figure 4.13(b) shows the comparison in block groups in peripheric block groups. The multi-mode method has a higher estimation than a single method. The difference is remarkable when the transformed accessibility index is medium (3.0 - 4.0). As the accessibility index increases, the difference became minimal. For the log-transformed values larger than 5, block groups with a multi-mode method all have a higher estimation, leading to that no estimation fall below the reference line 1:1.

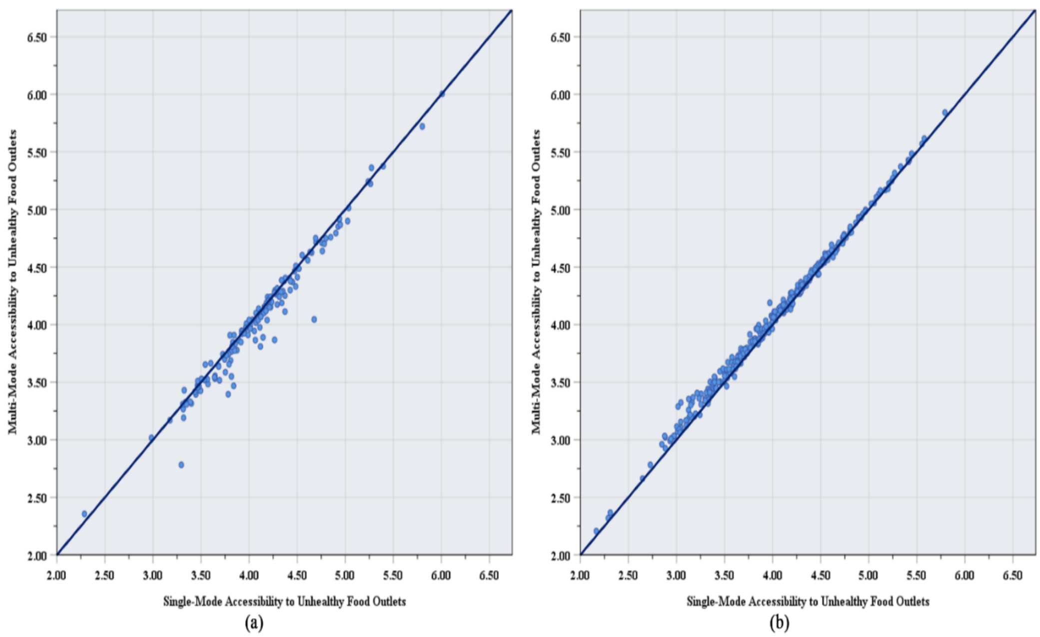


Figure 4.13 Comparison of the multi-mode and single-mode Huff-based 2SFCA on the $LnSAI_U$ in (a) urban core block groups (b) peripheric block groups.

Figure 4.14 shows the spatial patterns of accessibility to unhealthy food outlets for the two methods. I multiplied the value in each block group by 1,000 and categorized their values into six intervals (i.e., < 3, 6-9, 9-12, 12-15, and > 15). It can be seen that multi-mode and single-mode methods exhibit a similar spatial pattern of accessibility values (Figure 4.14(a) and 4.14(b)). The block groups in the downtown Austin, northwestern and northeastern Austin have high accessibility index (> 12), whereas rural areas have low accessibility index (< 3). Nevertheless, there is a dissimilarity between the two maps. The two inset maps show that in the urban core the multi-mode method generates more high accessibility values (> 20), while the single-mode method produces more medium accessibility values (10 - 20).

The differential map in Figure 4.14(c) shows the magnitude and direction of the percent difference between the two methods with β 1.5. It can be seen that the negative percent difference (< -10, or blue colors on the map), in general, could be observed at University of Texas at Austin (close to downtown Austin), as well as in the mid-north and mid-south of Austin along the IH-35, indicating that the multi-mode method generates more than 10% lower accessibility index than the single-mode counterpart. Whereas in most of the peripheric areas, the percentage difference is positive (0% - 10%). It suggests that the multi-mode method produces less than 10 % higher accessibility index than the single-mode method in these areas. Some block groups that approximate to urban center have a high positive percent difference (> 15) between the multi-mode and single-mode methods because in these areas the majority of people (more than 96%) own personal vehicles and are able to drive to food stores for foods.

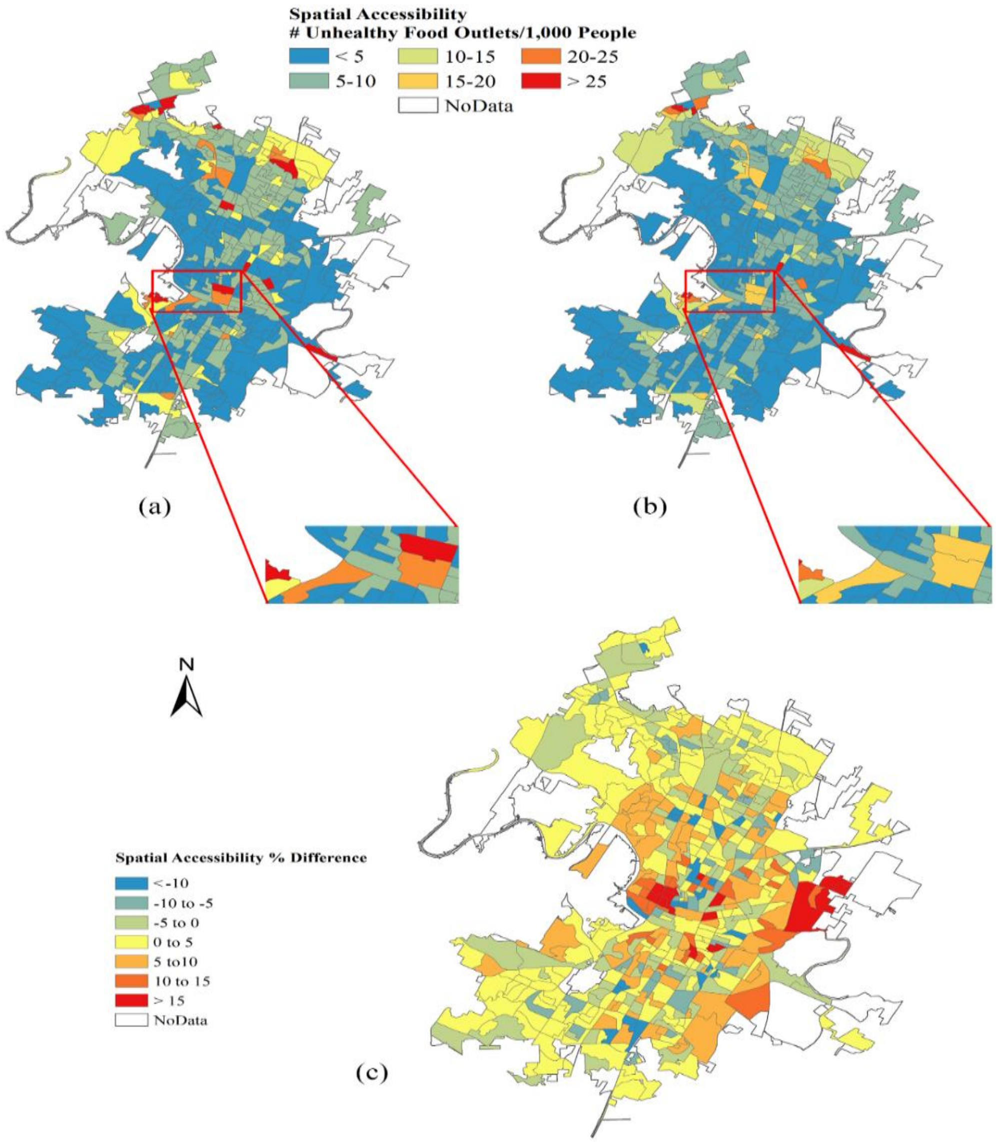


Figure 4.14. Spatial distribution of SAI_U with (a) multi-mode and (b) single-mode; (c) the percent difference between multi-mode and single-mode Huff-based 2SFCA.

Discussion and Conclusion

This part of the dissertation proposes a new method, multi-mode Huff-based 2SFCA, to overcome the over-estimating population demand issue in previous models. I applied this novel method in Austin, Texas to measure both healthy and unhealthy food accessibility. It can effectively minimize the overestimation problem in urban areas. Therefore, it exhibits a more accurate picture of spatial access to food outlets than its alternatives.

By using the proposed method, I found that the spatial accessibility to both healthy and unhealthy food with different travel modes reveals a clear pattern in the urban core and peripheral areas. The urban core areas have the best access, whereas many block groups in peripheral regions in Austin have inadequate access to food stores. This result is consistent with previous findings (Apparicio et al., 2007). The urban core- peripheral disparity is also seen in the single-mode Huff-based 2SFCA method. This phenomenon can be explained as grocery stores, convenience stores, and other food outlets are mainly concentrated in urbanized areas (Kuai 2017). Food retail business usually chooses to operate in urban core areas because a dense population density in these areas can ensure a high shopping volume and revenue (Kuai 2017). Therefore, the distribution of food retailers (or spatial advantage/disadvantage) could be a facilitating or impedimental factor to procure foods.

I compared the proposed method with its single-mode alternative. The primary advantage of the proposed method is that it differentiates the population with and without vehicles. It uses different transportation modes as a constraint or weight to further adjust the overestimation demand in both steps and leads to a much more reasonable result than the single-model method. The two methods are generally consistent with each other in terms of the urban core- peripheral disparities regardless of impedance coefficients. The result of paired t-test also supports this finding because the two methods exhibit insignificant difference with most of the impedance coefficients (except for β = 1.4

(healthy foods) and β = 1.5 (unhealthy foods)). I also found that the multi-mode method estimates an overall higher variability than the single-mode one, which is coincident with our expectation because a population with various transportation modes can maximize the heterogeneity of the measurement and thus produce a higher standard deviation.

Since the composition of transportation modes reveals a notable difference in urban and peripheral areas, I compared the multi-mode and single-mode methods in these areas separately. In most of the peripheral block groups, for both healthy and unhealthy food accessibility measures, the single-mode method produces a lower value than multi-mode one. It can be explained as follow. The single mode assumes that all people purchase foods by automobile, leading to that more people access to food stores to compete for food resource, which results in a larger denominator in Eq. 4-11 and thus a lower supply-to-demand ratio R_j. At the last step of the single-mode method (Eq. 4-12), it sums all $Prob_{ij}^{H} R_j W_{ij}$ in its catchment. In peripheral areas more of households (more than 96%) have vehicles, there is no much difference between the single mode (Eq. 4-12) and multiple modes (Eq. 4-15) at this step. Because R_j tends to be lower in Eq. 4-11, the single-mode method is anticipated to generate lower accessibility indices in peripheral areas. The single-mode Huff-based 2SFCA method tends to underestimate the accessibility values in peripheral regions. Therefore, the single -mode method has a high likelihood of overestimating food deserts (defined as underserved areas for healthy foods) and underestimating food swamps (defined as overexposed areas for unhealthy foods) in peripheral block groups (Table 4.11). This finding is essential to food stakeholders and health policymakers for interventions. Studies that use the single method for estimation would identify larger food deserts and smaller food swamps than they are (see Table

4.11). It means that the population in peripheral areas ask for more intervening resources for food deserts and fewer intervening resources for food swamps than they need. Consequently, food departments and health planners should reduce the amount of funding and interventions for food deserts but increase funds for the elimination of food swamps in peripheral areas when they review proposals and reports with single-mode method (by car) (Table 4.11).

Table 4.11 Consequences of using single-mode method to estimate food deserts and food swamps.

Classification	Food Deserts	Food Swamps
Peripheral areas	Overestimate Intervention: reduce resources	Underestimate Intervention: increase resources
Urban core	Underestimate Intervention: increase resources	Overestimate Intervention: reduce resources

In most of the inner-urban block groups, for both healthy and unhealthy food accessibility measures, the multi-mode method produces a lower value than the single - mode one, which could be explained as follows. The multi-mode method supposes that people procure foods by various transportation means, and thus fewer people compete for foods, which results in a higher R_j (Eq. 4-14). In the urban core, only a certain percentage of people (i.e., % 80) own cars and personal vehicles due to the developed public transportation systems and severe traffic and parking issues. At the last step of the multi-mode method (Eq. 4-15), the lower driving percentage in urban areas discounts on $Prob_{ij}^{H} R_j W_{ij}$ greatly, which decreases the overall accessibility. As a result, the single-mode method produces higher accessibility than the multi-mode one in the metropolitan core area. Therefore, it is very likely for the single -mode method to underestimate food deserts and overestimate food swamps in urban core block groups. This finding indicates that studies with a single method would identify smaller food desert areas and larger food swamp areas than they are (Table 4-12). It means that the population in urban core areas need more resources to deal with food deserts and fewer interventions to tackle food swamps than they ask for. In this sense, stakeholders should increase funding to facilitate access to healthy foods but reduce resources to limit access to unhealthy foods with the single-mode method (Table 4-12).

Despite the advantages of the proposed method, the results should be interpreted with caution. There are several points needed to be paid attention to, and they are summarized below. First, I used 10-min walking, 15-min driving, and 30-min public transit to generate the catchments for each transportation mode in this study. It is a reasonable assumption because people would like to select nearby food stores. However, the breaking travel time points were based on empirical data and seemed arbitrary. I do not have detailed information on how much time consumers are willing to spend on the commuting to food stores. Therefore, in the future study a customer survey may be needed to determine the most appropriate traveling breaking points for catchment size in each transportation mode (Wan et al. 2013).

Moreover, keep in mind that the catchment size could vary for different applications based on the neighborhood chrematistics and context (Wan, et al. 2012). For example, Mao and Nekorchuk (2013) set 30-min catchments for driving and bus, while (Kuai and Zhao 2017) used 15-min walking, 15-min (urban) and 25-min (rural) driving, and 15-min public transit to delineate the travel catchments. Ikram, Hu, and Wang (2015) calculated accessibility to pharmacies with multiple travel catchment sizes: 10 min, 15 min, 20 min, and 25 min. They found that increasing catchment-area size reduces the number of areas with zero (or low) accessibility. However, too large catchment size could lead to that the variability of accessibility is smoothed out, as well as the accessibility scores in most of the study area are close to the area average. Therefore, an optimal catchment size is vital to the final accessibility scores. Some research (i.e., McGrail and Humphreys (2009)) proposed methods to determine the optimal size, which might be incorporated into the proposed method to develop a more sophisticated model in a future study.

Moreover, I used constant catchment size in both steps. However, this is not necessary, and the size could vary in different steps as suggested by Wan, Zou, and Sternberg (2012). For instance, Luo and Whippo (2012) used variable sizes for demand catchment in E2SFCA method based on the number of opportunities at the facilities.

Second, population-weighted centroids were used to represent the population distribution of block groups. For a sparsely-populated block group, it may not produce accurate accessibility results. Future study may consider using finer-scale population data (i.e., parcel-level or grid level) to minimize the inaccuracy. Also, the use of population centroids assumes that home to store travel is the way most people access food stores. But people do not always travel from home to store. They go to work, schools, and churches and often purchase food on the way. In other words, we measured the potential access but not the realized access. Without detailed data of individuals' travel behavior, it is not likely to calculate the realized access to food stores.

Third, I used the percentage of people taking the three transportation modes to work as a proxy of food shopping. There might be a discrepancy between the two types of percentages; however, this was the best information we can obtain from the census. I simplified transportation into three modes. People in real life may use other modes (e.g., motorcycles) to purchase foods. Mao and Nekorchuk (2013) suggest that it is a good practice to use transportation simulation models to generate a dataset, simulating specific travel behaviors for particular population groups. Also, we did not consider the temporal dimension of accessibility. As temporal dynamics of food accessibility have drawn much more attention to scholars, the future study may integrate temporal dynamics (i.e., store opening hours) into the measurement.

Fourth, this study did not include bodegas, snack/beverage shops, discount stores, and farmer markets into food source; particularly for farmers' markets, and they are the important healthy food source for residents in the rural area of Austin. Future research should incorporate them into food access measures.

Lastly, I only considered spatial disadvantage (i.e., where are the food outlets) to quantify neighborhoods' accessibility to foods in this paper since this is primarily a spatial accessibility study. I did not incorporate non-spatial disadvantages such as socio-demographic factors into the measurement. These variables are equally essential to identify underserved areas and populations for procuring foods. Therefore, I may use some methods to combine spatial and non-spatial factors for a better delineation of regions in a future study (Kuai and Zhao 2017), which would be explored in Chapter Six since it is an important logical extension of current work (Wan, Zou, and Sternberg 2012)

5 EXPLORING ECONOMIC AND SOCIOCULTURAL DISPARITIES IN ACCESS TO FOOD OUTLETS

Introduction

The elevating prevalence of overweight and obesity cannot be fully explained by individuals' genetics and personal behaviors (Leia Michelle Minaker 2013; Luan 2016; Stein 2011; Coulter 2009). More recent studies have linked obesity with retail food environment (Witten 2016). Food access is an indispensable part of the retail food environment (Glanz, et al. 2005). It can shape individuals' food shopping and consumption behaviors, therein influence their health status (Morland, Roux, and Wing 2006).

Food access not only varies in spatial dimension but also differs in the nonspatial aspect (Dai & Wang 2011; Kuai 2017). For this reason, exploring nonspatial factors and their interaction with spatial access to food stores is equally important. Nonspatial factors usually are demographic and socioeconomic variables such as income, race or ethnicity, education, and employment (Dai and Wang 2011; Kuai and Zhao 2017). In the past two decades, researchers have extensively explored the relationship between food access and neighborhood deprivation (Morland, et al. 2002; Powell, Chaloupka, and Bao 2007; Pearce, et al. 2007; Zenk, et al. 2005b). Their research mainly focused on whether people in deprived areas have limited access to healthy foods or are overexposed to unhealthy foods (Galvez, et al. 2008; Kwate, et al. 2009; Lisabeth, et al. 2010), as well as these groups are more vulnerable to adverse health outcomes (Brown and Miller 2008; Powell, et al. 2007; Rundle, et al. 2009). However, findings were mixed in different studies. Some found that minorities neighborhoods and economically deprived neighborhoods had poorer access to foods (Fleischhacker, et al. 2011; Hilmers, Hilmers, and Dave 2012; Larsen and Gilliland 2008), while others reported that socially deprived or minority neighborhoods enjoyed better food access (Apparicio, Cloutier, and Shearmur 2007; Donkin, et al. 1999). The contradictory findings make it challenging to conclude that deprived neighborhoods have less access to healthy foods or more access to unhealthy foods. Part of the reasons could attribute to the weakness in characterizing neighborhood deprivation and drawbacks in traditional statistics.

One issue is that many studies associated several individual sociodemographic variables with food access (Dai and Wang 2011; Luan 2016). These individual variables are often highly correlated with each other. For instance, low-income neighborhoods tend to have a high percentage of people in poverty and a low proportion of people in employment. These variables represent the same deprivation dimension. Once they are entered in regression analysis, some variables could act as a proxy of the related ones, leading to the multicollinearity problem. Thus, a more sophisticated method is required to represent the real dimensions of socio-demographic factors. Such method is to construct a composite index (Matheson, et al. 2012; Larsen and Gilliland 2008). Factor Analysis (FA) and Principal Component Analysis (PCA) and are two commonly used methods for the construction of a composite index (Yu, et al. 2014; Matheson, et al. 2012; Zadnik and Reich 2006). For example, a well-cited and classic work Apparicio, Cloutier, and Shearmur (2007) developed a social deprivation index based on five variables: low income population (%), lone-parent families (%), unemployment rate (%), adults with low level of schooling (%), and recent immigrants (%). The spatial access to supermarkets was based on proximity, density, and competition. Then, the social deprivation index was combined with spatial access measures to identify food deserts. Gustafson, et al. (2012) estimated the Neighborhood Deprivation Index (NDI) in 14 counties in Kentucky with the following criteria: income below the poverty line, female-headed households, public assistance recipients, unemployment rate, population in management (%), education attainment, and families with at least two persons per room. Each factor was weighted to estimate the final deprivation scores via the PCA technique.

The other issue is to employ traditional or non-spatial statistic to analyze the association between neighborhood deprivation and food access. These traditional statistics include but not limited to Ordinary Least Square (OLS) for continuous scale data (Kuai and Zhao 2017), Poisson binomial regression for count data (Black, et al. 2011), logistic regression for binary data (Black, et al. 2011; Smoyer-Tomic, et al. 2008), etc. One of the important assumptions for these statistical methods is that each observation is randomly distributed over geographic space. However, food accessibility in essential is a ‘spatial' problem (Dai and Wang 2011; Kuai and Zhao 2017); residuals are likely to be spatially autocorrelated to invalidate the results of the analysis potentially. The use of spatial statistics could solve the deficiencies of traditional statistics.

Nevertheless, only a few studies used spatial statistics. For example, a spatial scan method was utilized to model the counts of food outlets with sociodemographic variables in St. Louis, Missouri (Baker, et al. 2006). A spatial Bernoulli model was employed by Lamichhane, et al. (2013) to analyze the relationship between access neighborhood deprivation and access to supermarkets and fast-food outlets in South Carolina. Luan (2016) extended Lamichhane and colleagues' (2013) work and utilized the spatial-temporal Bayesian hierarchical approach to model the count of food outlets with deprivation variables.

Of the spatial statistical approaches, spatial autoregressive models (SAR) have arisen researchers' attention. They were designated to deal with the problem of spatial dependence in the non-spatial analysis. There are two models constituted the SAR. One is a spatial lag model (SLM); it is a model that the dependent variable is affected by the values of the dependent variables in nearby places. The other one is a spatial error model (SEM), which is that some spatially clustered features that influence the value of the dependent variable and its neighbors but is omitted from the specification. There are only a few food studies that have utilized SLR and SEM. Dai and Wang (2011) employed SLM to account for the spatial autocorrelation effects of the food accessibility and socio-economic variables in South-west Mississippi. McKenzie (2014) associated neighborhood supermarket access with sociodemographic factors with the SEM in Portland, Oregon. Wang, et al. (2016) applied both SLM and SEM to analyze the relationship between spatial proximity (i.e., nearest distance and minimum travel time) to fresh food retailers and SES in two cities in Canada. However, the three studies are subject to two problems: 1) two out of the three used individual sociodemographic variables in the model, which potentially result in multicollinearity problem; 2) all of them only explored the relationships between sociodemographic variables with healthy food accessibility but failed to examine deprivation with unhealthy food accessibility (i.e., fast food restaurants).

Spatial dependence effects consist of spatial autocorrelation and heterogeneity (Zhang, Ma, and Guo 2009). SLR and SEM were sufficient to account for spatial autocorrelation but were insufficient to tackle the spatial heterogeneity problem (Zhang, Ma, and Guo 2009). The studies above only accounted for spatial autocorrelation in exploring spatial accessibility and socioeconomic factors but failed to consider spatial heterogeneity. Therefore, it is demanding to examine further how spatial heterogeneity influences the relationships between spatial accessibility and socio-economic factors.

Spatial heterogeneity, as opposed to spatial stationarity, measures structural instability of phenomenon by varying model parameters (Anselin and Griffith 1988). It means that we should effectively treat the spatial aspects of food accessibility and socioeconomic variables, and the spatial variation of their relationship. Geographically Weighted Regression (GWR) allows us to explore spatial non-stationarity process in food environment and social or economic factors. Several studies are using GWR to examine the relationship between food access, sociodemographic factors, obesity and other diseases (Alnasrallah 2015; Xu 2014; Ford and Highfield 2016). I acknowledge that GWR is helpful to examine the spatial variation of the relationship and to identify significant local patterns in these studies. However, GWR is too strict and rigid in parameter specification, meaning that it assumes that all coefficients of independent variables vary over space. Not all variables have varying relation with dependent variables. Semi-parametric GWR allows the flexibility to incorporate both fixed and geographically varying explanatory variables (Nakaya, et al. 2017). The advantage of using a semi-parametric approach is that including explanatory variables as fixed when they do not vary significantly over space, which can provide a more conceptually satisfactory model and improve the overall model fit. Such models have not previously been used in food environment studies.

Studies that explore the relationship between food access and socioeconomic deprivation are also can be used to identify food insecurity problems such as food deserts and food swamps. It is motivated by deprivation amplification hypothesis. Food deserts are areas that residents have barriers to access healthy foods in deprived neighborhoods; food swamps are where residents are overexposed to unhealthy foods in deprived communities. Most studies used traditional statistics to associate physical access with social-demographic variables. Take Gordon, et al. (2011) as an example, and they developed a Food Desert Index based on the competition of healthy and unhealthy foods in New York City. The food access index was ranked to a range of 3 (poor) to 9 (high). Then, the Food Desert Index and the demographic variables were combined to create a total food desert score. The areas with the highest food desert scores were identified as food deserts. It is a typical approach to identify food deserts and food swamps and is used in a well-cited work (Apparicio, Cloutier, and Shearmur 2007) and other studies. This approach is problematic in two aspects: 1) The classification of the low, medium and high is arbitrary. There are no standard thresholds in the literature; 2) The measure does not consider spatial associations between physical access and socio-demographic variables. In other words, it ignores the effect of spatial dependence in the relationship in the identification.

To better delineate the food deserts and food swamps with the consideration of spatial dependence in the study area, Exploratory Spatial Data Analysis (ESDA) is used. According to Anselin and Getis (1992), this methodology is useful to find patterns of spatial associations (i.e., clustering and dispersion), identify clusters and outliers, and define spatial instability such as non-stationarity. In general, ESDA could be classified into two categories: (1) global techniques, which focus on the entire study area and help to identify general spatial patterns such as cluster, dispersion, and randomness; (2) local techniques, which identify where the clusters are on the subset of the study area (Haining, Wise, and Ma 1998). The global statistics are mainly used to measure clusters in the study area; however, they fail to account for spatial autocorrelation occurring in neighboring units, particularly in a large dataset (Anselin 1995). Local indicators of spatial association (LISA) are often used to identify the local patterns and solve the spatial dependence issue in the dataset. It thus is capable of revealing the types of spatial correlation and the location of clusters and outliers. LISA consists of various indicators such as local Moran's I, local Geary's C, and Getis-Ord Gi*. For instance, Stein (2011) utilized a local bivariate Moran's I to analyze the relationship between two variables so that significant clusters and outliers in neighboring neighborhoods (Stein 2011). The specific method is to delineate food deserts (or food swamps) by identifying a Low-High (or High-High) relationship between healthy (or unhealthy) food accessibility in neighboring block groups. However, his study is subject to a serious problem --- the conceptual and applied definitions do not agree with each other. The bivariate Moran's *I* only concerns the relationships between food accessibility and social deprivation in neighboring units but ignoring the ones in the same unit, which deviates its conceptual definition. Nevertheless, the use of Getis-Ord Gi* (known as hot spot analysis) could solve this issue. It reveals spatial patterns in both neighboring units and its per se. The detailed explanation of this method would be present in the method part.

The objectives of this research are threefold. First, examining the relationship between sociodemographic factors and access to both healthy and unhealthy food by accounting for spatial autocorrelation. Second, exploring the relationships between the two domains (deprivation and food access) by considering spatial heterogeneity (non-stationarity). Lastly, identifying food deserts and food swamps by atoning for spatial dependence. The remaining of this chapter was organized as follows. The data source was present in 5.2; method section was followed. I demonstrated results in 5.4 and provided discussion and conclusion in 5.5.

Data

Economic variables data were obtained from the 2016 American Community Survey (ACS) 5-year estimations. There were four variables in the consideration (see Table 5-1). Harrington and Elliott (2009) quantified neighborhood economic condition using median household income, a proportion of households below the poverty level, and unemployment rates. Dai and Wang (2011) included household lacking complete kitchen facilities as an indicator of economic disadvantage. Anupama Joshi, the director of the National Farm to School Network, asserts "the lack of kitchen facilities or minimal kitchen facilities to prepare any fresh foods or cook from scratch, also contributes to obesity"[19].

Sociocultural variables data were also obtained from 2016 American Community Survey (ACS) 5-year estimations. Four variables were selected in this research (see Table 5.1). Brown, Perkins, and Brown (2004) articulated that home ownerships represent social cohesion. Home owners feel more attached to their places and neighborhoods than

[19] https://webspm.com/Articles/2011/06/01/Designed-to-Curb-Obesity.aspx?Page=1

home renters. Place attachments bond neighborhoods and their residents together; people can translate place attachments to feelings of pride and well-being, which could promote stability, familiarity, and security of neighborhoods and communities. Home owners stay longer in the neighborhood and are more willing to participate in community activities. By contrast, people in a rental-based neighborhood tend to less attach to their communities and less likely to invest in their direct surroundings because a transient and temporary residence does not bond them to the neighborhood and their neighbors. Harrington and Elliott (2009) included the percentage of people with their high school education as a sociocultural environmental factor. They argued that educational attainment can represent people's literacy on health, and it is likely to represent a neighborhood's attitudes and beliefs about obesity.

Ethnic group such as Hispanic/Latino have their own culture identities Hispanic groups are more vulnerable for obesity development (Galvez, et al. 2008). The Census defines linguistic isolation as all adults in the household speak a language other than English and no adults speak English "very well"[20]. Household linguistic isolation can reflect the broader sociocultural context in which built environment is situated, which has been used in many obesity and food studies to characterize neighborhood sociocultural characteristics. For example, Hsieh, et al. (2015) studied the associations between built environment and adiposity parameters among overweight and obese Hispanic adolescents; they used the proportion of Spanish speaking households with census tract-

[20] https://www.census.gov/hhes/socdemo/language/data/census/li-final.pdf

level linguistic isolation to measure neighborhood-level acculturation. Dai and Wang (2011) explored food stores accessibility in southwest Mississippi and used the linguistically isolated households as a component of sociocultural barriers to food stores. Texas has the highest numbers of linguistically isolated households. In Austin, 35% of people are Hispanic/Latino. More than half of linguistically isolated households (80 %) in Austin are Hispanic/Spanish speakers. Therefore, linguistic isolation is also included to measure socio-cultural environment.

Table 5.1 Economic and socio-cultural variables in the study area.

	Year of Data
Economic Variables	
median household income	
unemployment	
below the poverty level	
household lacking complete kitchen facilities	2016
Socio-cultural Variables	
rental home	
without higher education	
Hispanic population	
linguistic isolation	2016

Method

Construction of economic deprivation index and sociocultural deprivation index

The descriptive statistics of the selected socio-demographic variables were reported in the result section. Meanwhile, their spatial distributions were reported as well. Bivariate Pearson correlation was performed between the selected variables to examine the potential multi-collinearity issue. It turned out that some of the selected sociodemographic variables were highly correlated, and it is likely to cause information redundancy.

To eliminate multi-collinearity impact, factor analysis (FA) was performed in SPSS 25.0 to integrate the four economic (or sociocultural) variables into a single index. FA is a quantitative technique identifying a smaller number of uncorrelated components from a relatively larger set of observed variables without losing much information (Meyers, Gamst, and Guarino 2013; Smith, et al. 2002). This technique produces a weight for each variable according to its contribution in explaining the differences between analytic units. It constructs the component scores using a regression model; the component scores are dependent variables, and standardized observed values of the items in the estimated components are independent variables. Then, the percentage of variance associated with each principal component (based on eigenvalues of components greater than one) were obtained after running FA with Varimax rotation. Finally, component scores in each block group were multiplying by the explained variance percentage, then was summed up to generate a component index. The eight variables were reduced into two indices: Economic Deprivation Index (EDI) and Socio-cultural Deprivation Index (SDI).

Using spatial autoregressive model to examine associations between the indices

In this section, different spatial regression models were used to examine the relationship between food accessibility and social deprivation. One of the most critical assumptions of regression is the normality of the dependent variable (i.e., food accessibility). Therefore, we carefully examined the distribution of accessibility measure before performing any further analysis. It is found that food accessibility measures were not normally distributed. As a result, we have to do data transformation. The most widely used transformations are natural logarithm (ln), the logarithm (log10) and square root (sqrt). I performed three transformations. It turned out that the natural logarithm (ln) is the most appropriate one to use because it alleviates the abnormality of original data mostly. To avoid negative values of the two accessibility measures, we firstly inflated the two original measures by 10,000 and then performed the natural logarithm (ln) transformation. The two transformed ones are named as $LnSAI_H$ and $LnSAI_U$, respectively.

Bivariate parametric Pearson correlation was used to provide a preliminary insight on their crude associations. I then performed SAR model to examine a more accurate relationship between spatial food accessibility and social deprivation indices. The SAR model was implemented in GeoDa 1.1.2. The SLM accounts for the spillover effects of the dependent variable on the regress model. Its general form is:

$$\mathrm{Y} = \rho \mathrm{WY} + \mathrm{X}\beta + \mathrm{u} \qquad 5\text{-}1$$

where Y is the dependent variable, here is the logarithm transformed accessibility index ($LnSA_H$ or $LnSA_U$); W is the spatial weights matrix, WY together represents spatial lag; ρ is the spatial lag coefficient, X is a vector of the independent variables: EDI and SDI; β is a vector of the estimated coefficients, and u is the error vector.

When solving y in Equation 5-1, it is reduced to this form:

$$\mathrm{Y} = (\mathrm{X}\beta + \mathrm{u}) * (\mathrm{I} - \rho \mathrm{W})^{-1} \qquad 5\text{-}2$$

where $(I - \rho W)^{-1}$ is called spatial multiplier.

The SEM captures the effects of error terms and omitted variables, especially when these omitted variables exhibit spatial correlation. The form of its model is shown below.

$$\mathrm{Y} = \mathrm{X}\beta + \mathrm{u} \quad and \quad \mathrm{u} = \lambda W u + \varepsilon \qquad 5\text{-}3$$

where W is the spatial weights matrix, u is the error vector; Wu together represents spatial error; λ is the spatial error coefficient; other notations remain the same as the Eq. 5-1. U in Eq. 5-3 is solved to obtain its reduced form:

$$Y = X\beta + \ \varepsilon(I - \rho W)^{-1} \qquad 5\text{-}4$$

where $(I - \lambda W)^{-1}$ is also called spatial multiplier.

The use of SLM or SEM is contingent upon the diagnosis of the OLS regression model. Thus, OLS was also implemented. Its equation is:

$$\mathrm{Y} = \mathrm{X}\beta + \mathrm{u} \qquad 5\text{-}5$$

The notations remain the same as the Eq. 5-1. Compared with the SAR model,pparently the OLS neglects the spatial lag (or error) effects on the regression model.

Both Lagrange Multiplier (LM) statistics and Robust Lagrange Multiplier (RLM) statistics were examined after running the OLS regression model. The two statistics were used to determine which SAR model (SLM or SEM) to use in our research. Anselin (2005) proposed a diagram to depict the decision process. I briefly explain the procedure as follows. It is known that the null hypotheses of the SLM and SEM are $\rho = 0$ and $\lambda = 0$, respectively. If the LM test of the two models both fails to reject the null hypothesis (e.g., LM p-value > 0.05) after running the OLS, then an OLS model is appropriate to use. If the spatial lag LM statistic is significant, the SLM should be chosen; if the spatial error LM statistic is significant, the SEM should be selected. However, it is likely for LM statistics of two models to be significant. In this case, we need to check the RLM statistics further. If the spatial lag RLM statistic is significant, we use a spatial lag model; if the spatial error RLM statistic is significant, we use a spatial error model. If RLM statistics for both models are still significant, use the model with the largest value.

As per the Eq. 5-2 and 5-4, the weight matrix W is critical to the two models. Weights matrix measures the influence that closer things have more impact than things far away. When neighboring values are closer, the more weights are assigned to them. There are many options to create spatial weight matrix; for example, distance-based matrix and K neighbors. However, these weight matrixes are not contiguity-based. I need to use contiguity-based matrix because I would like to see whether block groups are next to tend to cluster with each other block groups in terms of spatial food access. The contiguity-based matrix includes two types: queen and rook contiguity. Queen contiguity means neighbors and their adjacent units sharing common boundaries and vertices that are similar to Queen Movement on a chess board. Whereas Rook contiguity similar to Rook Movement only considers an adjacent unit with a shared border as neighbors. In this study, we employed queen contiguity because the influence is going across all corners of the units. The order of contiguity must be specified as well. For instance, a first-order contiguity matrix would only include direct neighbors, but a second-order contiguity matrix would include weights for the neighbors' neighbors as well. To decide the best order of spatial weights, I tested from 1^{st} to 6^{th} order for queen contiguity. The 1^{st} order was chosen because the histogram distribution of connectivity approached normal using this order.

Several vital statistics such as R square of the SAR model would be reported. However, as Anselin (2005) claimed that the R square of the SAR is spatial pseudo-R^2, and it cannot be comparable with the R square of the OLS. More reliable measures of fitness are Log-Likelihood (LL), Akaike info criterion (AIC), and Schwarz Criterion (SC). A higher number Log-likelihood indicates a better model; while a lower number of

AIC and SC indicates better goodness of fit of the model. Lastly, we need to perform a diagnostic on the model. The Breusch-Pagan statistic is a measure of heteroscedasticity of the errors. A Likelihood Ratio (LR) tests the significance of spatial autoregressive coefficient. A high significance with LR value means the spatial effects in the data have not been removed completely (Anselin 2005).

Employing the semi-parametric Geographically Weight Regression

A classic (or full model) GWR model is written as:

$$y_i = \sum_k \beta_k (\mu_i, v_i) x_{k,i} + \varepsilon_i \qquad 5\text{-}6$$

where y_i is dependent variable at location i, $x_{k,i}$ is the kth independent variable at location i, ε_i is error term at location i; (μ_i, v_i) is the x-y coordinate of the ith location; and $\beta_k(\mu_i, v_i)$ are a set of varying coefficients of independent variables at location i.

Semi-parametric GWR is a critical extension of the classic GWR. It mixes globally fixed and geographically varying coefficients together. Its equation is updated as:

$$y_i = \sum_k \beta_k (\mu_i, v_i) x_{k,i} + \sum_l \gamma_l z_{l,i} + \varepsilon_i \qquad 5\text{-}7$$

where $z_{l,i}$ is the l^{th} independent variable with a fixed coefficient γ_l. Other notations remain the same with the Eq. 5-6.

The semi-parametric GWR is under the framework of geographically weighted generalized linear modeling (GWGLM), which consists of regular models such as Gaussian (continuous scale data), Poisson (count data), and logistic regressions (binary data). In in study, I used the Gaussian function model since the dependent variables are scale data. Model performance was evaluated based on the Akaike Information Criterion (AICc). The AICc method considers the fact that the degrees of freedom may vary among models, and thus it is a robust indicator to compare the goodness of fit in different models. A model with lower AICc indicates a better model fit. Its equation is shown below.

$$\mathrm{AIC}_c = 2nlog_e(\hat{\sigma}) + n\ log_e(2\pi) + n\left\{\frac{n+tr(S)}{n-2-tr(S)}\right\} \qquad 5\text{-}8$$

where n is the number of observations in the dataset, $\hat{\sigma}$ is the estimate of the standard deviation of the residuals, and tr(S) is the trace of the hat matrix.

To determine which parameters exhibited significant spatial variations and which ones did not the GWR4 software is designated to test the geographical variability of local coefficients. Suppose that I intend to test whether the k^{th} variable has a varying geographical coefficient, the following procedure would be iteratively taken: (1) It fits the model with all variables geographically varying (full GWR model) and calculates the model's goodness of fit (i.e., AICc value). 2) it then makes the k^{th} coefficient fixed while other coefficients are kept as they were in the Full GWR model, and this is called switched GWR model. It calculates the switched model's AICc value. If the switched GWR model outperforms the Full GWR model, indicating that the k^{th} variable is varying over space. It means that the switched GWR model has a smaller AICc value. It computes the difference of AICc values between the two models, and the difference is named as “Diff of Criterion” by its developers. The test repeats this comparison for each geographically varying coefficient and reports the “Diff of Criterion” value for each variable. If the value of “Diff of Criterion” for variables is negative, it indicates that there is significant spatial variability in the associated coefficient. Nakaya, et al. (2017) claimed that the model comparison is strongly supported when the absolute value of

“Diff of Criterion” is larger than 4. If the difference is positive, which means that there is no spatial variability in that variable; it would be better to represent this variable as a global or spatially fixed term in the model. Besides, GWR4 provides two routines to determine whether a semi-parametric approach is preferred; they are GtoL and LtoG. Simply speaking, the GtoL (geographically varying to fixed) approach begins with a full GWR model and then compares models to find the optimal combination of varying and fixed explanatory variables. Whereas the LtoG (fixed to geographically varying) one is the reversed procedure of the GtoL. Regarding the two approaches, one can refer to GWR4 manual guide for more information (Nakaya, et al. 2017).

I employed adaptive spatial kernels using a bi-square function, as defined by Eq. 5-9 (Nakaya, et al. 2017). The adaptive kernel uses the same number of observations at each regression point in the estimation. In some cases, one can choose a fixed bi-square kernel function when every expands in areas of sparse observations and shrinks in areas of dense observations. This research does not involve in sparse or dense observations, and it is secure to use an adaptive bi-square function. Meanwhile, the golden section search method was used to determine optimal bandwidths.

$$w_{ij} = \begin{cases} (1 - \frac{d_{ij}^2}{\theta_{i(k)}})^2 & d_{ij} < \theta \\ 0 & d_{ij} > \theta \end{cases} \qquad \text{5-9}$$

where w_{ij} is the weight value of the observation at location j for eastimating the coefficient at location i; d_{ij} is the Euclidean distance between i and j; θ is a fixed bandwidth defined by a distance metric measure; $\theta_{i(k)}$ is an adaptive bandwidth size defined as the kth nearest neighbor distance.

Identifying food deserts and food swamps with hotspot analysis

Figure 5.1 represents the specific steps to identify food deserts and food swamps. There are four steps. Step I, II, and III have been completed in previous sections. The last one is to use the hotspot analysis to achieve the objective.

As per the method, we used the hot spot analysis to define food deserts and food swamps. Its equation is:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \frac{\sum_{j=1}^{n} x_j}{n} \sum_{j=1}^{n} w_{i,j}}{\sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\frac{\sum_{j=1}^{n} x_j}{n}\right)^2} \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \qquad 5\text{-}10$$

where G_i^* is a statistic of hot spot analysis at location i; x_j is the value for feature j, $w_{i,j}$ is the spatial weight between feature i and j; n is the total number of features.

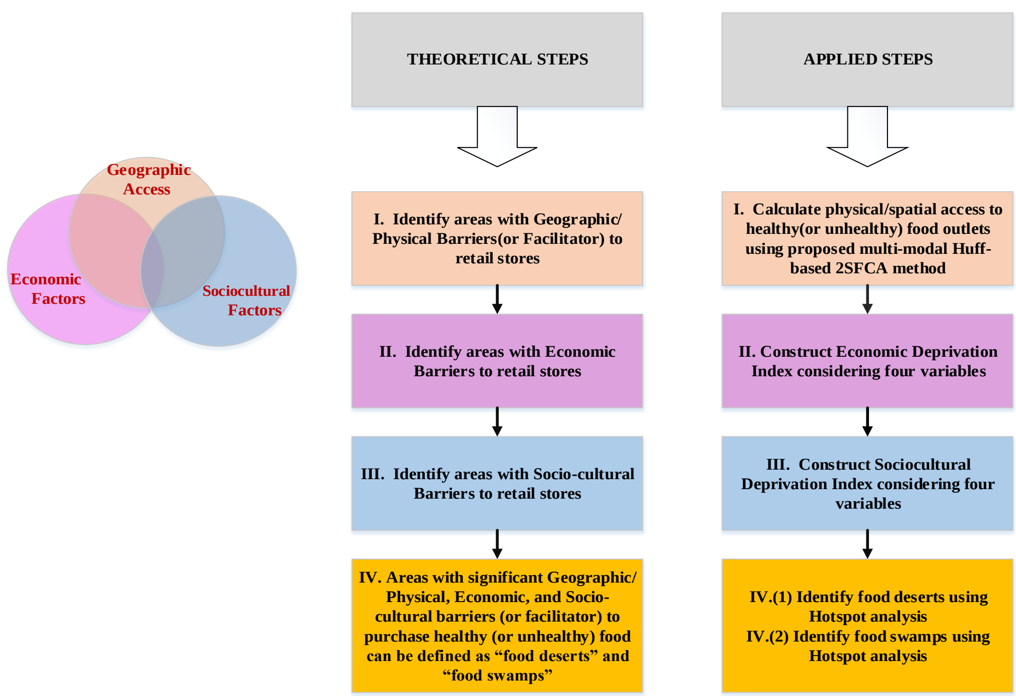


Figure 5.1 Theoretical and applied steps of identifying food deserts and food swamps.

Hot spot analysis was performed in ArcMap 10.5. Hot spot analysis maps were used to show where the statistically significant hot spots and cold spots are. The G_i^* statistic returns a z-score known as standard deviation. Three significance levels p (0.1, 0.05, and 0.01) were used. Based on the z-score and the significance levels p, block groups could be characterized by one of the three types (see Table 5.2). Block groups are identified as "Not Significant," meaning that there is no significant spatial clustering in these locations. While block groups in hot (or cold) spots indicate that high (or low) values are clustering in a specific area. For example, the hot (or cold) spots of $LnSAI_H$ depict that high (or low) access to healthy food outlets clusters in some block groups and their adjacent ones. The hot spots of EDI (or SDI) imply a severe economic (or sociocultural) deprivation in these areas.

For the delineation of the food deserts, I conducted three hot spot analyses on $LnSAI_H$, EDI, and SDI, respectively. The cold spots of $LnSAI_H$, hot spots of EDI, and hot spots of SDI are intersected to define food desert (Table 5.3). The concept of food oases is opposite to food deserts. They are areas where high access to healthy foods, low economic and sociocultural deprivation. Therefore, I intersected hot spots of $LnSAI_H$, cold spots of EDI, and cold spots of SDI to identify food oases (Table 5.3). For the identification of food swamps, provided that EDI is not a significant predictor of $LnSAI_U$ (the result of spatial error model and semi-parametric GWR), I did not consider the effect of the EDI on the $LnSAI_U$. Instead, I only two factors: SDI and $LnSAI_U$. These two variables were selected to define food swamps (Table 5.3).

Table 5.2 Potential results for hot spot analysis.

Relationship	**Explanation**
Not Significant	A Z score near zero and large p-value for a feature indicates no spatial clustering
Hot Spot	A high positive Z score and small p-value for a feature indicates a significant hot spot; the higher the Z score, the more intense the clustering
Cold Spot	A low negative Z score and small p-value indicate a significant cold spot; The lower the Z score, the more intense the clustering.

Table 5.3 Delineation of food deserts, food oases, and food swamps.

	Variables		**Intersect**
Food Deserts	$LnSAI_H$	Cold Spot	Yes
	EDI	Hot Spot	
	SDI	Hot Spot	
Food Oases	$LnSAI_H$	Hot Spot	Yes
	EDI	Cold Spot	
	SDI	Cold Spot	
Food Swamps	$LnSAI_U$	Hot Spot	Yes
	EDI	NA	
	SDI	Cold Spot	

Results

Result of the two indices: EDI and SDI

The summary of the four economic indicators is shown in Table 5.4. The distribution of the four indicators is depicted in Figure 5.2. Census block groups with low household income, high unemployment rate, a high proportion of households below the poverty line, and a high percentage of households without complete kitchen facility are mainly located in the East of IH-35.

Table 5.4 Descriptive statistics of the economic variables in the study area.

Variables	Min	1st Quartile	Median	3rd Quartile	Max	Mean	SD
Median Household Income ($1,000)	5.66	37.43	54.54	74.67	199.44	60.46	32.61
Unemployment Rate (%)	3.85	19.35	25.64	31.60	77.32	26.74	10.73
Below the Poverty Line (%)	0.00	5.14	11.79	21.80	100.00	16.28	16.58
Without a complete kitchen facility (%)	0.00	0.00	0.00	2.02	21.76	1.61	3.19

Table 5.5 shows the correlations between the four variables. Median household income and percentage of household below the poverty level were significantly correlated, and its relationship was larger than 0.50 (r = -0.635, p = 0.000); meanwhile, median household income was significantly associated with the percentage of households without complete kitchen facility (r = -0.109, p = 0.017). Unemployment rate and percentage of household below the poverty level were also significantly correlated (r=

0.447, p = 0.000). As a result, the multicollinearity problem is likely to present in the four indicators.

Table 5.5 Spearman correlations between economic variables.

	Median Household Income	**Unemploy ment Rate**	**Below the Poverty Line**	**Without Complete Kitchen Facilities**
Median Household Income	1	-0.078 (p = 0.087)	-0.635** (p = 0.000)	-0.109* (p = 0.018)
Unemployment Rate		1	0.447** (p = 0.000)	0.067 (p = 0.150)
Below the Poverty Line			1	0.046 (p = 0.316)
Without complete kitchen facilities				1

Table 5.6 Component score coefficient matrix of factor analysis.

	Economic Deprivation Index (EDI)
Median Household Income	-0.428
Unemployment Rate	0.309
Below the Poverty Line	0.508
Without complete kitchen facilities	0.112

Extraction Method: Principal Component Analysis.

As shown in Table 5.6, only one component was identified using factor analysis, and it is named as EDI. The eigenvalues suggest that this one factor accounts for more than 80% of the total variance of the four variables. The most important two variables contributing to this score were the proportion of household below the poverty level (factor loading FL = 0.508) and median household income (FL = -0.428). The final measure EDI of each block group was calculated using the coefficients in Table 5.6. The spatial pattern of the EDI is shown in Figure 5.3. The higher of the index means a higher deprivation, and the most economically deprived areas (i.e., >2) were on the campus of UT Austin (the University of Texas at Austin) and the East of IH-35.

The summary of the four sociocultural indicators is shown in Table 5.7. Figure 5.4 depicts the spatial pattern. The census block groups with a high percentage of home renters, high proportion of people without higher education, and high percentage of Hispanic people, high percentage of linguistical isolation were also located in the East of IH-35.

Table 5.8 shows Pearson correlations of the four sociocultural variables. These variables were all significantly correlated (i.e., $p = 0.000$). The percentage of Hispanic people and the proportion of people without higher education were highly correlated ($r = 0.836$, $p = 0.000$). Similarly, I performed factor analysis on these four variables in SPSS 25.0, and one component was identified: SDI. The most important two variables contributing to the SDI were the percentage of Hispanic people (FL = 0.359) and the proportion of people without higher education (FL = 0.345). The final measure of the SDI of each block group was calculated using the coefficients in Table 5.9. The spatial pattern of the SDI is shown in Figure 5.5. The most sociocultural deprived areas are also found in the East of the IH-35.

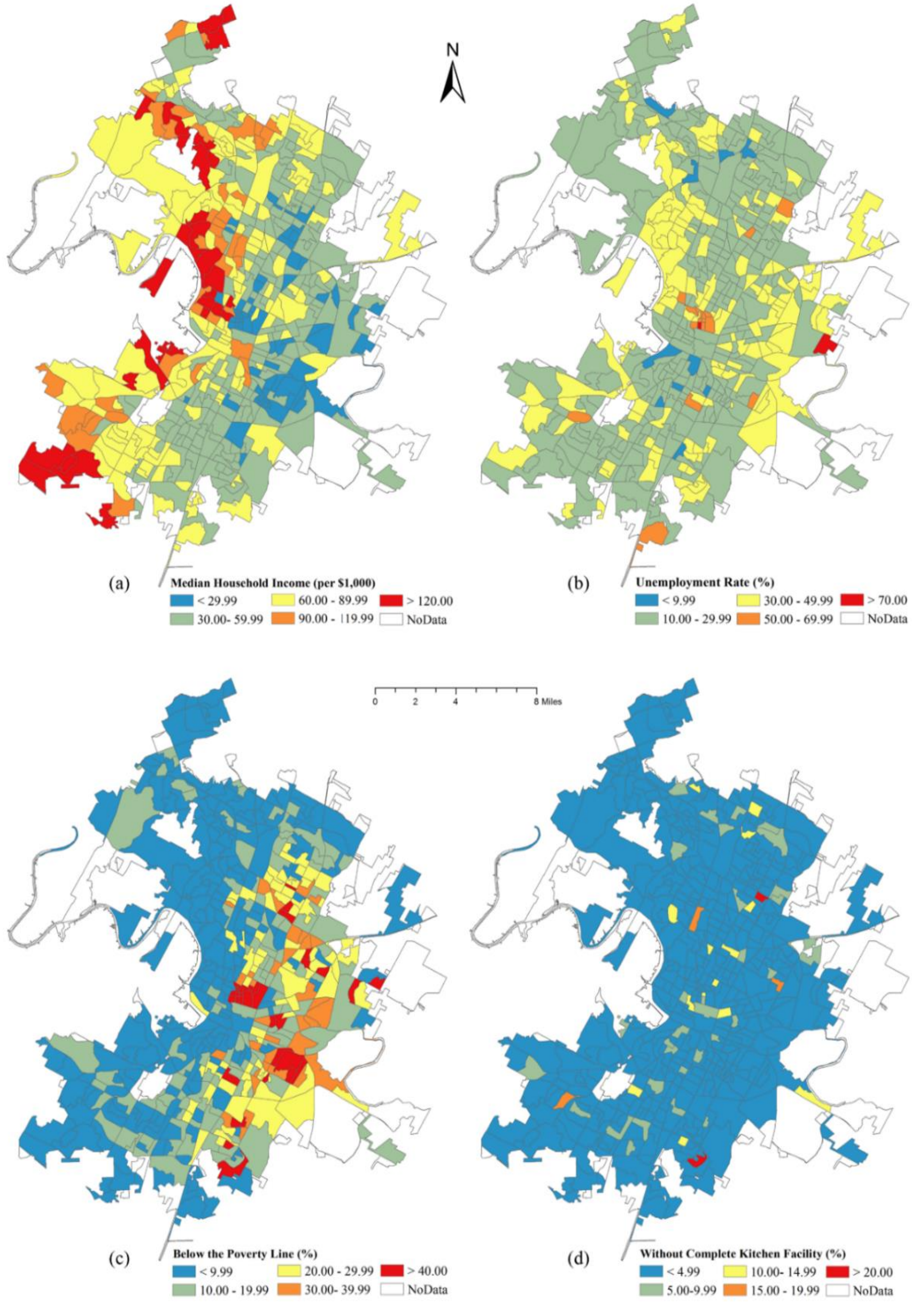


Figure 5.2 Spatial distribution of the four economic variables in Austin, Texas.

Table 5.7 Descriptive statistics of sociocultural variables in the study area.

Variables	Min	1st Quartile	Median	3rd Quartile	Max	Mean	SD
Home Renter (%)	0.00	29.47	51.03	74.06	100.00	52.50	28.63
Without Higher Education (%)	0.00	32.79	50.37	73.2	100.00	52.49	24.57
Hispanic People (%)	0.00	12.37	24.71	50.96	100.00	32.26	24.12
Linguistically Isolation (%)	0.00	0.00	3.43	8.76	46.62	6.81	9.38

Table 5.8 Bivariate Pearson correlations between sociocultural variables.

Variables	Home Renter	Without Higher Education	Hispanic People	Linguistically Isolation
Home Renter	1	0.233** (p = 0.000)	0.232** (p=0.000)	0.331** (p = 0.000)
Without Higher Education		1	0.836** (p=0.000)	0.618** (p = 0.000)
Hispanic People			1	0.715** (p = 0.000)
Linguistically Isolation				1

Note: **. Correlation is significant at the 0.01 level.

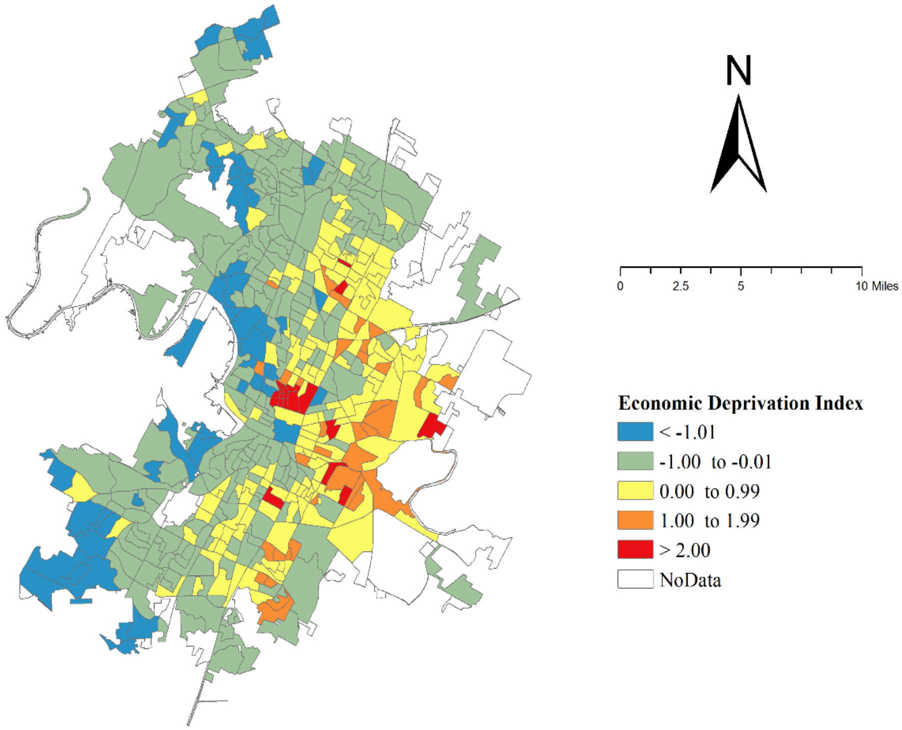


Figure 5.3 Spatial distribution of EDI in Austin, Texas.

Table 5.9 Component score coefficient matrix of factor analysis.

Variables	Sociocultural Deprivation Index (SDI)
Home Renter	0.174
Without Higher Education	0.345
Hispanic People	0.359
Linguistically Isolation	0.333

Extraction Method: Principal Component Analysis.

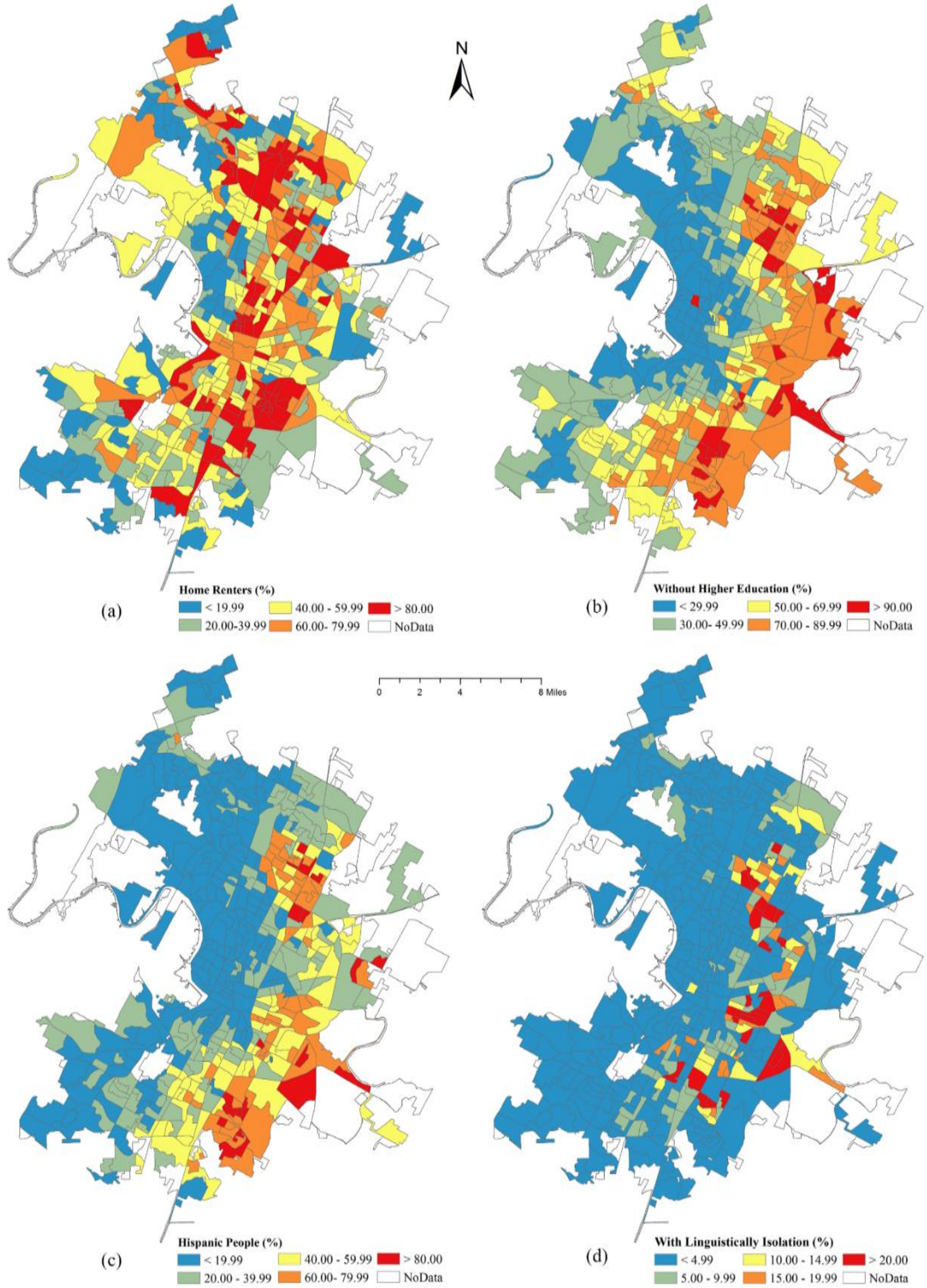


Figure 5.4 Spatial distribution of the four sociocultural variables in Austin, Texas.

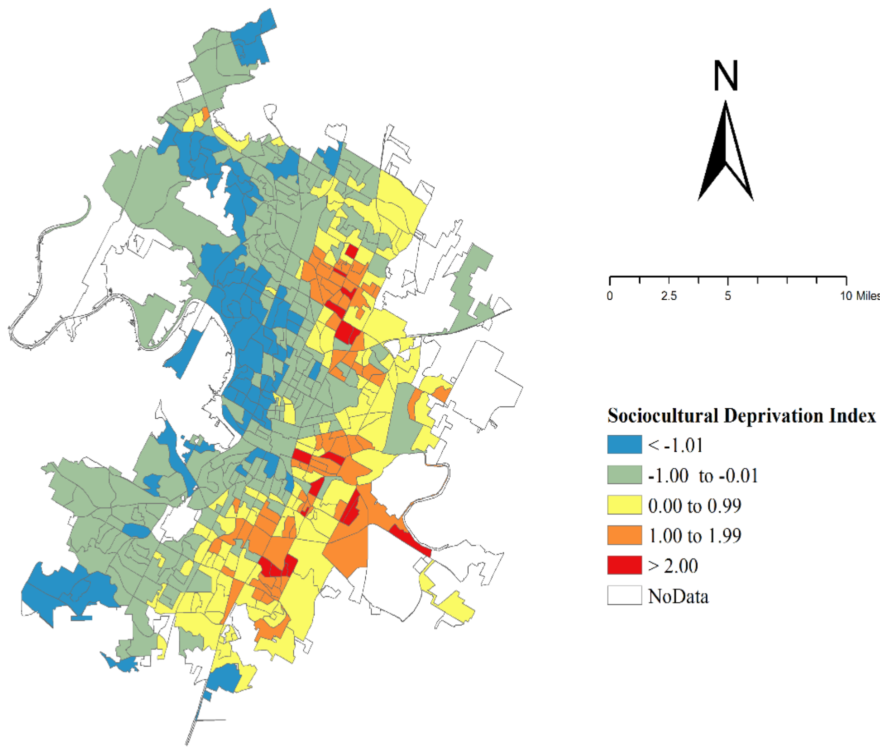


Figure 5.5 Spatial distribution of the SDI in census block groups in Austin, Texas.

Result of the SAR model

Figure 5.6 (a) and (c) show the histogram of the original spatial accessibility index to healthy and unhealthy food outlets; their distributions are not normal and right-skewed. They both look like a “bell shape” and became symmetric after the logarithm transformation (Figure 5.6 (b) and (d). I further tested the normality using skewness and kurtosis. Skewness measures the asymmetry of the distribution, and kurtosis measures the peakedness of the distribution. Table 5.10 depicts the skewness and kurtosis of the variables. As suggested by Kim (2013), when sample sizes greater than 300, either an absolute skewness value larger than 2 or an absolute kurtosis value larger than seven may be used as cutting points to determine substantial non-normality. The skewness and kurtosis of the two original spatial accessibility indices (i.e., 4.110, 25.692; 2.275, 7.167) are substantially abnormal. These values became smaller (i.e., 0.486, 1.279; 0.214,0.297) after the logarithm transformation, which indicates that no substantial non-normality present in the transformed data. Shapiro-Wilk test was used to test the normality of the data (Table 5.10). Note that the data is still not perfectly normal after the transformation because both p values of the Shapiro-Wilk test are smaller than 0.05 (p = 0.000 and p = 0.030). Nevertheless, I still used the transformed data since there is no substantial non-normality in the data, and the data with perfect normality is rarely encountered in the real world.

Table 5.10 Normality test using skewness, kurtosis, and Shapiro-Wilk test.

	Skewness		**Kurtosis**		**Shapiro-Wilk**	
	Statistic	**Std. Error**	**Statistic**	**Std. Error**	**Statistic**	**p**
SAI_H	4.110	0.113	25.692	0.226	0.641	0.000
$LnSAI_H$	0.486	0.113	1.279	0.226	0.974	0.000
SAI_U	2.275	0.113	7.167	0.226	0.795	0.000
$LnSAI_U$	0.214	0.113	0.297	0.226	0.993	0.030

Table 5.11 shows the statistics of the two log-transformed spatial food accessibility indices and deprivation indices. In general, the transformed data and deprivation data are in the same range, and they have comparable magnitude. The accessibility to unhealthy food outlets is larger than that of unhealthy ones. The two deprivation variables are the result of factor analysis, and they are standardized during the process. Thus, their means are zero, and standard deviations are 1.

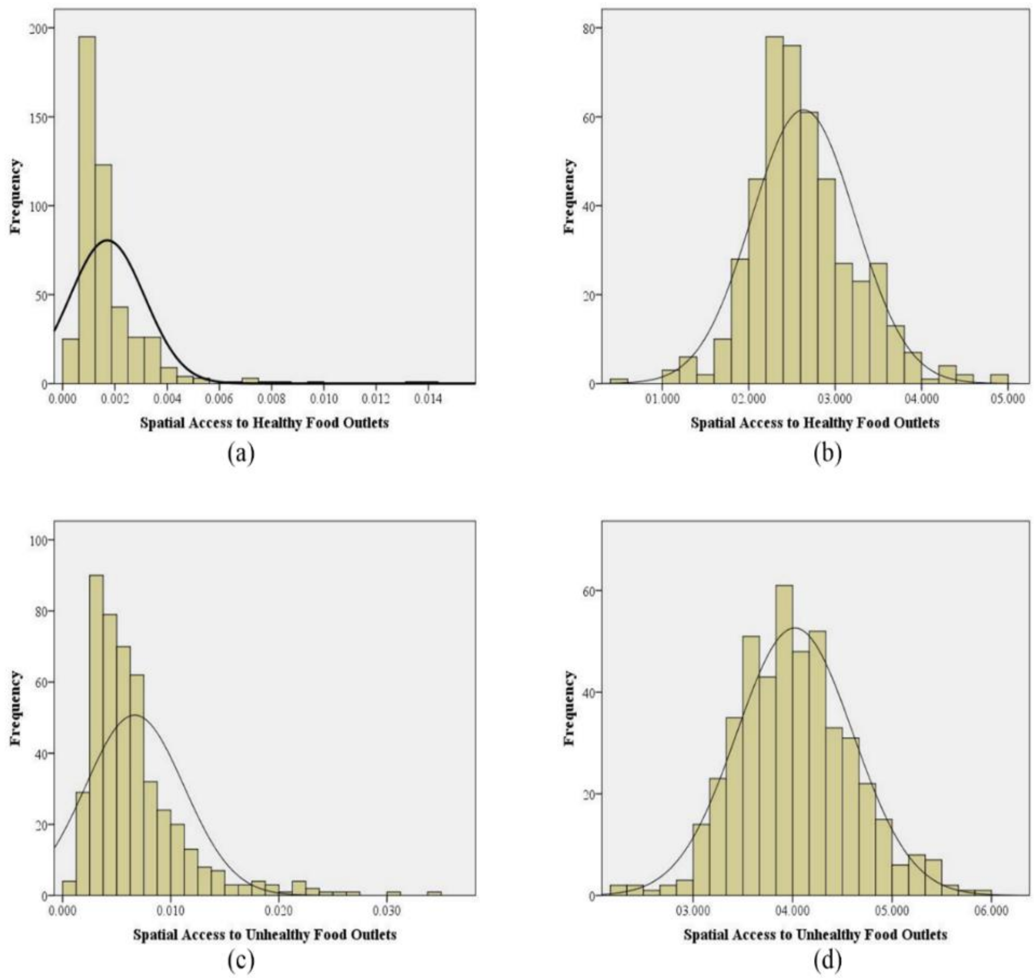


Figure 5.6 Histogram of the SAI_H before (a) and after (b) the logarithm transformation; histogram of the SAI_U before (c) and after (d) the logarithm transformation.

Table 5.12 shows that the two spatial accessibility measures ($LnSAI_H$ and $LnSAI_U$) have a significant correlation ($r = 0.682$, $p = 0.000$). SDI and EDI are significantly associated ($r = 0.548$, $p = 0.000$). $LnSAI_H$ is significantly associated with SDI ($r = -0.102$, $p = 0.025$) but insignificantly with EDI. $LnSAI_H$ and the both deprivation indices (SDI and EDI) are not significantly associated ($p > 0.05$).

Table 5.11 Descriptive statistics of the four indices in the study area.

	Min	**1st quartile**	**Median**	**3rd quartile**	**Max**	**Mean**	**SD**
$LnSAI_H$	0.000	2.238	2.542	2.940	4.942	2.562	0.731
$LnSAI_U$	0.000	3.578	3.970	4.379	6.000	3.925	0.860
EDI	-2.365	-0.642	-0.126	0.446	4.484	0.000	1.000
SDI	-1.626	-0.803	-0.216	0.541	3.819	0.000	1.000

Table 5.12 Bivariate Pearson correlations between the four indices in the study area.

	$LnSAI_H$	**$LnSAI_U$**	**EDI**	**SDI**
$LnSAI_H$	1	0.682** (p = 0.000)	-0.086 (p = 0.062)	-0.102* (p = 0.025)
$LnSAI_U$		1	-0.003 (p = 0.953)	0.079 (p = 0.086)
EDI			1	0.548** (p = 0.000)
SDI				1

**. Correlation is significant at the 0.01 level.

To specify which SAR model to use, we ran the OLS firstly. Table 5.13 shows the statistics after running the OLS. The columns on the left show the result of the dependent variable $LnSAI_H$. Both Lagrange Multiplier (lag) and Lagrange Multiplier (error) are significant with p = 0.000. Therefore, we need to further check Robust LM (lag) and Robust LM (error). Both Robust LM (lag) and Robust LM (error) show significant p-value (p =0.000 and p = 0.037). However, Robust LM (lag) has a larger value (18.013) than Robust LM (error) (4.343). As a result, I chose the SLM rather than the SEM. For the dependent variable $LnSAI_U$, we followed the same procedure to select the best model. The columns on the right show that Robust LM (error) has a larger value (74.494) than that of Robust LM (lag) (0.958). Therefore, I selected the SEM other than the SLM.

Table 5.13 Statistics of Lagrange Multiplier (lag) and Lagrange Multiplier (error).

Test	MI/DF [a]	Value [a]	P [a]	MI/DF [b]	Value [b]	P [b]
Moran's I (error)	0.384	13.746	0.000	0.450	16.072	0.000
Lagrange Multiplier (lag)	1.000	193.471	0.000	1.000	173.158	0.000
Robust LM (lag)	1.000	18.013	0.000	1.000	0.958	0.328
Lagrange Multiplier (error)	1.000	179.800	0.000	1.000	246.694	0.000
Robust LM (error)	1.000	4.343	0.037	1.000	74.494	0.000
Lagrange Multiplier (SARMA)	2.000	197.814	0.000	2.000	247.652	0.000

Note: **a** indicates that the dependent variable is $LnSAI_H$; **b** indicates that the dependent variable is $LnSAI_U$.

The result of SLM for $LnSAI_H$ is present in Table 5.14. The dependent variable is the $LnSAI_H$, and the independent variables are the two deprivation indices. The spatial lag coefficient is 0.671 and is highly significant (t= 16.612, p =0.000). The constant term is significant as well (p = 0.000). The coefficient of the EDI is -0.054 with p-value 0.037, which indicates that EDI has a significant negative effect on spatial access index to healthy food outlets. Whereas the coefficient of the SDI is 0.025 with p-value 0.339 (> 0.05), meaning that the SDI does not have a significant effect on healthy food outlets accessibility. In addition, the absolute value of the coefficient EDI is larger than the SDI, and it depicts that economic deprivation has more influences on healthy food accessibility. For a comparison purpose, we also report the result of OLS in the same table. Similar to the SLM, the constant term in the OLS is significant (p = 0.000). However, the confidence of the EDI and SDI are not significant (p = 0.135 and p =

0.167). The absolute value of the EDI coefficient is less relative to the one in the SLM, but for the SDI its absolute value of the coefficient is more compared with the SLM.

Table 5.14 The estimation of SLM for LnSAIH and the comparison between SLM and OLS.

	Spatial Log Model (SLM)					**Ordinary Least Square (OLS)**				
	Coef.	**t**	**p**	**Pse. R^2**	**Moran 's I**	**Coef.**	**t**	**p**	**Adj. R^2**	**Moran 's I**
W_LnSAH	0.671**	16.612	0.000							
I	0.865**	8.002	0.000			2.634**	95.160	0.000		
EDI	-0.054*	-2.089	0.037	0.390	0.004	-0.049	-1.496	0.135	0.015	0.384
SDI	0.025	0.958	0.339			-0.045	-1.384	0.167		
LL	-329.508					-415.509				
AIC	667.017					837.018				
SC	683.568					849.432				
BP	0.508 (p = 0.776)					1.082 (p = 0.582)				
LR	172.002 (p = 0.000)									

Note: C: Constant; LL: Log Likelihood; AIC (Akaike Info Criterion); SC: Schwarz criterion; BP: Breusch-Pagan test; LR: Likelihood Ratio test.

Pseudo- R2 of the SLM is 0.390, while the adjusted R2 of OLS is 0.015. However, these two values are not comparable. LL, AIC, and SC are more appropriate to examine whether model fit has been improved or not. The log LL increases from -415.509 (the OLS) to -329.508 (the SLM), which indicates that there are improvements of model fit using the SLM relative to the OLS. Both AIC (from 837.018 to 667.017) and SC (from 894.432 to 683.568) of the SLM decrease relative to OLS, again confirming that the SLM improves the model fit. Two tests, BP test for heteroscedasticity and LR test on the spatial lag coefficient, are employed in the model diagnostics. The values of the BP test are 0.508 with p-value 0.776 (the SLM) and 1.082 with p-value 0.582 (the OLS), and it indicates that the heteroscedasticity is not an issue for the model. The result of the LR test is 172. 002 (p = 0.000). Also, we checked the spatial dependence of the residuals in SLM.

The Moran's I index is 0.004 (essentially zero), which indicates that including the spatial lag term in the model has removed all spatial autocorrelation or spillover. The results and diagnostics confirm that the SLM is much more appropriate than the OLS in our study.

Table 5.15 The estimation of SEM for $LnSAI_U$ and the comparison between SEM and OLS.

	Spatial Error Model (SEM)					**Ordinary Least Square (OLS)**				
	Coef.	**t**	**p**	**Pse. R^2**	**Moran's I**	**Coef.**	**t**	**p**	**Adj. R^2**	**Moran's I**
C	3.992**	59.980	0.000			4.022**	149.055	0.000		
EDI	-0.043	-1.262	0.206	0.409	-0.044	-0.071*	-2.220	0.027	0.014	0.450
SDI	0.160**	4.147	0.000			0.089**	7.011	0.006		
LAMBDA	0.693**	16.780	0.000			NA	NA			
LL	-312.262					-430.720				
AIC	630.524					813.440				
SC	642.937					825.853				
BP	3.285 (p = 0.193)					2.202 (p = 0.333)				
LR	182.916 (p = 0.000)									

The result of SEM for $LnSAI_U$ is present in Table 5.15. The dependent variable is the $LnSAI_U$, and the independent variables are the two deprivation indices. The spatial error coefficient LAMBDA is 0.693 and is highly significant (t= 16.780, p =0.000). The constant term is significant (t = 59.98, p = 0.000). The coefficient of the EDI is -0.043 with p-value 0.206, which indicates that the EDI has an insignificant negative effect on spatial access index to unhealthy food outlets. While the coefficient of the SDI is 0.160 with p-value 0.000, which means that the SDI does have a significant effect on unhealthy food outlets accessibility. In addition, the absolute value of the coefficient SDI is larger than the EDI, and it reveals that the sociocultural deprivation has more influences on unhealthy food accessibility. For the result of the OLS, the constant term in the OLS is significant ($p = 0.000$); the coefficients of the EDI and SDI are both significant ($t = -2.220$, $p = 0.027$; $t = 7.011$, $p = 0.006$). The absolute value of the EDI coefficient is less relative to the one in the OLS, but for the SDI its absolute value of the coefficient is more compared with the OLS. The LL increases from -430.720 (the OLS) to -312.262 (the SEM), which indicates that there are improvements of model fit using the SEM relative to the OLS. Both AIC (from 813.440 to 630.524) and SC (from 825.853 to 642.937) of the SEM decrease relative to the OLS, again confirming that the SEM improves the model fit. The values of the BP test are 3.285 with p-value 0.193 (the SEM) and 2.202 with p-value 0.333 (the OLS), suggesting that the heteroscedasticity is not an issue for the model. The result of the LR test is 172. 002 ($p = 0.000$). The Moran's I of the SEM is -0.044 (essentially zero), which indicates that including the spatial error term in the model has removed spatial autocorrelation compared with the OLS.

Result of the S-GWR model

Table 5.16 shows the results of the conventional full-model GWR with dependent variable $LnSAI_H$. It assumes that the coefficients of the SDI and EDI are varied across geographic space. The quartile values, the mean values, as well as standard deviations for each of the coefficients are seen in Table 5.16. I reported two measures of model fit: AICc and F statistic. The AICc in the table is 971.607 with F statistic 2.794 ($p = 0.000$). The positive ‘Difference of Criterion’ values indicates that there are no significant spatial variations in the effects of the SDI and EDI. Therefore, a semi-parametric (or partial) GWR model is needed to calibrate the classic GWR model

Table 5.16 Summarized results from the full-model GWR model for $LnSAI_{H.}$

Variables	**Min**	**1st quartile**	**median**	**3rd quartile**	**Max**	**Mean**	**SD**	**Diff of Crit.**
Intercept	1.300	2.523	2.685	2.885	4.054	2.743	0.359	-29.097
SDI	-1.001	0.028	0.154	0.374	1.447	0.186	0.315	4.973
EDI	-0.645	-0.252	-0.116	0.056	1.132	-0.088	0.236	15.114
AICc				971.607				
F statistic				2.794				

Note: positive value of Diff of Criterion AICc suggests no spatial variability in terms of model selection criteria.

I manually performed the LtoG selection method to determine global variable(s). The results are summarized in Table 5.17. The EDI turns into the global variable with a coefficient -0.165 and p-value 0.000. It indicates that the EDI has a constant significantly adverse effect on the dependent variable $LnSAI_{H.}$ Only the SDI is a geographically varying variable, which could be confirmed via the negative value of "Diff of Criterion" (i.e., -7.323). The model fit has been improved compared with the classic GWR as the value of AICc becomes smaller (i.e., 956.42 vs. 971.607), and the F statistic becomes larger (i.e., 3.376 vs. 2.794). It again confirms that the semi-parametric GWR is better than the classic (or full model) GWR. It is a preferred model in our study. The overall coefficient of determination (adjusted R^2) of the model is 0.251. Moran's *I* test (-0.004, p =0.771) suggested that the error terms were randomly distributed.

Table 5.17 Summarized results from semi-parametric GWR with LtoG variable selection in GWR 4 for $LnSAI_H$.

Variables		Min	1st quartile	Median	3rd quartile	Max	Mean	SD	Diff of Crit.
Local	Intercept	1.435	2.497	2.681	2.928	3.720	2.723	0.321	-35.514
	SDI	-1.055	0.048	0.198	0.363	1.854	0.232	0.341	-7.323
Global		Value	t value	p-value					
	EDI	-0.165	-3.343	0.000					
AICc					956.492				
F statistic					3.376				
Adj. R^2					0.251				
Moran's I					-0.004 (p =0.771)				

The maps in Figure 5.7 show the spatial distribution of the model parameters and model performance for the semi-parametric GWR analysis. Figure 5.7(a) illustrates the distribution of local R^2 from the model. The values of local R^2 varied over the study area, suggesting the prediction power of the model was not consistent across the different block groups in Austin. Generally, the block groups in the east and southwest had the best regression results (with local R^2 larger than 0.450), whereas the mid-north and mid-south had the worst outcomes (smaller than 0.149). Figures 5.7(b) illustrates the residuals from the GWR model; it indicates that more than 50% block groups were overestimated (negative values, in blue colors), and the remaining block groups were underestimated (positive values, in red colors). Meanwhile, the distribution of residuals seems random, echoing Moran's *I* result from Table 5.17 and indicating no spatial clustering in the residuals.

The EDI is a global (or fixed) variable in the estimation, which means that it has a constant relationship with the $LnSAI_H$. As a result, the distribution of its estimation was not mapped. Instead, I mapped out the estimations of intercept and SDI. Figure 5.7(c) shows the intercept term represented the distributions of $LnSAI_H$ when the EDI and SDI equaled zero. We observed that higher intercept values ($LnSAI_H$) were located in the urban center and northwest of Austin, whereas lower intercept values ($LnSAI_H$) were in the east of Austin. GWR4 reports the t values for the intercept. The range of the t values is 3.139 to 25.465. The critical t value is 1.96 when p = 0.05. Thus, it indicates that the intercept values were all significantly positive across the study area. Figure 5.7(d) illustrates the spatial variation in the association between SDI and $LnSAI_H$. The range of t values for the coefficients of SDI is -2.812 to 5.394. The coefficients with t values between -1.96 and 1.96 were not significant, which implies that the parameter estimations in these areas were not reliable. Thus, I masked these areas with gray grids. The remaining block groups had a significant association between SDI and $LnSAI_{H.}$ The distribution of the parameter SDI showed a distinctly spatially non-stationary pattern. Most of these block groups exhibit significantly positive relationship, especially in the southwest and mid-west of Austin. It implies that higher SDI tended to associate with higher $lnSAI_H$ in these areas. On the contrary, only a few block groups (in blue colors) exhibit significantly negative relationship, representing lower $LnSAI_H$ related to higher SDI.

Table 5.18 shows the results of the classic GWR with the dependent variable $lnSAI_U$. The quartile values, the mean values, as well as standard deviations for each of the coefficients are reported. We reported two measures of model fit: AICc and F statistic. The AICc in the table is 1148.110 with F statistic 2.502 (p = 0.000). The positive 'Difference of Criterion' values indicates that there are no significant spatial variations in the effects of the variable SDI and EDI. Therefore, a semi-parametric (or partial) GWR model is needed to calibrate the classic GWR model.

LtoG selection method was performed to determine global variables. The results are summarized in Table 5.19. The EDI becomes a global variable with a coefficient -0.092 and p-value 0.061. It indicates that the EDI has a constantly insignificantly adverse effect on the dependent variable $lnSAI_{U.}$ The SDI is a geographically varying variable, which could be confirmed via the negative value of "Diff of Criterion" (i.e., -0.395). The model fit has been improved compared with the classic GWR since the value of AICc becomes smaller (i.e., 1133.907 vs. 1148.110), and the F statistic becomes larger (i.e., 2.951 vs. 2.502). It again confirms that the semi-parametric GWR is better than the classic GWR. The overall coefficient of determination (adjusted R^2) of the model is 0.214. Moran's *I* test (-0.010, p =0.291) suggested that the error terms were randomly distributed.

Table 5.18 Summarized result s from the classic GWR model for LnSAIU.

Variables	**Min**	**1st quartile**	**median**	**3rd quartile**	**Max**	**Mean**	**SD**	**Diff of Crit.**
Intercept	2.897	3.878	4.042	4.294	5.022	4.071	0.33	-0.742
SDI	-1.009	0.09	0.209	0.423	2.198	0.256	0.369	13.521
EDI	-0.912	-0.232	-0.067	0.077	2.31	-0.05	0.314	14.203
AICc				1148.11				
F statistic				2.502				

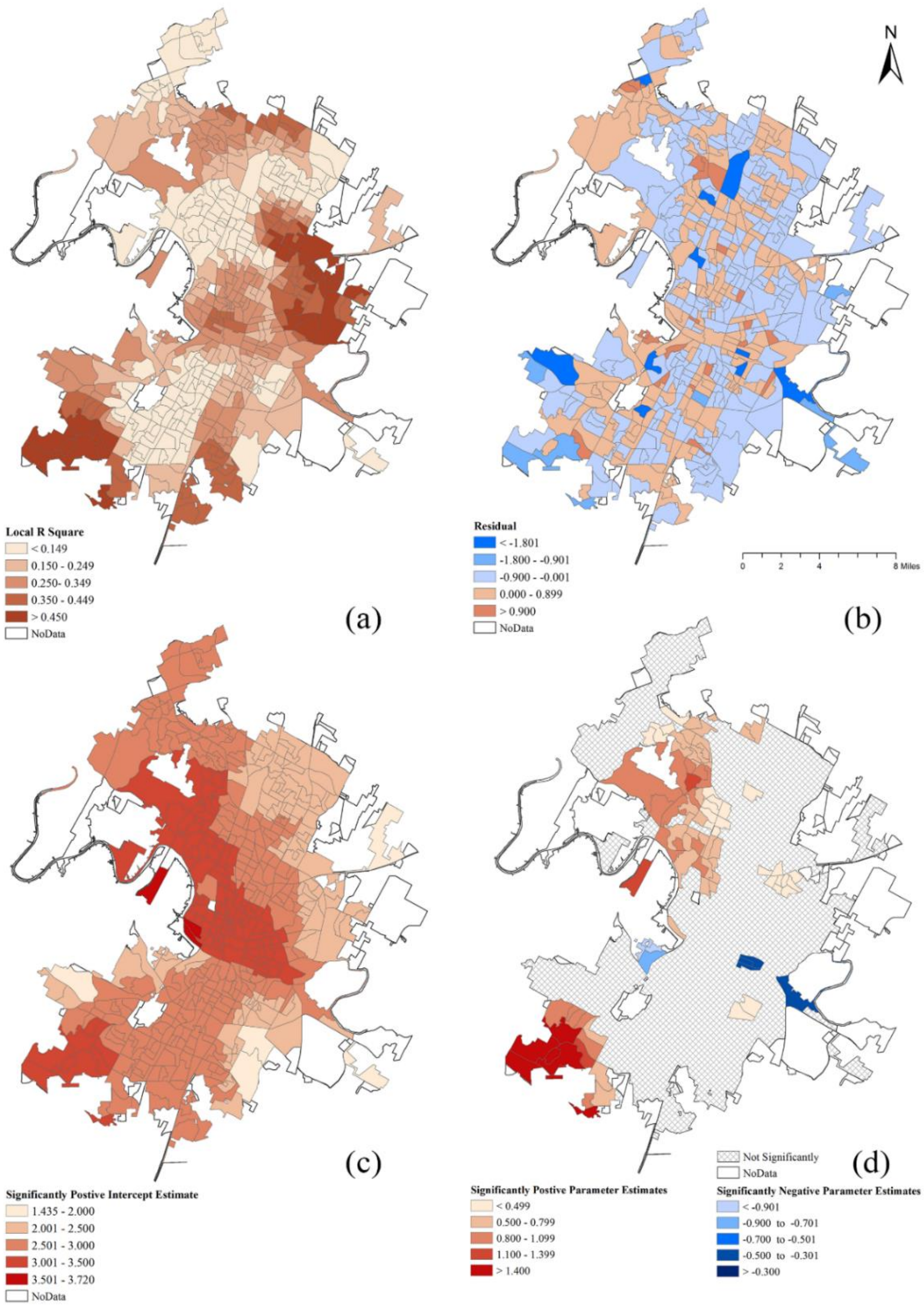


Figure 5.7. Result of S-GWR with $LnSAI_H$ as the dependent variable: (a) local R^2; (b) residual; (c) spatial variation in the coefficients of intercept; (d) spatial variation in the coefficients of SDI.

Table 5.19 Summarized results from partial GWR with Local to Global selection in GWR 4 for $LnSAI_U$.

Variables		Min	1st quartile	median	3rd quartile	Max	Mean	SD	Diff of Crit.
Local	Intercept	2.861	3.905	4.07	4.359	4.925	4.105	0.328	-14.585
	SDI	-1.394	0.086	0.215	0.529	2.683	0.317	0.445	-0.395
Global		Value	t value	p-value					
	EDI	-0.092	-1.554	0.061					
AICc					1133.907				
F statistic					2.951				
Adj. R^2					0.214				
Moran's I					-0.010 p = 0.291				

Figure 5.8 (a) illustrates the distribution of local R^2 from the model. The block groups in the northeast and southwest had the best regression results (with local R^2 larger than 0.400), whereas the mid-south had the worst outcomes (smaller than 0.099). Figures 5.8(b) illustrates the residuals from the GWR model; block groups with negative values (blue colors) were overestimated while with positive values (red colors) underestimated. Figure 5.8(c) shows the intercept term represented the distributions of $LnSAI_U$ when the EDI and SDI equaled zero. Higher intercept values ($LnSAI_U$) were located in the north and southwest of Austin, whereas lower intercept values ($LnSAI_U$) were in the mid-south and southeast of Austin. Figure 5.8(d) illustrates the spatial variation in the association between SDI and $LnSAI_U$. The range of t values for the coefficients of SDI is -2.814 to 6.704. We masked out the areas where the coefficients with t values between -1.96 and 1.96. The remaining block groups were having a significant association between SDI and $LnSAI_U$. Most of these block groups exhibit significantly positive relationship, which implies that higher SDI tended to associate with higher $LnSAI_U$ in these areas. Only a

few block groups (in blue colors) exhibit significantly negative relationship, representing higher $LnSAI_U$ related to lower SDI.

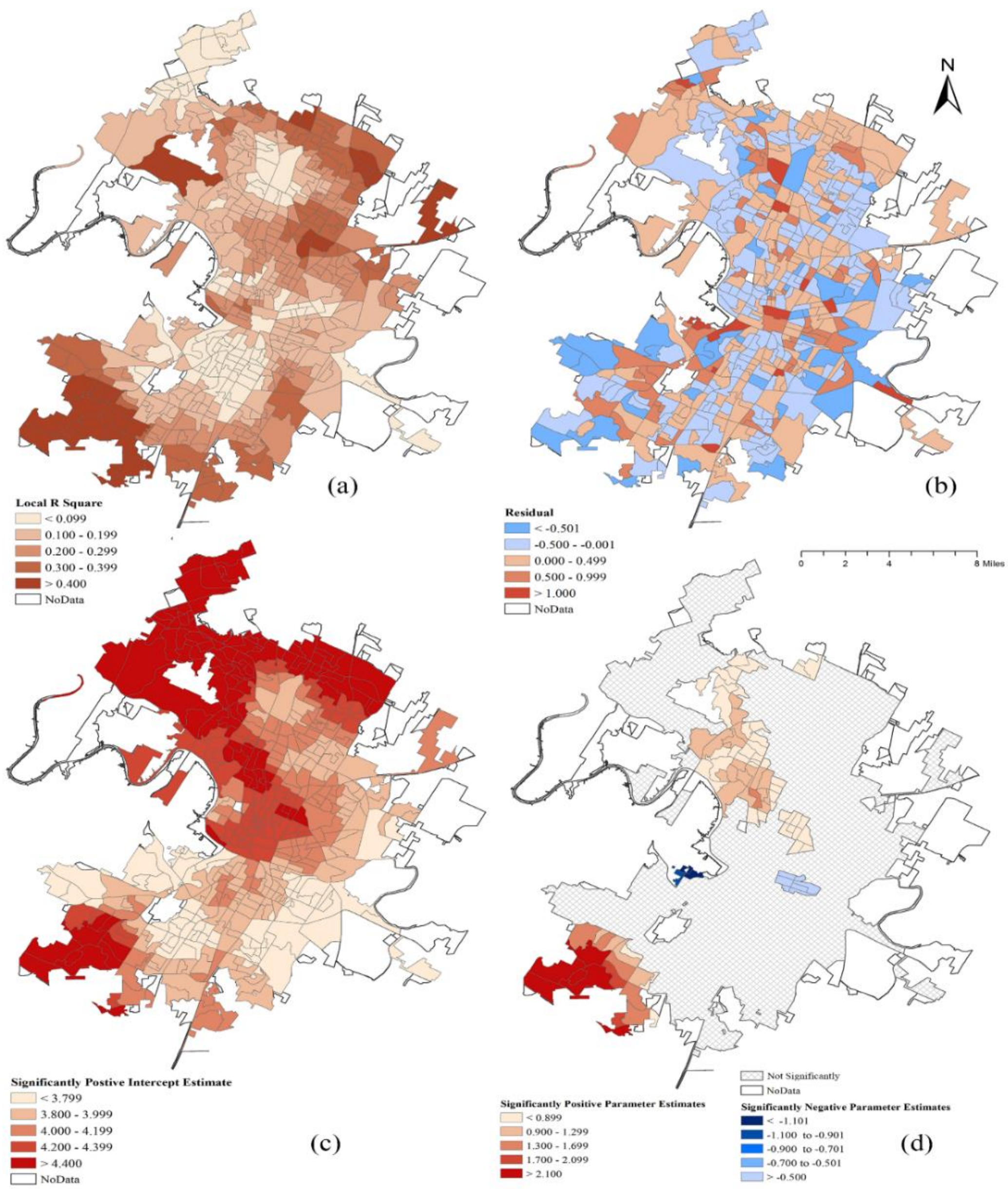


Figure 5.8 Results of S- GWR with $LnSAI_U$ as the dependent variable:(a) local R^2; (b) residual; (c) spatial variation in the coefficients of intercept; (d) spatial variation in the coefficients of SDI.

Delineation of food Deserts and food swamps

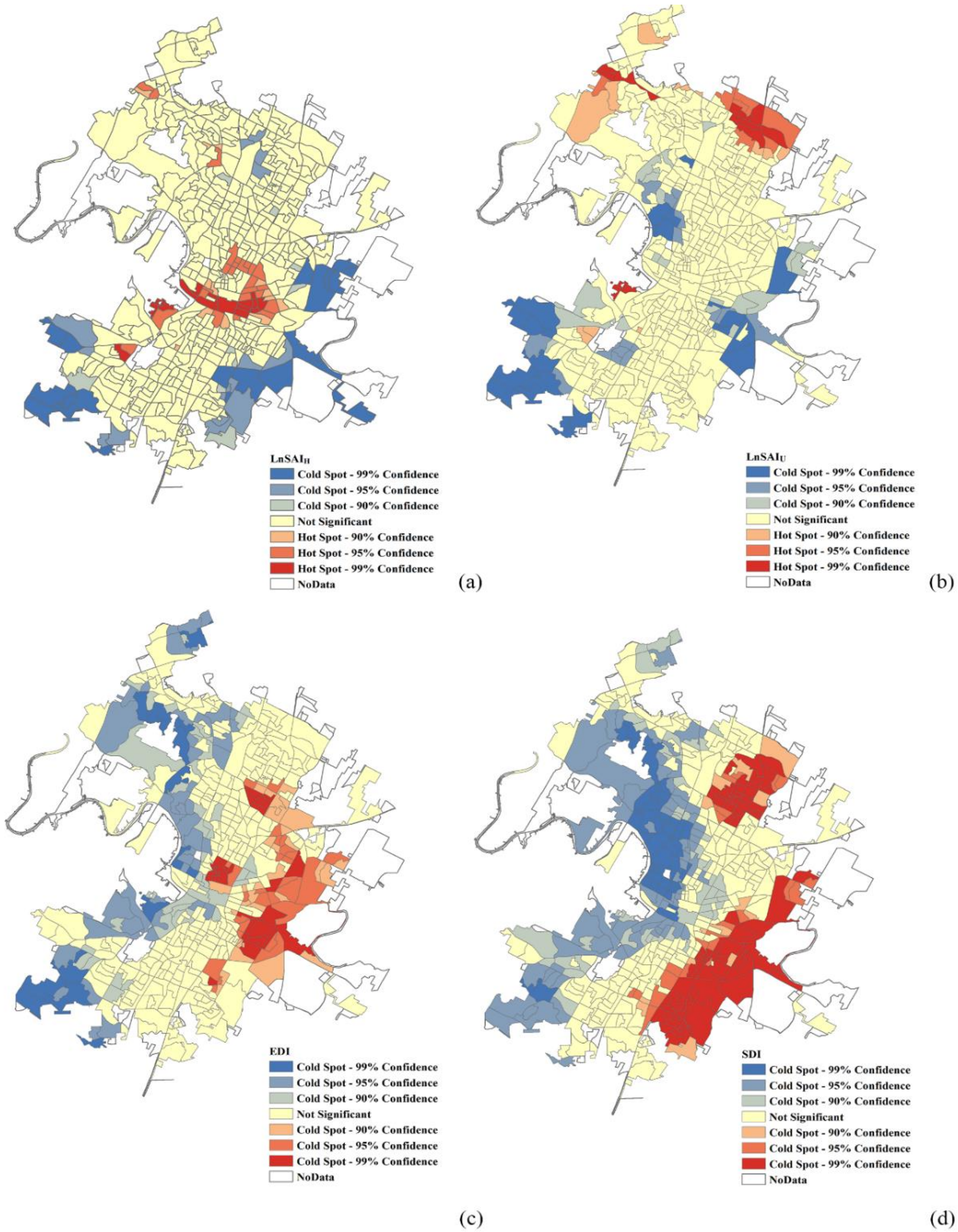


Figure 5.9 Result of hot spot analysis on (a) LnSAI$_H$ (b) LnSAI$_U$ (c) EDI (d) SDI.

Figure 5.9 (a) shows the result of hot spot analysis on the accessibility to healthy food outlets (LnSAI$_H$). It is observed that most of the block groups exhibit insignificant spatial relationships. Hot spots of healthy food accessibility are clustered in the city center and the southwest of Austin, while cold spots are found in the periphery a1nd east of Austin. Three significance levels are used to relax the delineation. I am interested in the block groups classified as cold spots since they indicate that these areas and their adjacent units have high low access to healthy food outlets.

Similarly, I performed the hot spot analysis on the accessibility to unhealthy food outlets ($LnSAI_U$, see Figure 5.9 (b)). The hot spots of unhealthy food accessibility are mostly located in the northeast of Austin. Figure 5.9 (c) shows the hot spot analysis on the economic deprivation index(EDI), and the hot spots are mainly located in the east of Austin. The hot spots of the sociocultural deprivation index (SDI) are in the eastern part of Austin as well (Figure 5.9 (d)).

I intersected the three domains to define food deserts as per the above proposed conceptual definition. In other words, the cold spots of $LnSAI_H$ and the hot spots of EDI and SDI are intercepted to generate food deserts. The result is illustrated in Figure 5.10, and food deserts are mainly located in the east of Austin near Austin International airport (i.e.., the block groups in orange color). As per the Table 5.3, I intersected the block groups with hot spots in Figure 5.9 (a) and cold spots of Figure 5.9 (c) and Figure 5.9 (d) to define potential food oases, which are shown in Figure 5.10 in green color. Food oases are mainly found in the urban center with some isolated ones in the north or northwest of Austin. The same method is applied to delineate food swamps in Austin. The hot spots of $LnSAI_U$ and SDI were intersected to delineate food swamps. The delineated food swamps are mainly located in the northeastern tip of Austin approximating to Walnut Creek Metropolitan Park, which is shown in red color in Figure 5.10.

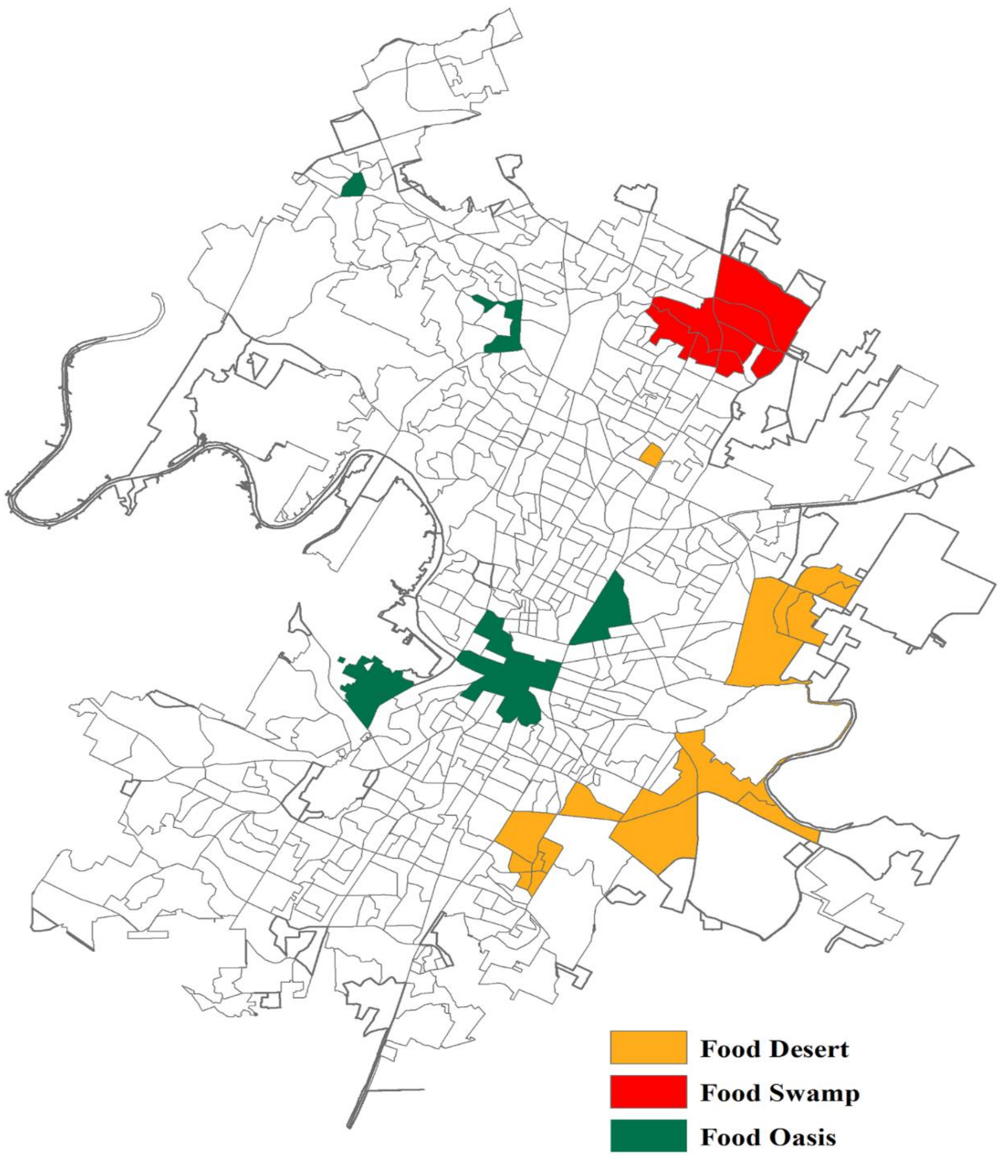


Figure 5.10 Identified food deserts, food oases, and food swamps in Austin, Texas.

Discussion and Conclusion

Demographic variables were scaled down to economic and socio-cultural deprivation indices by factor analysis. The results showed that the East of IH-35 was experiencing both economic and sociocultural deprivation. This pattern might be related to Austin's history and political policy and racial divide. In the late 1880s and the early 1900s, African-Americans and Hispanics spread across the city of Austin. In 1928, the creation of a "Nego District" made the East Austin the only place for African-Americans to access to public schools and other services. The segregation has been further reinforced along the IH-35 since the1940s. A large number of Hispanics/Latinos chose to remain the East of IH-35, while most of the Anglo population moved to other parts (especially in the West of IH-35) of the city. The deepened segregation prevents the people in East IH-35 from economic opportunities (e.g., income and employment) and cultural anchors (i.e., education), which makes this area the most deprived and marginalized in Austin. In one word, the economic and sociocultural berries reflect the population settlement unique to the Austin area.

I examined the association between inequality in spatial accessibility to food outlets and the two socio-demographic indices. This research focused on the use of spatial statistic methods. By using non-spatial (or traditional) statistic such as OLS, we did not find any significant relationship between access to healthy food outlets and the two sociodemographic factors in the city of Austin, Texas. By contrast, the use of a spatial lag model revealed a different relationship: access to healthy food outlets in Austin was significantly associated with economic deprivation but insignificantly associated with sociocultural deprivation. The magnitude of economic deprivation was stronger than the OLS. The relationship is negative (-0.054), indicating that residents with a high economic disadvantage (i.e., low EDI) would have reduced access to healthy food outlets. This finding aligns with many previous results that residents from low-income and high-poverty neighborhood were less accessible to healthy foods (Beaulac, Kristjansson, and

Cummins 2009; Canto, Brown, and Deller 2014; Drewnowski and Specter 2004; Leibtag and Kaufman 2003; Chung and Myers 1999; Morland, et al. 2002; Larsen and Gilliland 2008; Fleischhacker, et al. 2011; Hilmers, Hilmers, and Dave 2012). But it does not support Dai and Wang's (2011) finding that areas with higher urban economic disadvantages had better spatial food accessibility to healthy food stores. The possible reason could be ascribed to that Dai and Wang (2011) differentiated urban and rural economic disadvantages.

When it comes to the relationship between access to unhealthy food outlets, the OLS model depicted that access to unhealthy food outlets was significantly related to both economic and sociocultural deprivation. However, only the sociocultural factor was a significant predictor of unhealthy food access after the spatial error model was used. The magnitude of the sociocultural deprivation was stronger than the OLS as well. The positive relationship (0.160) suggests that neighborhoods in high sociocultural deprivation (i.e., high SDI) had better access to unhealthy foods. The influence of race/ethnicity on unhealthy food access has been widely recognized in America and other countries (Galvez, et al. 2008; Hargreaves, Schlundt, and Buchowski 2002; James 2004; Kumanyika, et al. 2007; Lisabeth, et al. 2010; Morland, et al. 2002; Pearce, et al. 2007); many of them showed that neighborhoods with predominantly minority (i.e., Hispanic or African America) and mixed races had higher convenience store or fast-food restaurants accessibility (Morland, et al. 2002; Pearce, et al. 2007; Galvez, et al. 2008; Lisabeth, et al. 2010). The findings agreed with these studies that Hispanic/Latino neighborhoods (Galvez, et al. 2008; Lisabeth, et al. 2010) and low education neighborhoods (Barker, et al. 2008; Lawrence, et al. 2009) were more exposed to unhealthy foods.

I further examined the spatial heterogeneity of the relationships between food access and neighborhood deprivation with semi-parametric GWR. In comparison with the classic GWR, this method allows specifying fixed and varying variables together in one model, which is likely to improve model fit. Our findings support previous research that demonstrates significant disparities in food access across socioeconomic groups and minority population. It revealed that economic deprivation was a global significantly negative predictor of healthy food access but not for unhealthy food access, meaning that economic deprivation imposed a globally persisting restriction on healthy food accessibility rather than on unhealthy food accessibility. Sociocultural deprivation only exhibited a significantly negative and positive effect on healthy (or unhealthy) food access in certain block groups in East Austin and north Austin. Acknowledging this unique local relationship is critical for regional policy making and resource allocation. It implies that designing programs to improve economic status such as household income and reducing poverty are effective ways to increase the possibility of procuring healthy foods, but they do little or no effect on unhealthy food access. Effective interventions aiming at reducing unhealthy food access could be through modifying sociocultural aspects such as improving education level, promoting Hispanic/Latino residents' healthy eating habit, and overcoming language barriers; especially in the areas with significant relationships.

Furthermore, I identified food deserts and food swamps in Austin, Texas with hot spot analysis. It revealed that food deserts were located in East Austin, and food swamps were found in the Northeast Austin. Both food deserts and food swamps were along the east side of the major highways IH-35. Food desert areas have not only scarce healthy food outlets but also have high economic and sociocultural deprivation. In these areas, the less-dense population makes it hard to make a profit from opening new grocery stores and supermarkets. The alternative is to encourage residents to operate farmers' markets, develop community gardens, and plant fruits and vegetables in backyards. Also, governments should invest more economic opportunities for people living in East Austin and help them improve income and eliminate poverty so that they have the affordability to purchase healthy foods. It is worthy to note that UT Austin campus was not identified as a food desert by our method because there are many grocery stores near it. Students in this area have the advantage of enjoying better spatial access to healthy food stores. However, I questioned whether this spatial advantage could indeed transfer to them. Students often have less mobility because of lower car ownership, and poor food affordability due to low income. They are likely to purchase cheap and unhealthy diets nearby because of the burden of carrying heavy grocery bags via public transit or walking. Therefore, food issues on the UT Austin campus should be concerned by some agencies. For food swamp neighborhoods, food authorities should make zoning law or policies to limit the development of fast-food restaurants and convenience stores. Moreover, the sociocultural disadvantage in food swamp areas may result in inadequate nutrition intake and impose a health disadvantage, especially for Hispanic and linguistically isolated families with low-level education. Therefore, intervention programs that help marginalized groups adopt a healthy lifestyle, as well as enhance their literacy and knowledge on food nutrition are more useful to deal with food swamp issue in these areas.

This research proves that spatial accessibility is not the only factor to limit the healthy food access or promote unhealthy food access; deprived economic and sociocultural conditions can exacerbate the situation. Deprived conditions also can influence people’s dietary habit; disadvantaged people may have no choice but to purchase energy-dense, nutritionally inferior but cheap food, even with sufficient healthy food supply around their neighborhoods (Helling and Sawicki 2003; Larson, Story, and Nelson 2009). Furthermore, this paper aims to address methodological gaps in previous research on the food environment in Austin, Texas. It emphasizes the importance of examining spatial effects in the study areas. Ignoring spatial effects in food environment assessment could lead to biased results. In the present study, the absolute values of coefficient estimates are generally larger in spatial models than the OLS model. The underestimated impacts of some economic and sociocultural deprivation could mislead policy, recommendations, and interventions; so does the identification of food deserts and food swamps. Failing to control for the spatial dependence could result in many small but discontinuous neighborhoods, which are not appropriate for food planning and policy implementation.

Despite the significant findings in this research, there are several limitations, which might be explored in future work. First, I separated two-factor analyses for two indices rather than a group of variables for two factors. The risk of doing this is that there might be a multicollinearity problem in my study; future research should rescale the eight variables into major factors. Second, this research emphasizes the importance of spatial statistics; however, the technique of constructing the composite index (i.e., factor analysis) is still a non-spatial approach. Luan (2016) applied its alternative spatial approach — spatial latent factor analysis to account for the spatial dependence of associated constructs. Future study would be beneficial from this approach. Second, we did not investigate food access for several population groups that potentially are vulnerable to procure foods. These groups include but not limited to the senior population and African American groups. Last, future studies should include more sociodemographic variables, which could guarantee that neighborhood marginalization is fully represented (Luan 2016).

6 EVALUATING THE CONSUMER NUTRITION ENVIRONMENT IN FOOD DESERTS AND FOOD SWAMPS

Introduction

According to Texas Health and Human Services, the prevalence of overweight and obesity for adults in Texas has increased from 15.9% in 1995 to 33.87% in 2010, which had placed Texas 8th nationally. It is estimated that the share of overweight and obese people will be 75% by 2040 (Gloria and Steinhardt 2010). People in Austin, Texas also suffer from a high prevalence of obesity epidemic. In 2011, 37.1 % of residents were overweight, and 27.0 % of residents were obese. The percentages have risen by 21% from 2011 to 2016[21]. The consequences of overweight and obesity are remarkable. On the one hand, it could result in a number of health risks and diseases, such as hypertension, cardiovascular disease, type II diabetes, stroke and some cancers (Brown, Donato, and Obarzanek 1998; Haffner, et al. 1991; Bostick, et al. 1994; Chute, et al. 1991). On the other hand, a considerable economic cost on medical care relates to obesity and overweight; in 2018 the total health care on obesity-related diseases would be $ 344 billion.

Researchers and epidemiologists are passionate about uncovering determination to reduce overweight and obesity. Early studies focused on individual behaviors to improve individuals' physical activity and diet (Coulter 2009; Glanz, et al. 2005; Luan 2016). Unfortunately, it proved that the effect of modifying personal behaviors is limited (Coulter 2009). Research has begun to shift their attention to the effect of environment on

[21] https://data.austintexas.gov/stories/s/Healthy-Austin/78uy-qt4w/

obesity intervention. It was found that diet has more effects determining weight outcomes than physical activity (Franco, et al. 2008; Galvez, et al. 2008). Therefore, in the past decade researchers have extensively explored the effect of the food environment on dietary behaviors and weight status (Lopez 2007; Mobley, et al. 2006; Macdonald, et al. 2011; Powell, Chaloupka, and Bao 2007; Wang, et al. 2007). A growing body of nutrition environment frameworks has been proposed during this period. Glanz and colleagues' nutrition environment framework is the most renowned one. It suggests that the nutrition environment should be dived into two sets of factors: community and consumer nutrition environments. The community consumer environment focuses on effect of number and location of food outlets on health outcomes, which has been explored by many scholars and researchers (Cooksey-Stowers, Schwartz, and Brownell 2017; Lopez 2007; Mobley, et al. 2006; Macdonald, et al. 2011; Wang, et al. 2007). The consumer nutrition environment refers to consumers' experience in food stores, including food availability, food affordability, food quality, and other in-store characteristics. The consumer nutrition environment has been less explored compared with the community nutrition environment (Glanz, et al. 2005).

Food availability measures the presence or absence of food items sold in the stores (Donkin, et al. 1999). It is found that greater availability of healthy foods (e.g., fruits and vegetables) is positively related to the consumption of these healthy items (Farah and Buzby 2005). Food affordability mainly refers to food cost relative to an individual's or household's income. Food cost was identified as an essential barrier to healthy eating in low-income communities (Chung and Myers 1999) and households (Cassady, Jetter, and Culp 2007). Food quality measures the quality characteristics of foods in food retailers.

Poor food quality (e.g., withered or bruised fresh produce, rotting meat, and expired canned foods) could deter food purchasing behaviors (Brown 2014) and therein impose an adverse effect on diet quality and health outcomes. Besides, in-store characteristics, such as food labeling, could provide detailed information on a food's nutrient content. Reading food labels is especially important when consumers have certain health conditions such as high blood pressure. It facilitates consumers to compare labels of different food items to make the best decision. It is found that food labels positively predicted dietary quality and dietary behaviors (Cooke and Papadaki 2014; Satia, Galanko, and Neuhouser 2005). It is also reported that food labeling could project a decrease in long-term body weight (Bollinger, Leslie, and Sorensen 2011). However, in comparison with the other three domains, food labeling has been less explored in consumer nutrition environment assessment. Hence, food labeling was included as one of the critical in-store characteristics in this study.

Studies of the consumer nutrition environment have often concerned whether people from deprived neighborhoods and minority-dominated communities experienced lower availability of healthy foods, higher food price, and lower food quality for the different type of stores (Coulter 2009; Glanz, et al. 2005; Gloria and Steinhardt 2010; Woodham 2011). These studies have been motivated by the deprivation amplification hypothesis (Macintyre 2007). Therefore, most of the studies have explored the consumer nutrition environment in different store types for two contrasting neighborhoods, and they are either high-income vs. low-income or white-dominated vs. African American (or Hispanic) dominated neighborhoods. As the food insecurity has been emerging as an impelling issue in the U.S and other countries, however, to date no studies have explored the consumer nutrition environment in food insecure (food desert and food swamp) vs. food secured (i.e., food oasis) neighborhoods. In addition, most studies compared the differences of consumer nutrition environment by store types and income (or minorities status) using simple comparison methods without exploring their interaction effect, such as t-test (e.g., (Coulter 2009)) and one-way ANOVA (e.g., (Glanz, et al. 2005)). Some research (i.e., (Gloria and Steinhardt 2010)) used complex methods such as two-way ANOVA, but their exploration of interaction effect was incomplete.

High-quality food auditing instruments are critically important to evaluate the consumer nutrition environment (Coulter 2009). A common way to measure consumer nutrition environment in early studies was to use a standard food basket. A series of standard food baskets were developed to serve different research agenda (Block and Kouba 2006; Anderson, et al. 2011; Ling 2005; Cassady, Jetter, and Culp 2007; Harrison, et al. 2007; Palermo, et al. 2008; Bovell-Benjamin, et al. 2009). However, these food basket instruments are absolute measures, which focus on specific food items (e.g., healthy foods) in a store. They do not allow to compare food price between healthy and regular options.

Moreover, most of them did not report the reliability of their instrument (Cassady, Jetter, and Culp 2007; Harrison, et al. 2007; Palermo, et al. 2008; Bovell-Benjamin, et al. 2009). NEMS (Nutrition Environment Measures Survey) has been developed as an observational tool to assess consumer nutrition environments (Glanz, et al. 2007). It consists of a serial of different instruments including NEMS-S (for stores) and NEMS-R (for restaurants), NEMS-CS (for corner stores), NEMS-V (for vending machines), and NEMS-P (for perceived nutrition environment). The NEMS is a relative tool, which can solve the problem of food basket instruments. Although NEMS has an excellent reputation for its high reliability, it might not be appropriate to other states or cities since it did not include some culturally appreciated foods that are important to residents. Take Texas as an example: 35 % of the population is Hispanic/Latino. An ethnic group such as Hispanic/Latino people have their own cultural identities. As a result, NEMS needs to be tailored to make it more adaptive to local communities. Physical Activity, and Obesity Prevention (NPAOP) developed such a tool and named it as "Texas Nutrition Environment Assessment in Stores (TxNEA-S)"(Gloria and Steinhardt 2010). It is an adaptation of the NEMS tool that included additional foods that are culture-important to Hispanics and other minorities in Texas.

Although TxNEA-S is useful and meaningful in Texas, it needs to be tailored for specific projects or research. One good example is TxNEMS-WIC[22]Store audit tool. Its objective is to measure the impact of WIC package on the availability of WIC items to WIC participants. The foods include fresh, frozen, and canned fruits and vegetables, bread, cereals, and milk. TxNEA needs to be customized due to several reasons. First, TxNEA does not include beverage such as coke, which contains high calories to make people more obese potentially. Second, it does not contain meat such as ground beef and chicken, which are essential to obtain protein. Third, TxNEA-S is excessively lengthy (i.e., 24 pages) and required a lot of time to complete. It would take about two and a half hours to finish one grocery store. Therefore, it is necessary to erase some less-important

[22] http://www.dshs.texas.gov/Obesity/TXNEAS/

food categories such as convenience-added produce and bulk dry grains, which have not often used in obesity-related environment assessment.

There are two purposes of this research: (1) developing an appropriate tool to measure consumer nutrition environment in grocery (or supermarket) and convenience stores in central Texas; (2) conducting a survey in Austin, Texas to examine whether there is a difference in the consumer nutrition environment from grocery and convenience stores between food desert, food swamp, and food oasis neighborhoods. I hypothesized that grocery stores and supermarkets would have greater availability of healthy foods, lower price of healthy options, higher food quality, and a higher percentage of food that are labeled than convenience stores. Moreover, stores from food desert and food swamp neighborhoods would have lower availability of healthy foods, a higher price of healthy options, lower food quality, and lower portions of labeled foods than those from food oasis neighborhoods.

Method

Design of M-TxNES-S auditing instrument

M-TxNES-S food auditing tool was designed by customizing the TxNES-S. The primary food categories of the M-TxNES-S instrument include fresh fruit and vegetables. Meanwhile, canned and frozen fruits and vegetables are also added because low-income people tend to purchase non-refresh alternatives. Milk and cheese, grains, meat and alternatives, beverages, and snacks (i.e., chips and pretzels) are included as well. The specific food survey instrument can be seen in Appendix B. The tool consists of 10 pages and 93 food items. The survey starting and ending time, the number of cash registers, WIC and food stamps certificates, and types of food stores were recorded on the cover page. Each food item in the list was examined by availability, price, and labels. In addition, quality was examined only for fresh fruits and vegetables. Since the M-TxNES-S survey instrument was adopted from TxNES-S, which was taken from NEMS-S, we designed the survey instrument guideline following the general protocols of TxNES-S and NEMS-S. One can refer to Appendix C for more information.

Texas State University institutional review board (IRB) have approved this survey (Appendix D). Upon entering the stores, the raters introduced themselves to store managers or cashiers and asked their permission to conduct the survey. The verbal consent form is attached in AppendixE. The raters need to explain the study purpose, the anonymity, and voluntariness of participation. They also need to present a letter (see Appendix F) to the manager/owner with the principal investigator's contact information in case of further questions.

Selection of neighborhoods and food stores

Food survey audit was conducted in three neighborhoods located in the city of Austin. Two criteria were used to select block groups in each neighborhood: 1) they were classified as food deserts, food swamps, and food oases in Chapter five; 2) they must be continuous in geographic space[23]. The distribution of the selected neighborhoods is illustrated in Figure 6.1. The food desert neighborhood was on the eastern side of IH-35 and close to Austin international airport. The food swamp neighborhood was across the

---

[23] Note: there is no standard definition of what constitutes an individuals' food environment and therefore no standard boundaries for neighborhoods. However, a neighborhood should be continuous and no gaps in it.

IH-35 and in the northeastern corn of Austin; whereas the food oasis neighborhood was in the western section of IH-35 and near Austin downtown.

Table 6.1 shows the population, density, and socio-demographic information of the three neighborhoods. The three neighborhoods had a comparable population, but the land sizes varied. The size of the food desert neighborhood was twice as large as the one in food oasis, leading to that population density in food oasis was highest (i.e., 4437.834 people/ sq. mi). It is evident that residents' economic and sociocultural statuses were low in food desert neighborhood but high in food oasis neighborhood (Table 6.1).

Table 6.1 General characteristics in the three neighborhoods in Austin, Texas.

	Food Desert		**Food Swamp**		**Food Oasis**	
# of Block Groups	13		10		12	
Total population	18173		17439		16945	
Size (square miles)	8.849		6.479		3.818	
Density (# people/ sq. mi)	2053.571		2691.683		4437.834	
Economic Characteristics	Mean	SD	Mean	SD	Mean	SD
Household income ($1,000)	33.531	8.314	49.400	6.784	75.694	16.113
Below Poverty line (%)	31.607	9.842	8.280	3.524	6.412	4.712
Unemployment rate (%)	32.119	7.380	21.200	4.241	16.467	4.144
Lack of ketch facility (%)	2.143	3.406	1.170	2.487	2.519	3.853
Sociocultural Characteristics	Mean	SD	Mean	SD	Mean	SD
Without Higher Education (%)	89.106	9.429	61.500	12.383	18.745	4.648
Hispanic people (%)	79.324	13.972	35.800	15.095	9.203	3.799
Home rental (%)	58.282	19.306	61.300	24.826	61.108	13.043
Language isolation (%)	0.148	0.086	0.060	0.038	0.02	0.038

I employed two methods to identify retail food stores in the three selected neighborhoods. One method is through the online directory databased Reference USA to look for all food stores in the neighborhoods, and there were 42 food stores identified. Then I combined with Google Maps to search keywords including grocery store, supermarket, supercenter, convenience store, mini-mart, food store, retail food store, food mart, corner store, mom and pop store, and bodega. It returned 50 food stores using Google Maps. The other method is to conduct fieldwork to verify the locations of the stores.

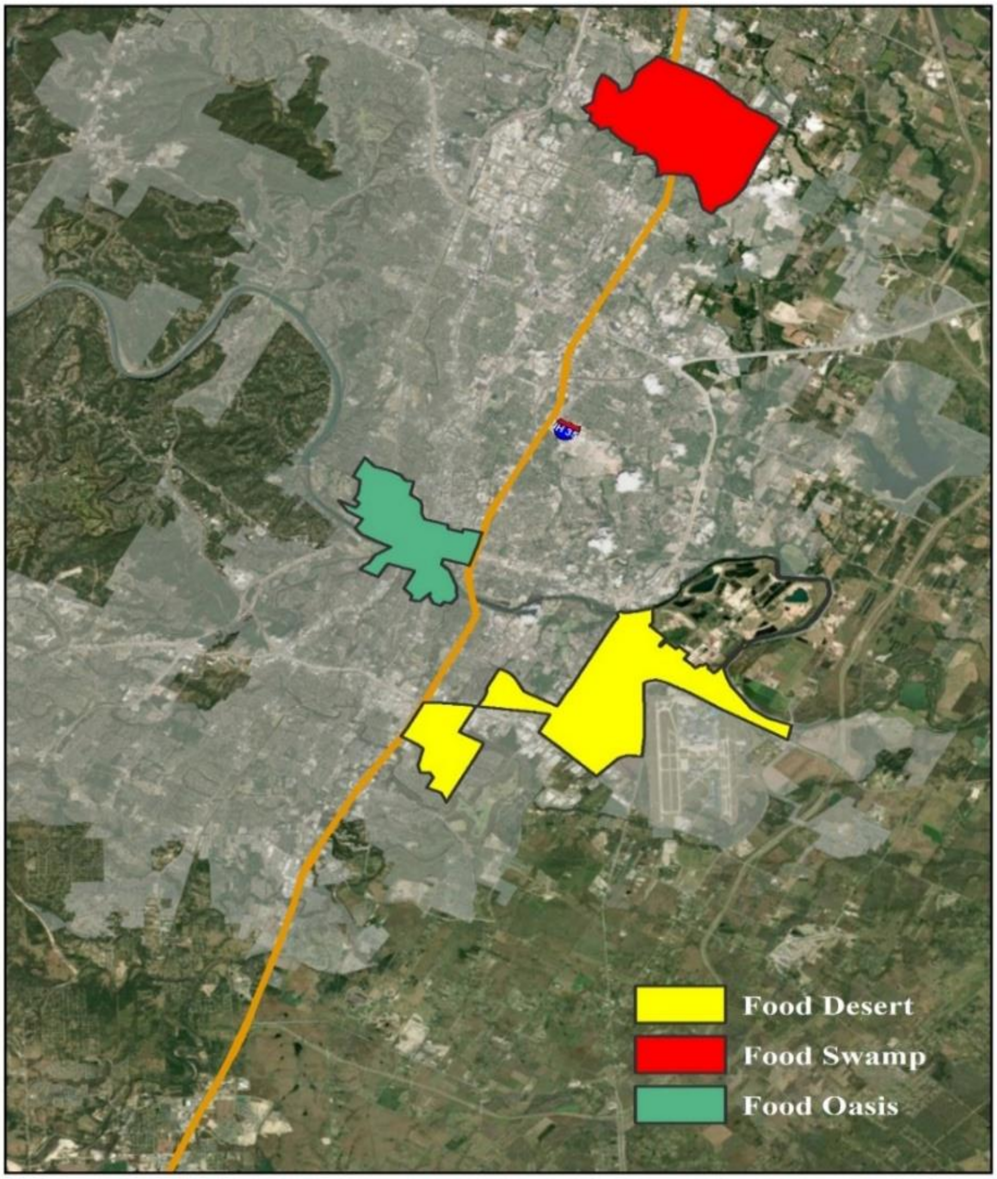


Figure 6.1 The three selected neighborhoods for food survey in Austin, Texas.

Table 6.2 The number of food stores were identified in the three selected neighborhoods.

	Grocery stores & supermarkets	Convenience stores	Total
	Reference USA & Google Map: 3 stores;	Reference USA & Google Map: 9 stores	12
	Filed trip: 1 additional grocery store, 2 grocery stores, and 1 convenience stores	Field trip: 9 convenience stores	
Food Deserts	Identify: **3** grocery stores & supermarkets	Identify: **10** convenience stores	**13**
	Reference USA & Google map: 5 stores;	Reference USA & Google map: 12 stores	17
	Field trip: 5 stores (4 grocery stores & supermarkets and 1 convenience store)	Field trip: 12 convenience stores	
Food Swamps	Identify: **4** grocery stores & supermarkets	Identify: **13** convenience stores	**17**
	Reference USA & Google map: 12 stores	Reference USA & Google Maps:10 stores	22
	Field trip: 8 stores (7 grocery stores & supermarkets + 1 convenience stores)	Field trip: 8 convenience stores + 1 cloth shop + 1 went out of business	
Food oases	Identify: **7** grocery stores & supermarkets	Identify: **9** convenience stores	**16**

It can be observed that the store identification between Reference USA & Google Maps and field trip almost corresponds to each other in both food desert and food swamp neighborhoods (Table 6.2). However, there are relatively large discrepancies for grocery stores & supermarkets in food oasis neighborhood. For instance, two grocery stores were moved out of the neighborhood; one was remodeled as a BBQ house; one was a convenience store and had been shut down, and one was under construction. In addition, some grocery stores & supermarkets identified by Reference USA & Google Maps turned out convenience stores via the field trip (Table 6.2); we thus classified them into convenience stores for a more accurate assessment of food environment. As a result, a total of 46 food stores were selected to be surveyed within the three neighborhoods (see Table 6.2). Among these 46 stores, 13 ones were in the food desert, 17 ones were in food swamp, and 16 ones were in food oasis, respectively. As suggested by Glanz et al. (2007), the sample size within each neighborhood should meet the minimum size of at least 15 stores. Note that in our research the number of stores surveyed in food desert neighborhood does not meet this criterion (n =13). However, 13 is the maximum number of the existing food stores found in that neighborhood. In essence, it makes sense because the areas classified as food deserts are inherently short of food stores.

Rater recruitment and training

To assure inter-rater reliability of the survey two raters were needed to conduct the survey simultaneously. If the inter-rater reliability for some stores is below the threshold (i.e., 80%), one of the raters conducted a second-time survey in these stores for the verification purpose. The raters are not necessarily from the nutrition program, and they are not required to have prior knowledge of food nutrition. A flyer has been designed to recruit raters (See Appendix G). Approximately 20 copies of the flyers were put up in Department of Geography, Alek Library, and LBJ student center on Texas State University campus. Fortunately, I was able to recruit one rater to assist me in surveying the selected three neighborhoods. I then gave the rater a professional training about NEMS-S and TxNES-S. The training materials were derived from NEMS online training[24] and Department of State Health Services. Then the rater and I conducted a pilot assignment in a grocery store in San Marcos after the training. Once our inter-rater reliability achieved 80% and above, this training was complete.

The two raters conducted during the week of Aug. 17 - Aug. 24, 2018. The two raters entered each store at the same time on the same day but conducted the survey independently by checking up food items in different orders. The prices for some food items in certain ethnic grocery stores and convenience stores were not available, and the raters asked cashiers for prices. Once both raters have finished all pages in one store, they scrutinized the survey instrument to make sure that the survey was100% completed.

Data analysis

All data analysis was conducted in SPSS 25.0. The percentages of agreement for food availability, price, quality, and labeling were calculated, respectively, as suggested by Glanz, et al. (2007). It is an Item-by-item agreement between the two raters was compared, and percentage agreement – the frequency of correctly matched responses divided by the total number of observations. Then, the availability of healthy foods score was calculated using the percentage values (i.e., the number of healthy food observations is divided by 93 food items and then multiply 100). I used Two-way ANOVA to examine whether the availability of healthy foods is different between store type (grocery stores & supermarkets and convenience stores) and neighborhood food environment (food desert, food swamp, and food oasis).

[24] http://www.med.upenn.edu/nems/onlinetraining.shtml

Food items in grocery stores usually are sold by the pound, but in convenience stores, some foods (i.e., bananas) are sold by the piece. To make their prices comparable, we converted the price by per piece to by per pound based on Appendix H. Price comparisons was calculated as a percentage, based on the average price for a healthy item compared to its regular counterpart. For quality analysis, we calculated the percentage of F&V in acceptable quality (i.e., the number of F&V in acceptable quality is divided by all observable F&V in a store and then multiply 100). The histogram of quality data was severely left-skewed and was not appropriate to perform Two-way ANOVA on it. Instead, we applied nonparametric one-way ANOVA (Mann-Whitney U test, and Kruskal-Wallis H test) to the quality data to examine whether ST and NNE have effects on it. Two-way ANOVA analysis was also used to compare food labeling in the three neighborhoods among two different types of stores.

Results

The total number of food stores we visited was 46. However, five stores (one Hispanic grocery store and four convenience stores) refused to participate in the survey, resulting in 13 grocery stores & supermarkets, and 28 convenience stores for analysis (Table 6.3). The cross-tabulation of all grocery and convenience stores within the neighborhood samples revealed that the food swamp neighborhood had 150% as many convenience stores as those in food oasis neighborhood, and food desert neighborhood had only less than 30% as many grocery stores & supermarkets, as compared with food oasis neighborhood. The completion rates for grocery stores & supermarkets and convenience stores were 92.89 % an 87.5%, respectively. 62% (11 out of 13) of the grocery stores & supermarkets had three or more cash registers, and small grocery stores (with 1 or 2 cash registers) were more common in food oasis neighborhood whereas 95% of convenience stores had no more than 3 cash registers. Approximately 70% of food stores accepted food stamps, but only a few stores (i.e., 30%) accepted WIC. The mean time to complete the measures were 35.077 (SD 29.029) minutes for grocery stores & supermarkets and 6.839 (SD 2.449) minutes for convenience stores.

Table 6.3 Summary of cover-page information for the food survey instrument.

ST	**NNE**	**N**	**complete rate (%)**	**>2 cash register**	**food stamp**	**WIC**	**completion time (min)**
Grocery Store & Supermarket	Food Desert	2		1	2	0	11.250 (SD:1.768)
	Food Swamp	4		4	3	2	28 (SD: 22.275)
	Food Oasis	7		6	4	3	45.929 (SD: 32.814)
	Total	13	92.89%	84.62%	69.23%	38.46 %	35.077 (SD: 29.029)
Convenience Store	Food Desert	8		2	8	2	7.313 (SD: 3.494)
	Food Swamp	12		2	12	3	7.125 (SD: 2.090)
	Food Oasis	8		2	3	3	5.938 (SD:1.657)
	Total	28	87.50%	21.43%	71.88%	25.00 %	6.839 (SD: 2.449)

Inter-rater reliability

Inter-rater reliability of food availability, price, quality, and labeling is shown in Table 6.4. The agreement rates on the availability measure were consistently high, ranging from 91.70% to 100%. Food price had a moderate agreement rate, with a mean 86.08 % and stand deviation of 5.78%. Food quality and food labeling both had a high agreement rate.

Table 6.4 Inter-rater reliability for the M-TxNEA-S food instrument.

Store ID	Store type	Neighborhood	Inter-rater reliability			
			availability	price	quality	label
001-001-01	C	FD	0.952	0.857	1.000	0.905
001-002-01	G&S	FD	0.963	0.852	1.000	0.963
001-002-02	G&S	FD	0.920	0.920	0.920	0.960
001-004-02	C	FD	1.000	1.000	1.000	1.000
001-004-03	C	FD	1.000	0.895	1.000	0.947
001-004-04	C	FD	0.933	0.867	1.000	0.933
001-004-05	C	FD	0.933	0.867	1.000	0.867
001-004-06	C	FD	1.000	0.808	1.000	0.923
001-004-07	C	FD	1.000	0.929	1.000	0.929
001-004-09	C	FD	0.950	0.900	1.000	0.950
002-001-01	G&S	FS	0.978	0.870	1.000	0.935
002-001-02	G&S	FS	0.950	0.850	1.000	0.850
002-001-03	G&S	FS	1.000	0.886	0.950	0.977
002-001-04	C	FS	0.917	0.750	0.833	0.917
002-003-01	G&S	FS	0.989	0.860	1.000	0.957
002-004-01	C	FS	1.000	0.889	1.000	0.944
002-004-02	C	FS	1.000	0.889	NA	1.000
002-004-03	C	FS	1.000	0.846	1.000	0.923
002-004-05	C	FS	1.000	0.846	1.000	0.923
002-004-06	C	FS	1.000	0.857	1.000	0.857
002-004-07	C	FS	1.000	0.889	NA	1.000
002-004-08	C	FS	0.923	0.846	NA	0.923
002-004-09	C	FS	1.000	0.875	NA	1.000
002-004-10	C	FS	1.000	0.824	1.000	0.824
002-004-11	C	FS	0.933	0.800	NA	0.933
002-004-12	C	FS	0.923	0.846	NA	0.923
003-001-02	G&S	FO	0.987	0.803	1.000	0.868
003-001-04	C	FO	1.000	0.867	NA	0.933
003-001-05	G&S	FO	1.000	0.905	1.000	0.952
003-001-06	G&S	FO	1.000	0.839	1.000	0.968
003-001-07	G&S	FO	0.971	0.794	1.000	0.941
003-001-08	G&S	FO	0.987	0.787	1.000	0.960
003-001-09	G&S	FO	0.988	0.800	1.000	0.976
003-001-11	G&S	FO	1.000	0.798	1.000	0.989

Table 6.4-countinued

003-004-02	C	FO	0.933	0.867	1.000	0.800
003-004-03	C	FO	1.000	0.875	NA	1.000
003-004-04	C	FO	0.944	0.889	NA	0.833
003-004-05	C	FO	1.000	0.750	NA	1.000
003-004-07	C	FO	1.000	1.000	NA	1.000
003-004-08	C	FO	1.000	1.000	NA	1.000
003-004-10	C	FO	1.000	0.800	NA	1.000

Note: G&S: Grocery Stores & Supermarkets; C: Convenience Stores; FD: Food Desert; FS: Food Swamp; FO: Food Oasis; NA: Not Applicable due to few (i.e., <= 3) fresh fruits and vegetables in that store.

Healthy food availability

There are 93 food items in the M-TxNEA-S instrument. The mean number of healthy food items across the 41 food stores is 17.15 with standard deviation 21.938 (Table 6.5). The mean percentage of healthy food availability is 6.452 (SD = 18.817%). The histogram (Figure 6.2) shows that the distribution of healthy food availability (%) is right skewed, and most of the stores have a lower percentage of healthy food availability. These stores are mainly convenience stores.

Table 6.5 Descriptive statics on the availability of healthy food items.

	Min	**1st Quartile**	**Median**	**3rd Quartile**	**Max**	**Mean**	**SD**
Number of healthy food items	1.000	4.500	6.000	17.520	69.000	17.150	21.938
% availability of healthy foods	1.075	4.838	6.452	18.817	74.193	18.436	23.589

The two-way ANOVA analysis was conducted to examine the effect of store type (ST) and neighborhood nutrition environment (NNE) on healthy food availability. Residual analysis was performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot; normality was assessed using Shapiro-Wilk's normality test for each cell of the design and homogeneity of variances was assessed by Levene's test. There were two outliers, residuals were not normally distributed (p = 0.038), and the homogeneity of variances was rejected (p = 0.035). In spite of the violations of two-way ANOVA assumption, we still used the two-way ANOVA because the p values were so close to the critical value 0.05.

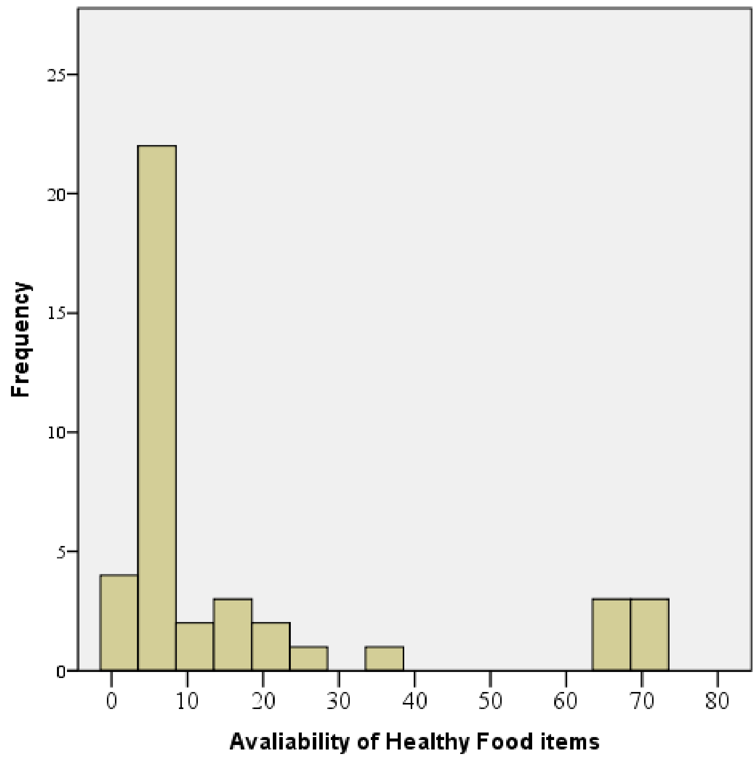


Figure 6.2 Histogram of % availability of healthy foods.

There was a statistically significant interaction between the effects of ST and NNE on the availability of healthy foods, F (2, 35) = 4.462, p = 0.019, partial $\eta^2 = 0.203$. The interaction indicates that the mean differences between healthy foods availability in different nutrition neighborhoods is dependent on store types (ST). It seems that there were significant main effects of ST (i.e., F (1, 35) = 42.915, p = 0.000, partial $\eta^2 = 0.551$) and NNE (i.e., F (2, 35) = 3.473, p =0.042, partial $\eta^2 = 0.166$) on the healthy food availability (Table 6.6). However, according to this professional website[25] , reporting the main effects can be misleading when there was a statistically significant interaction. Therefore, I conducted a separate simple main effect analysis on ST and NNE.

Table 6.6 Results of two-way ANOVA on healthy foods availability.

Source	**type III Sum of squares**	**df**	**mean square**	**F**	**p-value**	**partial η^2**
Corrected Model	15645.05	5	3129.01	16.560	0.000**	0.703
Intercept	15167.09	1	15167.1	80.271	0.000**	0.696
ST	8110.512	1	8110.51	42.925	0.000**	0.551
NNE	1312.372	2	656.186	3.473	0.042*	0.166
ST * NNE	1686.316	2	843.158	4.462	0.019*	0.203
Error	6613.158	35	188.947			
Total	36194.94	41				
Corrected Total	22258.21	40				

R Square = 0.760 (Adjusted R Square =0 .660)

Note: * significant at α = 0.05 level; ** significant at α = 0.01 level.

An analysis of simple main effect for store type (ST) at each neighborhood was performed with statistical significance receiving a Bonferroni adjustment and being accepted at the p < 0.025 level[26]. All pairwise comparisons were run for each simple main effect with reported 97.50 % confidence intervals and p-values Bonferroni-adjusted within each simple main effect. As shown in Table 6.7, in food desert neighborhood there was no significant difference in healthy food accessibility in the two types of stores (F

[25] https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics.php

[26]. This analysis involves the testing of multiple simple main effects. A common method is to apply a Bonferroni adjustment to the declared statistical significance (i.e., p= 0.05). We can do this by dividing the current level (p = 0.05) by the number of simple main effects (in this case, 2). Thus, we would only declare a simple main effect as statistically significant if p < 0 .05 ÷ 2 = 0.025).

(1,35) = 0.707, p = 0.406, partial η2 = 0.020). A statistically significant difference was observed in the mean percentage of healthy food availability between grocery stores & supermarkets and convenience stores in food swamp neighborhood (F (1, 35) = 34.075, p = 0.000, partial η2 = 0.493) (Table 6.7). In this neighborhood, the mean percentage of healthy food availability for grocery stores & supermarkets was 51.613 ± 6.873 and 5.287 ± 3.968 for convenience stores (Table 6.8), a statistically significant mean difference of 46.326 (97.5 % CI, 27.740 to 64.912). In food oasis neighborhood, I also observed a statistically significant difference in mean percentage of healthy food availability between grocery stores & supermarkets and convenience stores (F (1, 35) = 38.704, p = 0.000, partial η2 = 0.525) (Table 6.7); for grocery stores & supermarkets the healthy food availability was 49.770% ± 5.195 and for convenience store with 5.511% ± 4.860, a statistically significant mean difference of 44.259 (97.5 % CI, 27.598 to 60.920) (Table 6.8).

Table 6.7 Simple main effects of store type (ST) within each NNE.

NNE		**Sum of Squares**	**df**	**Mean Square**	**F**	**p**	**Partial η2**
Food Desert	Contrast	133.657	1	133.657	0.707	0.406	0.020
	Error	6613.158	35	188.947			
Food Swamp	Contrast	6438.341	1	6438.341	34.075	0.000**	0.493
	Error	6613.158	35	188.947			
Food Oasis	Contrast	7313.019	1	7313.019	38.704	0.000**	0.525
	Error	6613.158	35	188.947			

Table 6.8 Multiple comparisons between different pairs by ST in each NNE.

NNE	(I) ST	(J) ST	Mean Difference (I)-(J)	SE	p value	97.5% CI for difference [a]
Food Desert	Grocery Store Mean: 16.667 SE: 9.720	Convenience Store Mean: 7.527 SE: 4.860	9.140	10.867	0.406	(-16.310, 34.590)
Food Swamp	Grocery Store Mean: 51.613 SD: 6.873	Convenience Store Mean: 5.287 SE: 3.968	46.326	7.936	0.000 **	(27.740, 64.912)
Food Oasis	Grocery Store Mean: 5.511 SE: 5.195	Convenience Store Mean: 5.511 SE: 4.860	44.259	7.114	0.000 **	(27.598, 60.920)

Note: ** the mean difference is significant at the 0.025 level; a: adjustment for multiple comparisons: Bonferroni.

Figure 6.3 is a clustered bar chart to illustrate the simple main effect of store type in different neighborhoods. Error bars in the figure show the upper and lower 97.5% confidence intervals that extend above and below the mean column. Moreover, the error bar indicated whether several groups of values are statistically different along with the upper and lower bounds of the mean. That is, if there's no overlap in confidence intervals, the differences are statistically significant at the level of confidence (in most cases). It was evident that grocery and convenience stores had substantial differences in food swamp and food oasis neighborhoods (i.e., non-overlapping CI) in terms of healthy food availability, while the difference was trivial in food desert neighborhood and the two value bars had overlapping CI. The graph confirms the appropriateness of analyzing the simple main effect of store type (ST) in this research.

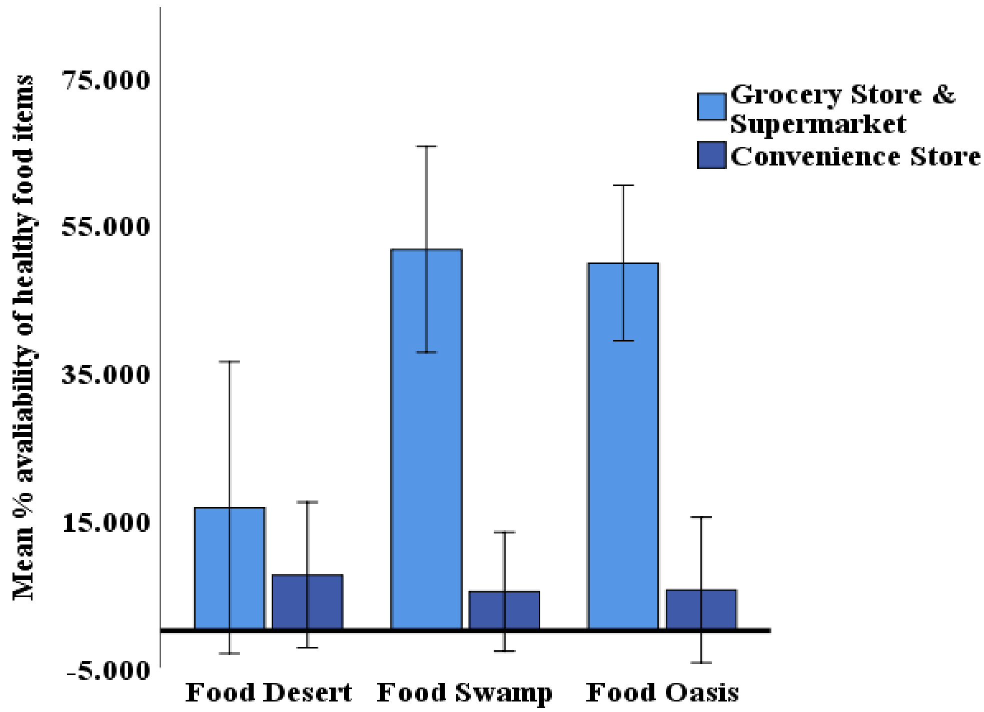


Figure 6.3 Difference in the mean percentage of healthy food availability by ST and NNE.

Table 6.9 The simple effect of NNE within each store type.

ST		Sum of Squares	df	Mean Square	F	p value	Partial η2
Grocery Store	Contrast	1938.946	2	969.473	5.131	0.011*	0.227
	Error	6613.158	35	188.947			
Convenience Store	Contrast	26.668	2	13.334	0.071	0.932	0.004
	Error	6613.158	35	188.947			

An analysis of the simple main effect for NNE at each store type was also performed with statistical significance with a Bonferroni adjustment ($p < 0.025$) level. A statistically significant difference was found in mean percentage of healthy food availability between food desert, food swamp, and food oasis for grocery stores & supermarkets ($F(2, 35) = 5.131$, $p = 0.011$, partial $\eta 2 = 0.227$) (Table 6.9), but there was no significant difference for convenience stores across the three neighborhoods ($F(2, 35) = 0.071$, $p = 0.932$, partial $\eta 2 = 0.004$) (Table 6.9). Mean percentages of healthy food availability for grocery stores & supermarkets in the food desert, food swamp, and food oasis neighborhoods were 16.667 ± 9.720, 51.613 ± 6.873, and 49.770 ± 5.195 (Table 6.10), respectively. Grocery stores & supermarkets in food desert had a statistically significantly lower mean percentage of healthy food availability than those in food swamp, 34.946 (97.5% CI, -68.238, -1.655), $p = 0.018$ (Table 6.10). Grocery stores & supermarkets in food oasis also had a statistically significantly higher mean percentage of healthy food availability than those in the food desert, 33.103 (97.5% CI, 2.281 to 63.925), $p = 0.015$, (Table 6.10). Figure 6.4 shows the simple main effect of NNE in different types of food stores. Grocery stores in food swamp and food oasis neighborhoods had substantial differences regarding healthy food availability since there were non-overlapping CI, which also justifies the appropriateness of examining the simple main effect of NNE.

Table 6.10 Multiple comparisons between different pairs in each NNE by ST.

ST	(I) NNE	(J) NNE	Mean Difference (I)-(J)	SE	p value	97.5% CI for difference
Grocery Store	Food Desert Mean: 16.667 SE: 9.720	Food Swamp Mean: 51.613 SE: 6.873	-34.946*	11.904	0.018	(-68.238, -1.655)
	Food Swamp Mean: 51.613 SE: 6.873	Food Oasis Mean: 49.770 SE: 5.195	1.843	8.616	1.000	(-22.251,25.938)
	Food Oasis Mean: 49.770 SE: 5.195	Food Desert Mean: 16.667 SE: 9.720	33.103*	11.021	0.015	(2.281, 63.925)
Convenience Store	Food Desert Mean: 7.527 SE: 4.860	Food Swamp Mean: 5.287 SE: 3.968	2.24	6.274	1.000	(-15.306,19.786)
	Food Swamp Mean: 5.287 SE: 3.968	Food Oasis Mean: 5.511 SE: 4.860	-0.224	6.274	1	(-17.770,13.322)
	Food oasis Mean: 5.511 SE: 4.860	Food Desert Mean: 7.527 SE: 4.860	-2.016	6.873	1	(-21.137,17.205)

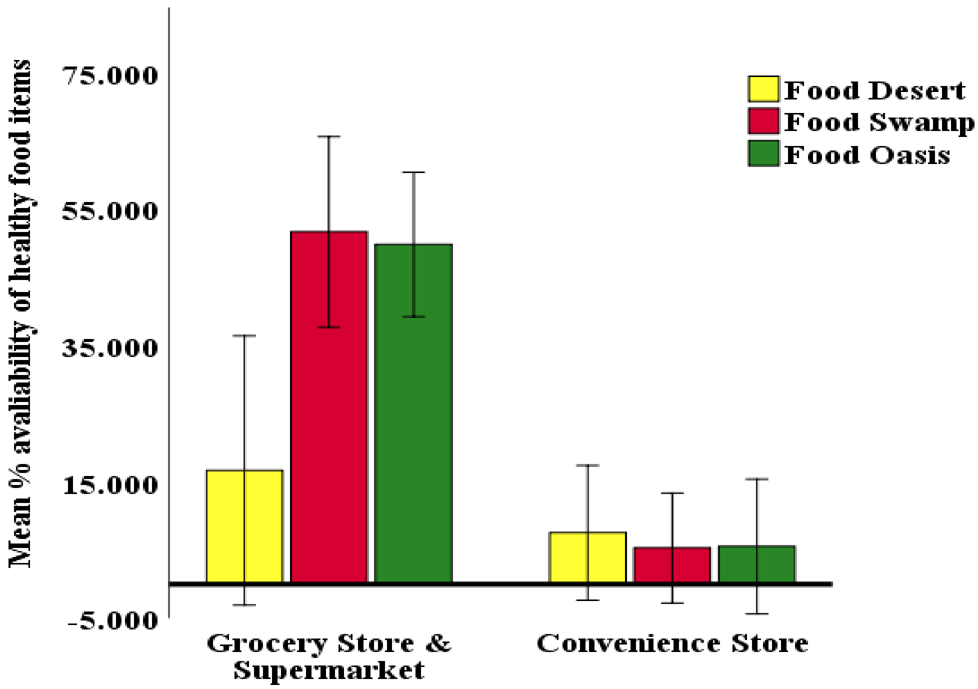


Figure 6.4 Difference in the mean percentage of healthy food availability by NNE and ST.

Food price

The paired t-test was conducted to compare the price between healthy items and regular alternatives. Table 6.11 shows that the prices for most healthy (lower fat, lower calorie, and whole grain) items were not significantly different from the comparable regular items (14 out of 19 pairs, 73.684%). For these 14 pairs, I found similarly priced healthier and standard options for pear in syrup, milk (128 oz, 64 oz, and 32 oz), cheerio, coke, and juice (i.e., 90% -110% of regular, $p > 0.05$). The prices for certain pairs of healthy vs. regular options, such as peach in syrup, lactose-free milk, yogurt, cheese, tortilla, pasta, and pretzel, were not significantly different but marked (i.e., < 90% or > 120 % of healthy vs. regular). The mean prices for mixed fruit in light syrup was 109.574 % of the heavy syrup ($p = 0.048$); whole wheat bread was 106.298% of its white counterpart (p = 0.025). The most significantly substantial differences were found in the higher cost of chicken breast skinless (153.347 % of regular, p = 0.035), lean ground beef (135.507 % of regular, p = 0.002), and chips (124.915 % of regular, p = 0.005).

Table 6.11 Paired t-test for healthy items vs. regular ones.

	Healthy Item (I)		**Regular item (J)**			**Difference of (I) –(J)**		**p-value**
	Mean	**SD**	**Mean**	**SD**	**Percentage** [a]	**Mean**	**SD**	
Pear in syrup	0.097	0.031	0.089	0.038	108.989	0.008	0.008	0.238
Peach in syrup	0.094	0.036	0.076	0.031	123.09	0.018	0.018	0.423
Mixed Fruit in syrup	0.116	0.041	0.106	0.044	109.574	0.01	0.008	0.048*
Milk (128 oz)	3.909	0.871	4.084	0.92	95.715	-0.175	0.689	0.19
Milk (64 oz)	2.919	0.939	3.031	0.859	96.305	-0.112	0.392	0.166
Milk (32 oz)	1.919	0.41	1.935	0.421	99.173	-0.016	0.08	0.327
Lactose Free (64oz)	4.571	1.501	3.769	0.777	121.279	0.803	0.995	0.077
Yogurt	0.200	0.247	0.198	0.245	80.916	0.002	0.032	0.881
Cheese	0.151	0.126	0.120	0.057	126.087	0.031	0.07	0.374
Tortilla	0.148	0.064	0.167	0.082	88.599	-0.019	0.056	0.196
Bread	0.148	0.062	0.139	0.06	106.298	0.009	0.014	0.025*
Pasta	0.289	0.412	0.228	0.314	126.948	0.061	0.105	0.141
Cheerio	0.235	0.118	0.261	0.086	90.051	-0.026	0.039	0.368
Beef	0.318	0.071	0.235	0.055	135.507	0.083	0.026	0.002**
Chicken Breast	0.206	0.075	0.134	0.054	153.347	0.072	0.061	0.035*
Coke	0.078	0.035	0.078	0.035	100	—[b]	—[b]	—[b]
Juice	0.123	0.065	0.128	0.07	96.184	-0.005	0.027	0.412
Chip	0.532	0.192	0.425	0.15	124.915	0.106	0.155	0.005**
Pretzel	0.420	0.221	0.52	0.227	80.802	-0.1	0.164	0.059

Note: a percentage was calculated based on the average price for a healthy item compared to its regular alternative; b: comparison cannot be made due to that the price between diet coke and regular coke is exactly same for each store.

The price of healthy food and regular options was averaged to calculate their ratio. If the ratio is larger than 1, representing that healthy food is much more expensive than a regular one in that store. Two convenience stores did not have any single healthy options, and they were excluded in the analysis. Two-way ANOVA was also performed on the ratio. As per the result in Table 6.12, there were no significant main effects of ST (F (1,33) = 0.812, p = 0.374) and NNT (F (2,33) = 0.848, p = 0.437) on the dependent variable, neither did the effect of their interaction (F (2,33) = 0.636, p = 0.536). Even though the difference was not significant, I observed that on average healthy options were priced higher than regular ones in both grocery and convenience stores (Table 6.13). Healthy food priced lower than regular one was observed in the food desert neighborhood (ratio = 0.999). In the other two neighborhoods, food was priced higher than its regular counterpart, especially in food oasis.

Table 6.12 Results of two-way ANOVA on healthy vs. regular price ratio.

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**p**	**Partial η2**
Corrected Model	0.089	5	.018	1.035	0.413	0.136
Intercept	30.048	1	30.048	1756.184	0.000	0.982
ST	0.014	1	0.014	0.812	0.374	0.024
NNE	0.029	2	0.015	0.848	0.437	0.049
ST* NNE	0.022	2	0.011	0.636	0.536	0.037
Error	0.565	33	0.017			
Total	43.724	39				
Corrected Total	0.653	38				
R Squared = 0.176 (Adjusted R Squared = 0.107)						

Table 6.13 Descriptive statistics of price ratio by ST and NNE.

	Mean	**SD**	**N**
ST			
Grocery store	1.099	0.147	12
Convenience Store	1.030	0.120	27
NNE			
Food Desert	0.999	0.013	10
Food Swamp	1.062	0.157	14
Food Oasis	1.075	0.145	15

Food quality

Table 6.14 shows that the percentage of fruits and vegetables (F&V) were in acceptable quality in the 35 food stores ranging from 50% to 100%. However, the values of 1st quartile, median, and 3rd quantile were all 100% (Table 6.14). It indicates that most of the food stores had a high percentage of acceptable F&V, which can be justified by its frequency bar (Figure 6.5). The histogram was severely left-skewed, and 30 out of 35 stores had 100% F&V in acceptable quality.

Table 6.14 Descriptive statics on percentage of F&V in acceptable quality.

	Min	1st Quartile	Median	3rd Quartile	Max	Mean	SD
% foods were in acceptable quality	50.000	100.000	100.000	100.000	100.000	95.210	12.382

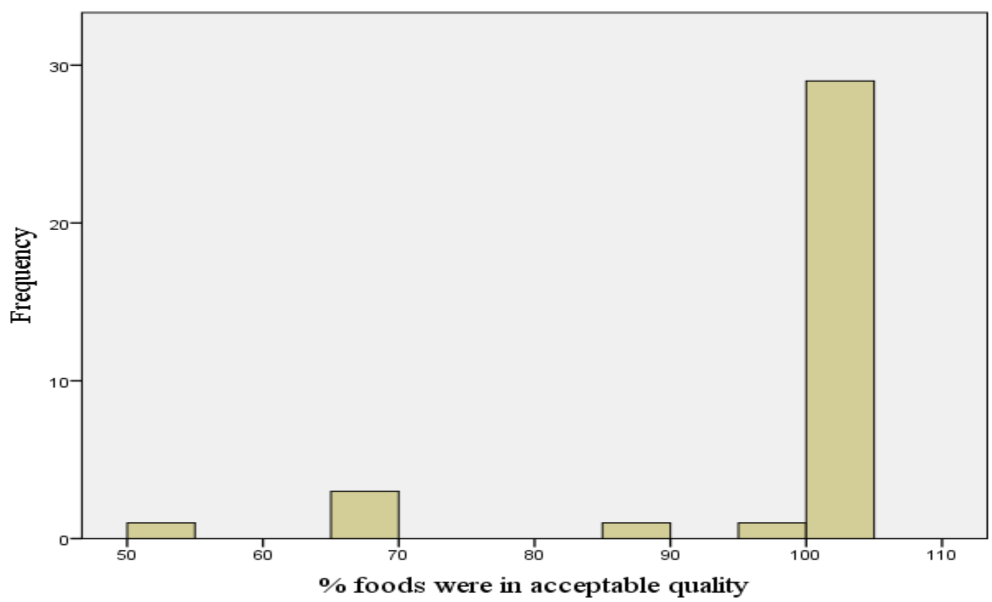


Figure 6.5 Histogram of percentage of F&V in acceptable quality.

I initially performed a two-way ANOVA on the percentage of F&V in acceptable quality by ST and NNE. Nevertheless, the dependent variable was severely left-skewed

(Figure 6.5) and was not normally distributed (Shapiro-Wilk test = 0.444, df = 38, p = 0.000). I applied different transformation techniques (square root, natural logarithm, logarithm, and reciprocal) to the dependent variable. Neither of them alleviated the abnormality issue. For this reason, the use of nonparametric two-way ANOVA seems plausible since it does not require data distribution. The fact is that nonparametric two-way ANOVA does not exist. As a result, we performed nonparametric one-way ANOVA on the percentage of F&V in acceptable quality by ST and NNE, respectively. Therefore, a Mann-Whitney U test was conducted on the quality percentage by ST, and a Kruskal-Wallis H test was performed by NNE. Table 6-15 shows that there was no significant difference between ST (e.g., p = 0.801) and NNE (e.g., p = 0.272) in terms of quality percentage. Despite no significant difference across store type and neighborhoods, the descriptive statistic in Table 6-16 informs us that grocery stores (98.654% ± 3.625%) and food oasis neighborhood (97.436% ± 9.245%) had highest acceptable quality of foods, whereas F& V in convenience stores and food desert neighborhood were in lowest quality on average.

Table 6.15 Results of Mann-Whitney U Test and Kruskal-Wallis Test.

Null Hypothesis	**Test**	**p-value**	**Decision**
The distribution of Quality Percentage is the same across categories of ST	Mann-Whitney U Test	0.801	Retain the null hypothesis
The distribution of Quality Percentage is the same across categories of NNE	Kruskal-Wallis H Test	0.272	Retain the null hypothesis

Table 6.16 Descriptive statistics of percentage of F&V in acceptable quality by ST and NNE.

	Mean (%)	SD (%)	N
ST			
Grocery Store	98.654	3.625	13
Convenience Store	93.182	15.135	22
NNE			
Food Desert	89.352	18.530	9
Food Swamp	97.051	9.233	13
Food Oasis	97.436	9.245	13

Food labelling

40 out of 41 food stores had more than 75% of the foods that were labeled (Figure 6.6). The store with the value 62.5% seems an outlier. Since the data is not normally distributed, I used a Friedman two-way ANOVA to perform the analysis, and it revealed no significant differences in food labeling by store type or neighborhood (Chi-square = 8.386; p= 0.078). In addition, the histogram was normally distributed (Shapiro-Wilk test = 0.954, df = 40, p = 0.108) after eliminating this outlier. I performed a two-way ANOVA analysis on the percentage of food labeling for 40 stores by ST and NNE. Table 6-17 shows that there was neither significant interaction of ST and NNE (F (2,34) = 0.654, p = 0.527) nor the main effects on food labeling (F (1, 34) = 0.897, p = 0.350; F (2, 34) = 0.457, p = 0.457) (Table 6.17). The same results held by a parametric two-way ANOVA that omitted an outlying observation. Although there was no significant difference of food labeling across store types and neighborhoods, I observed that on average grocery stores and food oasis neighborhood had the highest percentage of labeled foods (Table 6.18), while convenience stores and food swamp neighborhoods were in lowest percentage of food labels (Table 6.18).

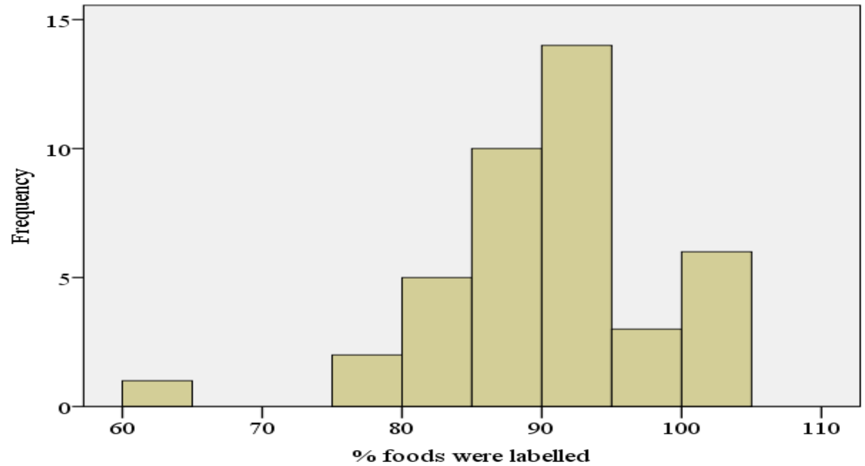


Figure 6.6 Histogram of percentage of food items with labelling.

Table 6.17 Results of two-way ANOVA on percentage of labelling foods.

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**p**
Corrected Model	117.707	5	23.541	0.590	0.708
Intercept	242733.407	1	242733	6081.957	0.000
ST	35.786	1	35.786	0.897	0.350
NNE	63.996	2	31.998	0.802	0.457
ST* NNE	52.173	2	26.087	0.654	0.527
Error	1356.954	34	39.910		
Total	332174.961	40			
Corrected Total	1474.661	39			
R Square = 0.180 (Adjusted R Square = 0.076)					

Table 6.18 Descriptive statistics of percentage of labelled foods by ST and NNE.

	Mean (%)	**SD (%)**	**N**
ST			
Grocery Store	92.170	4.123	13
Convenience Store	90.326	6.908	27
NNE			
Food Desert	90.953	6.452	10
Food Swamp	89.716	7.394	16
Food Oasis	92.288	4.21	14

Discussion and Conclusion

I developed the M-TxNEA-S food instrument by customizing the original TxNEA-S to measure nutrition environment in Austin, Texas. Even though there are many other nutrition environment instruments available, the M-TxNEA-S instrument is the most appropriate to use in this study because of three advantages. First, our instrument, like the TxNEA-S tool, contains a comprehensive scope as well as a series of food items that are culturally important to ethnic subgroups such as Hispanics/Latinos in Texas, which adequately reflects residents' dietary behaviors and food preferences. Second, I reduced the redundancy of the original TxNEA-S by removing several food items that are relatively less important to Texans’ diet. Meanwhile, we added meats (i.e., beef and chicken breast) and beverages (i.e., cokes) that are consumed by Texans on a daily base. Third, in addition to the assessment of food availability, price, and quality that are commonly evaluated by existing food instruments (e.g., NEMS-S and TxNEA-S), I included food labels in the instrument, which are believed to impose a positive influence on people’s food purchasing behaviors. These are the three merits that make our instrument stand out.

M-TxNEA-S has a high inter-rater reliability, which indicates that the protocols and instructions of my instrument are valid and are sound enough to prepare data raters to collect reliable data. It also complies with Glanz, et al. (2007)’s recommendation that the modifications of NEMS-S and its variants must evaluate the reliability. Compared to the studies that did not test reliability (Cassady, Jetter, and Culp 2007; Harrison, et al. 2007; Palermo, et al. 2008; Bovell-Benjamin, et al. 2009), the use of inter-rater reliability in this study has a distinct advantage. My survey tool (M-TxNEA-S) was compared with the

TxNEA-S; it turned out that in general, the reliability scores of my tool were higher. In addition, I reported inter-rater reliability for food availability, price, quality, and labels, while TxNEA-S only reported food availability inter-rater reliability. Previous studies often focused on whether healthy foods are less available in low-income or minority neighborhoods (Horowitz, et al. 2004; Jetter and Cassady 2006; Coulter 2009; Woodham 2011); no studies to date have paid attention to the differences in food availability by neighborhood nutrition environments (NNE, different levels of accessibility to food outlets). To the best of our knowledge, this is the first to examine the consumer nutrition environment in the food desert, food swamp, and food oasis neighborhoods. The result shows that the effect of NNE on healthy food availability was determined by store types. Specifically, grocery stores from food swamp and food oasis neighborhoods had significantly greater healthy foods availability than convenience stores, but this was not the case for food desert neighborhood. The reason is that the grocery stores in food deserts are small in size and have a limited range of healthy foods. Although these small groceries are named as grocery stores, they essentially more like convenience stores. Therefore, the difference by store type does not apply to food desert neighborhood.

Provided that food desert was in low SES, food swamp was in medium SES, and food oasis was in high SES. The findings thus support the previous results that healthy foods are less available to grocery stores in a low-income neighborhood (Coulter 2009; Glanz, et al. 2005; Gloria and Steinhardt 2010; Woodham 2011). However, it does not align with Gloria and Steinhardt's (2010) finding that income has a positive relationship with healthy food availability. In our study food swamp neighborhood had a low to medium SES but had the highest mean healthy foods availability in grocery stores. As its name indicates, food swamp neighborhood should have excessive access to unhealthy foods, and it thus was hypothesized to have a low to moderate amount of healthy foods. Nevertheless, our finding does not correspond to this hypothesis.

The possible explanation for the discrepancy above can be attributable to the following perspectives. In the previous two studies, I used grocery stores (supermarkets) and convenience stores (fast food restaurants) as proxies of healthy and unhealthy food entities to measure food accessibility and delineate food deserts and food swamps. The use of store types to represent food healthfulness could be misleading and result in incorrect desalination of food swamp. The results inform us that it is inadequate to consider food access to stores and neighborhood deprivation when identifying food insecure areas. It should conduct a food survey and calculate its healthfulness score in a neighborhood before the delineation; the healthfulness score then can be combined with access to food stores and sociodemographic deprivation for the identification. A methodology as such is capable of capturing a completer neighborhood food characteristic and should be considered to use in future studies. In 2016, the sustainability office of the city of Austin begun to adopt such a method from the Johns Hopkins Center for a Livable Future program. Seven data collectors visited the stores and objectively assessed the types and quantities of food available. This information was used to create a Healthy Food Availability Index (HFAI) score for each store that is based on the availability of foods in the following categories: protein, dairy, produce, and grains. HFAI scores were averaged at the block group, and a block group that did not meet a certain average score was considered not healthy. These areas were overlaid on a map with information on household income, physical proximity to healthy food stores, and vehicle availability. Any place where all four variables overlapped completely was identified as Healthy Food Priority Areas. Their analysis is still ongoing, and the final report will be released by the end of 2018. The problem with this method is that it requires a substantial investment of time to assess each food store. Therefore, compromising between the accuracy of results and efforts investment could be a critical question for researchers.

Regarding the price comparison between healthy and regular options, our result was consistent with Glanz, et al. (2005)'s finding that most healthy choices were not significantly different from their regular alternatives. The present study also found that low-fat meats (i.e., beef and chicken breast), whole wheat bread, and low-fat chips had a significantly higher price than regular ones, which was corresponding to Glanz and colleagues' finding. However, our study did not find any significant price difference between 100% juice and juice drink. Healthy food was found to be cheaper in food desert neighborhood and convenience stores, which contradicts to our hypothesis. However, the price difference was minimal across store type and communities, since the average price ratio of healthy food to its regular option was not significantly different by store type and neighborhoods.

The quality of F&V across different stores was consistently high. It is consistent with previous research that produces is generally in high quality in stores (Coulter 2009; Cummins, et al. 2009). I also found that F&V quality did not significantly differentiate by store type and neighborhood nutrition environment. It is consistent with Coulter's (2009) finding that quality of F&V is not significantly different by either neighborhood-level race or income. However, it does not support Cummins, et al. (2009)'s results that medium-sized stores in a high-SES neighborhood had significant highest-quality fresh produce, while stores that are secondary in a deprived neighborhood had the lowest quality. It does not either align with Glanz et al. (2005)'s finding that grocery stores and high-income neighborhood had a high quality of foods. Our findings do not support earlier findings that fruit and vegetable quality was differentiated by store type and neighborhood marginalization (Horowitz, et al. 2004; Turrell, et al. 2002).

These findings reveal possible roles for healthy food availability, food affordability, food quality and labeling as a mediating micro-environmental variable that may affect individuals' purchasing behavior and diet choice. The healthy food availability did vary significantly by neighborhoods and store types, but food price, food quality, and food labeling did not. The possible explanation is that the aggregation data may smooth out the variation. However, the interpretation of this finding is challenging because the deprivation amplification theory does not seem to be held in healthy food availability in food swamp neighborhood and on food price. The specific items sold in different types of stores, urban or rural setting of neighborhoods, and chain stores' reputation and marketing share all could influence these domains.

This study has several limitations. First, the present study has a limited sample size for food stores. As a rule of thumb, a sample with more than 30 observations is considered a reliable size to perform statistical analysis (Hogg and Tanis 2009). Our data sample (i.e., 41 stores in total) meets this criterion. However, the sample size of grocery stores (13 stores) is only 46% of the convenience stores (28 stores), leading to the interpretation power of the findings not as compelling as I anticipated. Second, the survey was only conducted one time. Some studies conducted the survey twice, and the second survey is held two or three weeks later after the first one (Glanz, et al. 2005; Gloria and Steinhardt 2010). The two surveys enable them to perform test-retest reliability, which tests whether the food availability, price, quality, and labels in a food store change dramatically for several weeks. A high test-retest reliability suggests a sound design of the survey instrument. I did not conduct the second survey due to that there is not sufficient financial support for this project. Future studies would conduct two surveys when external funding is available. Third, the survey was done only in the summer season. It cannot capture the price and quality fluctuation and variations in different seasons. It is a cross-sectional study, making it difficult to conclude that the within-store characteristics (i.e., healthy food availability) consistently differ in grocery stores across the neighborhoods. Fourth, our instrument m-TxNEA-S, like the renowned NEMS, lacks measures to assess other in-store characteristics and the physical environment. For instance, how does the interior (i.e., lighting quality and cleanliness) of the store look like? How many bus stops near the store? Is the store located in a busy street? Are there any sidewalks or bike lanes located adjacent to the store? Do they have defined lines? All these features significantly affect the possibility of food shopping at the store and their consumers' shopping experience. Future studies should consider these factors and develop a more robust instrument to assure a high-quality food environment assessment. Lastly, unhealthy food entities are constituted of convenience stores and fast food restaurants. I did not examine the in-restaurant features in this study since it uses a distinct food survey instrument from food stores. NEMS-R and TxNEA-R are two common instruments used in the U.S and Texas. I may examine fast food restaurants nutrition environment in future studies.

7 CONCLUSIONS

Texas is the second largest state in the U.S., and more than one-third of adults are either overweight or obese. Austin, the capital of Texas, also is facing an elevating prevalence of obesity. Retail food environment plays an essential role in shaping individuals' eating behaviors and diet-related health outcomes. Therefore, an increasing number of studies have endeavored to link the food environment (such as access to food stores) with health disparities. My dissertation reflects such a research frontier. It is composed of three parts of inter-related studies (presented in Chapter Four to Six) that indicate one common theme — food insecurity. Two components of food insecurity, such as food desert and food swamp, were examined. The food desert is created by the suburbanization of grocery stores (or supermarkets), leading to that local residents lack access to healthy foods. Food swamp is the result that the rapid growth of convenience stores and fast food industry makes residents over-exposed to unhealthy food options. With this central theme in mind, I examined food access, economic and sociocultural marginalization, and consumer nutrition environment in Austin, Texas. Results are beneficial for stakeholders and planners to tackle food insecurity issue. This chapter summarizes the findings and policy implications from my analyses. Then, the contributions of this dissertation are highlighted. Last, the related limitations are discussed, and future directions are proposed.

Findings and Implication

The three chapters produced a series of findings, and they are listed in the below. These findings can provide useful policy implications for food planning and management.

Major findings

Chapter Four proposed a novel spatial access assessment method --- multi-mode Huff-based 2SFCA. I applied this method to food access study in Austin, Texas. Grocery stores & supermarkets and convenience stores & fast food outlets in Austin were used as proxies for simple classification of healthy and unhealthy foods. The significant findings are: (1) the accessibility index to both healthy and unhealthy food stores increase as the impedance coefficients β increase. (2) It reveals a distinct urban core – peripheric disparity for access to both food entities. Block groups in urban core areas enjoyed the best access, while the ones in peripheral areas were least accessible to food stores. (3) In comparison with the single-mode Huff-based 2SFCA, for most of the impedance coefficients the proposed method did not exhibit significant difference, except when β equals to 1.4 (access to healthy food) or 1.5 (access to unhealthy foods) (4) the proposed method produced low accessibility index in urban core areas and high accessibility index in peripheric areas than the single-mode method; (5) the multi-mode Huff-based 2SFCA method reveals more spatial variability than its single-mode counterpart in the study area. These results illustrate that the proposed method is an appropriate model to assess food accessibility in food environment studies. They also support that the food accessibility disparities between urban and peripheric areas persist regardless of the methods and that the interventions for procuring foods should be prioritized to the underserved regions of Austin, Texas.

Chapter Five explored how economic and sociocultural deprivation relates to food accessibility. Two deprivation indices (EDI and SDI) were constructed. The spatial lag model reveals that the access to healthy food entities was only significantly related to EDI; a neighborhood with low economic deprivation enjoyed better access to grocery stores and supermarkets. By contrast, the access to unhealthy food outlets (convenience stores and fast food outlets) was proved to have a significantly positive relationship with the SDI; residents from high sociocultural deprivation neighborhoods had better access to unhealthy foods. The semi-parametric GWR was used to explore the heterogeneity of associations between food access and the two deprivation indices. The result shows that the access to healthy foods was globally negatively associated with the EDI but was locally varied with the SDI, which indicates that the EDI is a constant stronger predictor of healthy food access than the SDI. Most of the block groups did not exhibit significant relation with the SDI except the ones in the Southwest and northwest Austin. The access to unhealthy food entities had a globally insignificant relationship with the EDI but a locally varying relation with the SDI. The significantly positive association was observed in Southwest Austin. The bivariate local Moran’s I was employed to identify food deserts and food swamps in Austin. The results show that food deserts were in East Austin where is close to Austin international airport, and food swamps were in Northeast Austin.

Chapter Six examined consumer nutrition environment in the food desert, food swamp, and food oasis neighborhoods. Consumer nutrition environment usually contains food availability, price, and quality. My research extended it and included food labels, which has been underexamined by other studies. I found that grocery stores had a significantly higher availability of healthy foods than convenience stores across the neighborhoods except for the food desert neighborhood. For grocery stores, the mean percentage of healthy food availability in food swamp and food oasis neighborhoods was significantly higher than the food desert; food swamp and food desert neighborhoods had the highest and lowest percentages of healthy food availability, respectively. For convenience stores, food desert neighborhood had the highest healthy food availability, but the difference was insignificant between the three communities. We also found that the price for most of the healthy items was not significantly different from their regular counterparts. For food quality and labeling, there was not any significant difference between store types and neighborhoods. In summary, store type and neighborhood nutrition environment are both related to healthy food availability, but not food price, quality, and labeling in our study.

Implications

Our research has important implications for policymakers to deal with the issues that emerged in the food environment and nutrition. The specific implications from each of the three parts of the analysis are summarized below.

Chapter four has two implications. First, it informs stakeholders that different interventions should be initialized in the urban core and peripheric areas. Most studies used single-mode (usually by automobile) to measure spatial food accessibility. The consequence of using single-mode method is to produce higher accessibility of both healthy and unhealthy foods in urbanized areas. The single -mode method would underestimate food deserts and overestimate food swamps in urban core areas. It indicates that most studies identify smaller food deserts and larger food swamps than they are, meaning that residents in urban center ask for fewer resources to fight for the limited access to healthy foods but more resources to restrict the excessive access to unhealthy foods than they need. Therefore, policymakers and authorities should increase the amount of funding to improve healthy food access but reduce resources to limit unhealthy food access. The situation in peripheric areas, however, is opposite to the urban core areas. As a result, policy in peripheric regions as opposed to that in urban core areas, that is --- reducing resource that promotes healthy food access but increasing interventions to fight against the overexposure to unhealthy foods.

Second, the finding of chapter four also implies that peripheric and urban core areas should focus on different aspects of interventions due to the core- peripheric disparities of food accessibility. In peripheric areas, residents often have poor access to healthy food outlets; interventions thus should aim at reducing travel barriers/impedance to grocery stores and other healthy food entities. A straightforward solution is to open new food stores. However, it is not practical to establish a large grocery store in peripheric areas because lower population density makes larger stores challenging to make a profit; smaller stores have more power to compete with larger ones, which may lead to a shutdown for the larger store in a short period. As a result, interventions should emphasize the increases of other healthy food resources such as community gardens and farmers' markets provided that these two entities were found to enhance the consumption of F&V. In urban areas, residents enjoy better access to grocery stores and supermarkets since many food entities prefer to be established there. However, some neighborhoods in the urbanized area such as the University of Texas campus potentially suffer inadequate access to healthy foods. Most of them are low-SES students; it is likely for them not to have their vehicles and have to rely on public transportation to commute. Therefore, policy should focus on increasing income, reinforcing vehicle ownership, and improving public transit systems, etc.

Chapter Five sought to understand the role of neighborhood deprivation on food accessibility, by examining their relationships in Austin, Texas using different sets of spatial statistical methods. The results indicate that the intervention on healthy and unhealthy food access should consider various aspects of neighborhood deprivation. The intervention on promoting healthy food access should be prioritized on the improvement of economic opportunities and food affordability, such as reducing the number of residents who are below the poverty line, increasing household income, and decreasing unemployment rates. By contrast, the intervention on restricting unhealthy food access should more emphasize socio-cultural factors. Hispanics/Latinos and their cultures are inclined to choose high-fat food; programs that educate them to adopt a healthier lifestyle gradually might be useful for them. Meanwhile, residents without higher education might lack awareness and knowledge on healthy and less healthy foods; the solution is to develop programs to reinforce their positive attitude on a healthy diet and negative view on high-fat and low-nutritious eating. Besides, the spatial heterogeneity of the relationship between food access and deprivation implies that food programs and policies should vary from place to place to response distinctive neighborhood conditions.

Chapter Six focused on in-store characteristics aiming at improving the consumer nutrition environment, and the findings are beneficial for designing intervention programs in food insecure neighborhoods. Specifically, healthy food availability was not significantly different in grocery stores and convenience stores. The fact is that the grocery stores are small groceries with carrying a limited range of fruits and vegetables, beverages, and milk, which are more like convenience stores. Food desert lacks larger grocery stores. The establishment of larger food stores in these areas might not be so rewarding. Establishing more small groceries with healthier options might alleviate food issues in the food desert. Also, it is informative to authorities that grocery stores in food desert neighborhood should increase their healthy food availability, keep their low food price, improve food quality and labeling. Grocery stores in food swamp neighborhood had the highest healthy food availability, which indicates that their residents in this neighborhood are likely to procure enough healthy foods from stores and eat healthily at home. The creation of food swamp in this area might be more related to the overwhelmingly unhealthy food environment such as convenience stores and fast food restaurants. Stakeholders should consider designing interventions on fast food restaurants and convenience stores instead of grocery stores. The possible solution is to limit the number of fast-food restaurants and convenience stores by zoning and policies. Grocery stores in food oasis neighborhood had high food availability, food quality, and labeling, but the averaged healthy food was priced higher than regular options. It indicates that policy that is aiming at reducing the healthy food price would be beneficial for residents even though this neighborhood did not have serious issues.

Contribution

This dissertation has contributed significantly to food environment research. It can be specified into theoretical contribution and methodical contribution, which are described in the below.

Theoretical contribution

This research theoretically contributes to our understanding of the food environment. It proposed a theoretical framework, which was shown in Figure 3-1. This framework was built upon the Glanz, et al. (2005)'s one that food environment assessment should examine both the community nutrition environment and consumer nutrition environment. More importantly, it extends Glanz, et al. (2005)'s framework by introducing the deprivation amplification hypothesis, which concerns that residents in deprived areas are more restricted by their environment and more vulnerability to adverse health. For example, how does access to foods change when considering marginalized population groups who lack personal vehicles? (Chapter Four); why do some neighborhoods suffer from the limited access to healthy food (or overexposed to unhealthy food) and neighborhood marginalization simultaneously, but others do not? (Chapter Five); Is healthy food less available, higher in price, lower in quality and labeling in food stores in deprived areas such as food desert and food swamp neighborhoods? (Chapter Six).

The results in the three chapters illustrate food environment research should consider how adverse environment could potentially affect residents' access to foods as well as consumers' experience in food stores. It helps to understand the local food desert and food swamp issue, and which neighborhoods should be prioritized for interventions. What is more, our study supports Macintyre, Macdonald, and Ellaway (2008) conclusion that the deprivation amplification hypothesis holds only in some neighborhoods in the study area. That is to say, residents from deprived neighborhoods are not necessarily always suffer from a poor food environment (Luan 2016), and some deprived neighborhoods have good access to healthy foods (i.e., food swamp neighborhood). This argument is still controversial in food studies since the various ways of characterizing food environment and using statistical approaches could alter their relationships even in the same neighborhood. However, the argument also motivates researchers to put more efforts into exploring in which extent environment influences residents' food access, fruit and vegetable consumption, dietary behaviors, and health outcomes.

Inspired by Luan (2016)'s framework that summarized five aspects (dimension, data, scale, strategy, and methodology) in the neighborhood food environment assessment, I further extended his framework via adding statistics. The new one (Figure 7.1) has six elements: method, statistic, dimension, data, strategy, and scale. I emphasize the importance of using spatial statistics to explore marginalization and food environment because many studies ignored the effect of spatial autocorrelation and spatial structure in relationships, which could result in bias even incorrect conclusions. Moreover, this dissertation has addressed the under-studied areas in food environment research. For instance, chapter four filled up the gap that few modeling approaches (i.e., accounting for realistic constraints) measure geographic food accessibility, and chapter five fixed the problems of using traditional statistics. Chapter Six is a direct food observational survey in food stores. The in-store survey examined the consumer nutrition environment, which has not been extensively explored relative to community nutrition environment. The survey compensates the deficiency of using secondary data and emphasizes the importance of collecting first-hand or primary data.

Meanwhile, the survey instrument I developed is a relative measure of consumer nutrition environment involving multiple types of food stores, both healthy and less healthy food items. It solves the issues of using an absolute assessment of the food environment in many existing studies. More importantly, as Luan (2016) indicated, the proposed six-element framework (Figure 7.1) emphasizes diversified aspects in the food environment. The framework in Figure 7.1, in conjunction with Glanz et al.'s conceptual framework and deprivation amplification hypothesis (Figure 3.1), provides a much more comprehensive assessment of the food environment in scope. Future research could benefit from these frameworks to enable scholars to quickly conceptualize their research design by referring to our proposed frameworks.

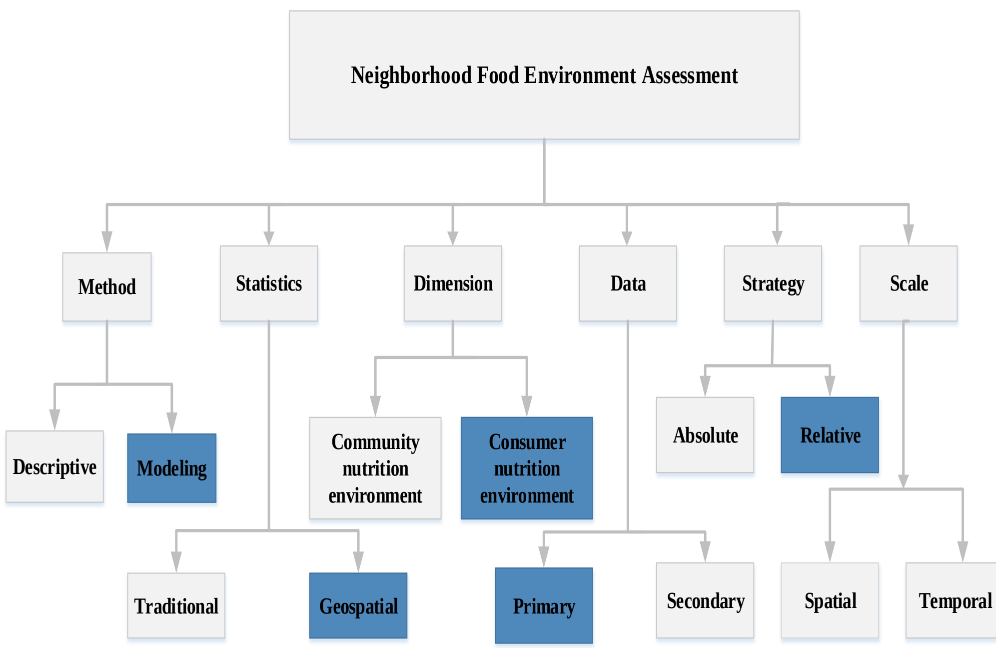


Figure 7.1 Six basic elements in food environment study.
Note: this figure modified is derived based on Luan (2016, 14). Blue boxes represent under-studied areas in food environment research and have been addressed by this dissertation.

Methodological contribution

The significant contribution of this dissertation is to address methodological gaps in existing quantitative research on retail food environment. For the study one (Chapter Five), my contribution is to propose a novel sophisticated modeling approach to measure food access. Compared with many researchers and agencies who used the descriptive approach (i.e., density and proximity), the use of the proposed method has merits. It accounts for various realistic restrictions including the distance-decay effect between catchments, store attractiveness, and competition between store supply and people demands, and multiple transportation modes. These merits can provide a more accurate picture of food access and reveal more variability of the food environment. Moreover, the proposed method can be applied in many other areas, especially in metropolitan areas with diversified means of transportation modes (e.g., New York City, San Francisco, Beijing, Shanghai, etc.)

The study two (Chapter Five) examined economic and sociocultural inequities in access to healthy and unhealthy food retailers though spatial statistics. The traditional statistics such as OLS and logistical regression used in many studies ignored the spatial dependence in the residuals, leading to the associations potentially invalid. Compared with the widely used traditional statics, the use of different spatial statistics, such as spatial lag model and spatial error model (spatial autocorrelation), semi-parametric GWR (spatial heterogeneity), hot spot analysis (spatial dependence), can capture different aspects of spatial extent problems in many geographic studies. It thus is beneficial for subsequent research such as the food survey since it helps to delineate more accurate surveying neighborhoods.

The food survey in the Chapter Six can further add to better understanding healthy food availability, food price, quality and labeling in food deserts and swamps. I used two-way ANOVA to analyze the interaction effect of store type and neighborhood nutrition environment. In comparison with other studies that only used one-way ANOVA, or used Two-way ANOVA but did not explicitly explore the interaction effect, our method thoroughly investigated the interaction effect and carefully examined the simple effects of store type and neighborhood nutrition environment on consumer nutrition environment. The interaction revealed that the impact of neighborhood nutrition environment on healthy food availability is dependent on store types, which have not been detected in other studies. Therefore, an entire exploration of ANOVA interaction assures a more robust result on relationships.

In summary, this dissertation research contributes to understanding the food insecurity challenges in Austin, Texas through exploring food access issues and identifying food deserts and food swamps in the study area. In addition to developing a new method to measure food accessibility and to examine how geographical accessibility is associated with the socioeconomic status of neighborhoods, my research contributes to the understanding of how the community food environments are different through in-store auditing of food availability, price, quality, and labels.

Limitation and Future Direction

Even though this dissertation conducted a systematically in-depth examination on food access, food dessert, food swamp, and consumer nutrition environment in the city of Austin, future work may advance the research in several directions. Figure 7.2 shows under-studied areas in food environment research but has not been addressed by this dissertation. The contents in the green boxes should be explored in future studies. The specific limitations are discussed, and corresponding future directions are proposed in the following section.

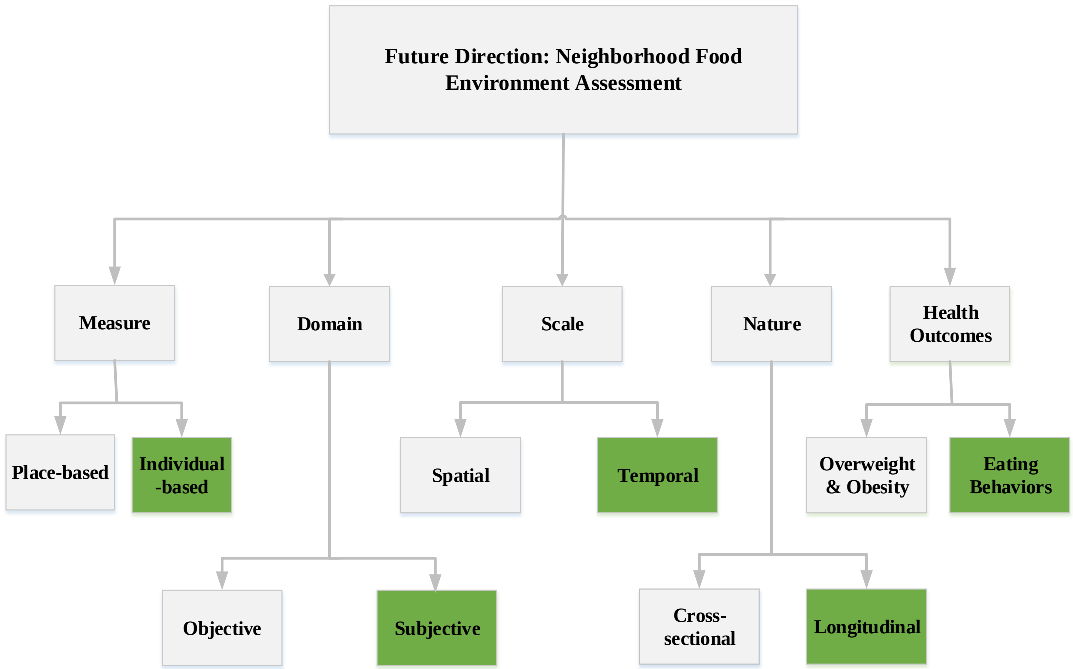


Figure 7.2 Potential directions for future research.
Note: green boxes represent under-studied areas in food environment research but have not addressed by this dissertation.

Focusing more on individual-based measure

This research (Chapter 4) measured the distance from census block group population-weighted centroid to a store, assuming home to store travel is the way most people access food stores. People, however, do not always travel from home to stores. They go to work, schools, and churches and often purchase food on the way. A consumer who likes eating at fast food restaurants or convenience stores may have high access to these stores but may pass by them on the way to a supermarket. In other words, the access I measured was area-based (known as potential access) rather than individual-based (known as realized access). The area-based accessibility shows where people in a neighborhood could shop and approximates the potential food exposures of a community, while the individual-based one reveals where an individual's travel path and quantifies individuals' actual food exposure through movement and mobility (Luan 2016). An Individual's movement in the course of the daily activities usually is termed as “activity space” (Li and Kim 2018). Many individual-based access studies have adopted this concept and quantified it in different manners (Li and Kim 2018; Kestens, et al. 2012; Kestens, et al. 2010). A typical one is Li and Kim (2018)'s work. The authors quantified an individual's healthy food accessibility by three activity space measures (e.g., transport mode-weight standard deviation ellipse (SDE), time-weighted SDE, and route network buffer). They found that the three activity space measures exhibited similar patterns, and individuals in high SES (i.e., high income and employment) had larger activity space and higher food accessibility. Research as such would provide robust measures of food access, and their association with demographic variables would be more authentic because multiple finer-scale (individual-level) measures were employed. As Luan (2016) claimed that measuring activity space is rewarding for food interventions since it would assist in identifying clusters for individuals who have restricted access to foods, therein the interventions for areas with clustering individuals could be prioritized.

The study employed area-based measures of food access, while future research could use alternative individual-level measures such as activity space to reveal more details in the quantifications. By doing this, we not only will be able to model the multiple locations to the exposures to food retailers beyond the residential areas but also measure actual food accessibility. Nevertheless, the individual-based measure requires an individual's travel dairy or survey data. While obtaining detailed travel survey data could be formidable due to the concern of privacy issue. Besides, the travel diary data usually is extensive, and processing a massive dataset requires high-computation computers and more complex algorithms. Therefore, using area-based or individual-based approach is contingent on specific research setting and data availability. If computation algorithms and data availability are not problems, I prefer to quantify food access with the individual-based measure. Meanwhile, in the future study, I may conduct a comparison between area-based and individual-based food environment measures to examine to what extent the two methods are differentiated from each other.

Moving forward to subjective assessment

This dissertation examined both community and consumer nutrition environments, which are classified as objective measures. The counterpart of an objective assessment is a subjective one, and it refers to people's perception of food environment (Glanz, et al. 2005; Leia Michelle Minaker 2013). In recent years, researchers have begun to realize the importance of perceived food environment and reached a consensus that food studies should include the subjective measure as an indispensable component. Several studies have measured the perceived food environment (Carbonneau, et al. 2017; Jilcott, et al. 2009; Moore, Roux, and Brines 2008). These studies often utilized a qualitative method such as survey questionnaires and semi-structured interviews to understand how residents perceived their food environment. The advantages of subjective assessment are multifaceted. One the one hand, it can avoid the problem of using secondary data that commercial addresses of food stores might be significantly different from their actual locations. On the other hand, residents' personal selection preferences and perceptions (i.e., the perceived opinions on food quality, acceptability, and affordability) may provide a complementary alternative of the objective assessment of food access.

The future study will be exploring subjective measures of the food environment. Different from the traditional survey and semi-structured interviews, I would use Twitter streaming data to extract information about their perceived healthy eating because of the emergence of using social media big data in GIScience. The following keywords would be searched in the streaming data --- healthy eating, healthy diet, healthy meal, healthy food, nutritious food, low-nutrition food, balanced diet, gluten-free diet, low-salt diet, salt-free diet, light diet, low-fat diet, high-vitamin diet, and a high-protein diet, etc. Meanwhile, I will compare the perceived measure with the objective measure to examine whether the two align with each other or not. Understanding the relationship between the two measures can help refine the assessment of the food environment and facilitate interventions. For example, if the objective food access is predicting residents' eating behaviors better than subjective measure, policy program aiming at improving population-wide eating behaviors and health outcomes should emphasize more on increasing spatial access to healthy foods (or reducing spatial access to unhealthy foods) in a neighborhood; otherwise, interventions should focus on reinforcing residents' awareness of creating a healthy diet environment in their neighborhoods.

In the well-cited Glanz and colleagues' nutrition framework (our framework is derived from this framework), people's perception is hypothesized to be a potential mediator of the association between objective food environment and diet-related outcomes. However, few studies have examined the mediating effect of perceived measure impose on objective measures and health outcomes. Future research should put effort into this direction, since examining mediating pathways could facilitate our understanding of which features of the food environment are influencing health outcomes therein more effective intervention programs could be developed and delivered.

Incorporating temporal analysis to food environment study

Spatial and temporal analyses are two themes of geographic studies. However, in this dissertation, we did not explore the temporal aspect of the food environment. Time plays an important role in shaping the availability of food outlets therein influencing food access. A few studies are focusing on the temporal analysis of food access (Chen and Clark 2013; Chen and Clark 2016; Shannon 2016; Zenk, et al. 2011). Their foci might be a little different, but all emphasize the importance of incorporating time as a constraint. For instance, food store opening hours can create an hourly and daily variation of availability to purchasing food (Chen and Clark 2013; Chen and Clark 2016). GPS units can be used to trace people's real-time movements over one or more days to understand how their movements influence food shopping behaviors of certain population groups (Shannon 2016; Zenk, et al. 2011). Public transit schedules and commuting flows into accessibility can be integrated to study the dynamics of food access, as well as how these dynamics affect shopping behaviors at daily scale (Widener, et al. 2015; Widener, et al. 2013). Other temporal examples contain the annual opening and closing of farmers markets and produce stands in urban spaces (Lucan, et al. 2015; Widener, Metcalf, and Bar-Yam 2011), which measured yearly and seasonal variations of food access in different locations. These studies have advanced our understanding of how time can shape food access and food environment.

For future study, I would be exploring the spatial-temporal trend of food insecurity issue in Austin, Texas. A food store data in the 20 years (1996-2016) will be utilized to examine the yearly dynamics of the food desert and food swamp. I intend to witness how the distribution of food desert and food swamp evolves in each year, and which issue is becoming more prevalent during the study period. Moreover, there are several farmer's markets in Austin, but we did not consider them when we assessed the food environment. Farmer's markets are often operated on a seasonal base. Future research may explore the seasonal variation of access to farmer's markets and whether these farmer's markets could alleviate the food desert problem in East Austin.

Exploring more on longitudinal study

Our study is cross-sectional in nature as only one-year data has been used. Extent food environment studies exclusively focused on cross-sectional study, part of because multiple-year data is not available. The drawback of cross-sectional study is that the causal relationship cannot be assured. In many studies, researchers attempted to establish a causal relationship between food environment, socio-demographic deprivation, and health outcomes. However, the cross-sectional data only can secure an association instead of causality between them. The present study explored the relationship food access and marginalization. Even if I employed advanced spatial statistical methods to populate their relationships, the causality cannot be warranted. As a result, the interpretation of our result should be taken with caution. Future study will incorporate multiple-year data as a longitudinal study to explore the causal pathway of food environment and marginalization on health. The longitudinal study in conjunction with spatial-temporal analysis can reveal more authentic relationship or causality between variables, therein more issues in food environment are anticipated to be detected and solved.

Connecting food environment to health outcomes

A growing number of studies have attempted to explore the association between food environment with the prevalence of weight status and cardiovascular diseases (Morland and Evenson 2009; Brown and Miller 2008; Rundle, et al. 2009; Powell, Chaloupka, and Bao 2007). However, our study did not examine the linkage between geographic distribution of food store disparities with diet-related outcomes. I proposed to associate food desert and food swamp with the prevalence of obesity in Austin, Texas. The obesity data was obtained from the Texas Bureau of Motor Vehicles. It contains approximately 900, 000 records for individual driver license adult holders in Austin. However, using driver license data was questioned by the issue that people often misreport their heights and weights on driver license (Ossiander, et al. 2004). My dissertation committee members recommended me to remove this part and explore this topic in my future study.

I definitely will continue my research along this direction when a sound dataset is available. In addition, according to the Social-Ecological Model, individual behaviors and neighborhood food environment jointly determine a specific group's health outcomes. However, current studies often lack individual behavior variables such as the consumption of fruits and vegetables in the analysis. Meanwhile, individuals might purchase nutritious food in convenience stores or buy low-nutrient dense junk food in supermarkets and grocery stores. Therefore, the use of detailed purchasing behavior data could dramatically improve the accuracy of the assessment. In summary, future studies would benefit from including information about individual behaviors to more fully understand how these healthy and unhealthy food outlets contribute to disparities in obesity and diet-related diseases in central Texas.

APPENDIX SECTION

Appendix A

Creating a Multi-modal Network Using GTFS Text File and Street Shapefile

Step 1: Generate transit routes and stations. GTFS text file contains latitude/longitude information of transit stations, and this information is read by **Generate transit lines and stops** tool in Add GTFS Data to a Network Dataset toolkit embedded in ArcGIS. A point shapefile that contains all transit stops in Austin is created to store the spatial information. Then it generates straight lines to connect two adjacent stops; lines are converted to line shapefiles (i.e., transit routes). In total, 2684 transit stops and 3232 transit route segments were generated.

Step 2: Create connectors between transit stops to street networks. Road network and transit stops (or transit lines) come from different resource; there might be gaps between transit stops and road network. It is impossible for people to across the gaps unless there is a “bridge” connecting transit stops and streets. The **Generate Stop-Street Connectors** tool can create a “connector” as a “bridge” to facilitate pedestrians to walk through. The “connector” is a short straight line and are perpendicular to streets, and it connects transit system and street network. The “connector” might not exist in real world but is an important step. By creating connectors, transit lines and street network only are connected at stops, which prevents pedestrians walking on transit lines.

Step 3: Create a multimodal transportation network. With “creating a multimodal network dataset” toolkit provided in ArcGIS 10.5 Network Analyst Extension, a multimodal transit network could be created. The setup of three transportation modes is shown in the below.

(3a) Transit mode. There is an assumption that people walk on the street to transit stops, then take transits to other transit stops to get off, and walk on street lines to arrive at destinations. I assume an ingress, egress, and transfer with a walking speed of 0.05 miles/minute. For the connectors created in step 2, I apply a delay of 0.5 minutes for transitions between streets and transit lines (to represent boarding a transit vehicle) and a delay of 0.5 minutes for transitions from transit lines to streets (to represent alighting). I also create a pedestrian restriction to prevent pedestrians walking on four types of roads: 1(Interstate, Fwy, Expy, and Toll), 2(US and State Highways), 15 (Private Road), and 17 (Platted Row/Unbuilt). The evaluator is vital for setting up the network because it determines how the network uses the fields in shapefile tables. For transit network, it uses TransitEvaluator from the Add GTFS to a Network Dataset toolkit to calculate transit travel time along transit lines. The TransitEvaluator determines the travel time across that transit line by looking up the available transit trips in the GTFS schedules at the appropriate time of day and summing the wait time for the trip plus the ride time from the current stop to the next. In my analysis I use a general Monday to calculate the transit travel time because I am not focusing on a specific timetable or schedule of the transits. A general workday like Monday can serve the analysis.

(3b) Drive mode. The setup of drive mode is not as complex as the transit mode. The street shapefile has a field “minutes”, which is the minimum travel time on each street segment. The evaluator uses the “minutes” to calculate drive time on street. In addition, the evaluator uses one- way (such as “B”, “FT”, and “TF”) field in street shapefile as one-way restriction.

(3c) Walk mode. The setup of walk mode is identical to the pedestrian part of transit mode. I assume that the walk speed is 0.05 miles/minute. The walk mode also uses the pedestrian restriction for four types of roads: 1, 2, 15, and 17.

Appendix B

Food Auditing Instrument

Modified Texas Nutrition Environment Assessment (M-TxNEA-S, Texas State University)
Cover Page

Rater ID: ________

Store ID: ________

Date: __/____/____

Start time: ____: ____ ○ am ○ pm

End time: _____:____ ○ am ○ pm

Number of cash registers:
(Including self-checkout)

○ 1-2 ○ 3-4 ○ 5-6 ○ 7+

WIC Store Certification:

○ Certified ○ Not Certified ○ Unknown

Food Stamp Store Certification:

○ Certified ○ Not Certified ○ Unknown

Type of Stores

○ Chain/Supermarket

○ Convenience store

○ Grocery store

○ Supercenter

Write down any comments that will help us understand the rating of any item in the food outlet

____

____

Modified Texas Nutrition Environment Assessment (M-TxNEA, Texas State University) **Measure # 1A Fresh Fruits**						
Rater ID: __________ Store ID: ___________ Date __/___/___						
Item	**Availability**	**# of Varieties**	**Lowest Price**	**Quality**	**Labeled**	**Comments**
1. Apple	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
2. Bananas	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
3. Cantaloupe*	O Yes O No	______	$_ ._ _/ O 1PC	O U O A	O Yes O No	________
4. Grapes	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
5. Grapefruit	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
6. Honeydew*	O Yes O No	______	$_ ._ _/ O 1PC	O U O A	O Yes O No	________
7. Mango	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
8. Oranges	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
9. Papaya	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
10. Peaches	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
11. Pears	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
12. Pineapple	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
13. Plum	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
14. Strawberry	O Yes O No	______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	________
15. Watermelon*	O Yes O No	______	$_ ._ _/ O 1PC	O U O A	O Yes O No	________
Total # of Types	______					
Total # of Varieties		______				

Modified Texas Nutrition Environment Assessment (M-TxNEA, Texas State University) Measure # 1B Fresh Vegetables						
Rater ID: _______		Store ID: ___________			Date __/___/___	
Item	Availability	# of Varieties	Lowest Price	Quality	Labeled	Comments
1. Avocado	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
2. Asparagus	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
3. Bell pepper	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
4. Broccoli	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
5. Corn	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
6. Cucumber	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
7. **Greens	O Yes O No ( )	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
8. ** Leaf Lettuce	O Yes O No ( )	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
9. Onion	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
10. Summer Squash (yellow)	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
11. Tomatoes	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
12. Potatoes	O Yes O No	_______	$_ ._ _/ O 1LB O 1PC	O U O A	O Yes O No	_______
13. Carrots	O Yes O No	_______	$_ ._ _/ O 1bg	O U O A	O Yes O No	_______
Total # of Types	_______					
Total # of Varieties		_______				

Modified Texas Nutrition Environment Assessment for Stores (M-TxNEA-S, Texas State University)
Measure # 2A Frozen Fruits & Vegetables

Rater ID: __________ **Store ID:** ___________ **Date** __/___/___

Item	Availability	Lowest Price	Labeled	Comments
Frozen Fruits				
1. Blueberries	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
2. Mango	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
3. Peaches	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
4. Strawberries	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
5. Mixed fruits	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
Frozen Vegetables				
Item	**Availability**	**Lowest Price**	**Labeled**	**Comments**
6. Broccoli	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
7. Carrots	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
8. Corn (white, yellow & mixed)	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
9. Green Beans	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
10. Spinach	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
11. Green Peas	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
12. Mixed Vegetables	O Yes O No	$_ ._ _/oz.	O Yes O No	_________

Modified Texas Nutrition Environment Assessment for Stores (M-TxNEA-S, Texas State University) Measure # 2B Canned Fruits & Vegetables				
Rater ID: __________ Store ID: ___________ Date __/___/___				
Canned Fruits				
Item	Availability	Lowest Price	Labeled	Comments
1. Pears in heavy syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
2. Pears in light syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
3. Peaches in heavy syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
4. Peaches in light syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
5. Mixed fruit in regular/heavy syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
6. Mixed fruit in light syrup	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
Canned Vegetables				
Item	Availability	Lowest Price	O Yes O No	Comments
7. Corn (plain, whole kernel)	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
8. Green Beans	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
9. Tomatoes (plain only)	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
10. Green Peas	O Yes O No	$_ ._ _/oz.	O Yes O No	_________
11. Mixed Vegetables	O Yes O No	$_ ._ _/oz.	O Yes O No	_________

Modified Texas Nutrition Environment Assessmentfor Stores (M-TxNEA-S, Texas State University)
Measure # 3 Dairy: Milk and Cheese

Rater ID: _______ **Store ID:** ________ **Date** ____/___/____

Milk				
Item	**Availability**	**Lowest Price**	**Labeled**	**Comments**
1.Whole Milk, Gallon (128 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
2.Whole Milk, Half Gallon (64 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
3.Whole Milk, Quart (32 oz)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
4.Reduced Fat Milk (2%), Gallon (128 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
5.Reduced Fat Milk (2%), Half Gallon (64 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
6.Reduced Fat Milk (2%), Quart (32 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
7. Low-Fat milk (1%), Gallon (128 oz)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
8. Low-Fat milk (1%), Half Gallon (64 oz)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
9. Low-Fat milk (1%), Quart (32 oz)	○ Yes ○ No	$_ ._ _/ea	○Yes ○ No	_________
10. Fat-Free Milk (Skim), Gallon (128 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
11. Fat-Free Milk (Skim), Half Gallon (64 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
12. Fat-Free Milk (Skim), Quart (32 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
13.Lactose-Free Milk, Whole, Half Gallon (64 oz.)	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
14. Lactose-Free Milk, Reduced (2%), Half Gallon	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
15. Lactose-Free Milk, Low-fat (1%), Half Gallon	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
16. Lactose-Free Milk, Fat-free (Skim), Half Gallon	○ Yes ○ No	$_ ._ _/ea.	○Yes ○ No	_________
How much shelf space is dedicated to 1% milk and skim milk of total shelf space in the milk area? Note: Measure Only if skim and/or 1% is available.		○< 25% ○> 25% and <50% ○> 50% and < 75% ○ > 75%		
Yogurt and Cheese				
Item	**Availability**	**Lowest Price**	**Labeled**	**Comments**
17.Regular Yogurt (flavored or plain)	○ Yes ○ No	$_ ._ _/oz.	○Yes ○ No	_________
18.Light Yogurt ("light", "nonfat" or "fat free)	○ Yes ○ No	$_ ._ _/oz.	○Yes ○ No	_________
19.Regular Cottage Cheese (full-fat, or 2% milkfat)	○ Yes ○ No	$_ ._ _/oz.	○Yes ○ No	_________
20. Light Cottage Cheese (1% or less and non-fat)	○ Yes ○ No	$_ ._ _/oz.	○Yes ○ No	_________
How much shelf space is dedicated to Light/n onfat yogurt and light cottage cheese of total shelf space in the yogurt/cottage cheese area?		○< 25% ○> 25% and <50% ○> 50% and < 75% ○ > 75%		

Modified Texas Nutrition Environment Assessment for Stores (M-TxNEA-S, Texas State University) Measure #5 Meats and Alternatives				
Rater ID: _______	Store ID: ___________		Date __/___/___	
Item	Availability	Lowest Price	Labeled	Comments
1.Ground Beef (> 10% fat)	O Yes O No	$_ ._ _/LB	O Yes O No	_______
2.Lean Beef (< 10% fat)	O Yes O No	$_ ._ _/LB	O Yes O No	_______
# of varieties of lean beef available	O 0 O 1 O 2 O 3 O 4 O 5 O > 6			
3.Chicken Breast Skinless	O Yes O No	$_ ._ _/LB	O Yes O No	_______
4.Regular Chicken Breast	O Yes O No	$_ ._ _/LB	O Yes O No	_______
# of varieties of Chicken Breast Skinless	O 0 O 1 O 2 O 3 O 4 O 5 O > 6			
Meat Alternatives				
5.Legumes (lentils, dry beans, peas)	O Yes O No	$_ ._ _/LB	O Yes O No	_______
# of varieties of legumes		O 0 O 1-2 O 3-4 O 5-6 O > 6		
6.Tofu	O Yes O No	$_ ._ _/LB	O Yes O No	_______
# of varieties of tofu		O 0 O 1-2 O 3-4 O 5-6 O > 6		

Modified Texas Nutrition Environment Assessment for Stores (M-TxNEA-S, Texas State University) Measure #6 Beverages				
Rater ID: _______	Store ID: __________		Date __/___/___	
Item	Availability	Lowest Price	Labeled	Comments
1. Diet Coke	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
1. Regular Coke	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
2. 100% Juice	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
3. Juice Drink	O Yes O No	$_ ._ _/oz.	O Yes O No	__________

Modified Texas Nutrition Environment Assessment for Stores (M-TxNEA-S, Texas State University)
Measure #7 Snacks

Rater ID: _______ Store ID: __________ Date __/___/___

Item	Availability	Lowest Price	Labeled	Comments
1. Low-fat chips (less than 3g fat/per 1 oz serving)	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
2. Regular chips (more than 3g fat/per 1 oz serving)	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
Total Varieties of low-fat Chips	O 0 O 1-2 O 3-4 O 5-6 O > 6			
3. Low-fat hard pretzels (less than 3g fat/per 1 oz serving)	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
4. Regular hard pretzels (more than 3g fat/per 1 oz serving)	O Yes O No	$_ ._ _/oz.	O Yes O No	__________
Total # varieties of low-fat pretzels	O 0 O 1-2 O 3-4 O 5-6 O > 6			

- **Please review this instrument to make sure all items are complete.**
- **Once each page has been verified as complete, return to the cover page to record any comments and enter your end time.**
- **Thank you very much for completing the survey.**

Appendix C

Training Guide for Surveyors

Background

M-TxNEA-S is a modification of the TxNEA (Texas Nutrition Environment Assessment for Stores), which was a modification of the NEMS (Nutrition Environment Measurement Survey. The objective of the NEMS was to measure differences in the nutrition environment (stores and restaurants) between different SES-level neighborhoods. The TxNEA was an adaptation of the NEMS tool that included additional foods that were culture-specific to the minority populations of Texas (Hispanic and African-American).

The M-TxNEA-S store audit tool, a modification of the TxNEA-S, was designed to measure the availability, accessibility, and affordability of foods in food outlets in Austin metropolitan area. The objective of this store audit tool is to measure the food accessibility and price in different food environment (e.g. food deserts, food swamps, and food oases). We also intend to compare the healthier foods with their regular counterparts to examine whether there is a difference between these items.

General Instructions

To prepare for every store audit be sure to have the following items with you:

- blank audit tools
- clipboard
- letter to store owners/managers
- maps
- pens/pencils
- site visit schedule
- student I.D.
- training guide

When you start your survey, be sure to obey the rules in the bellow.

- Permission should be obtained from store managers before data collection begins (present letter of explanation).
- Avoid early morning audits because it is likely to interfere with shelf stocking.
- Please do not interfere with the normal flow of business.
- Returning items to where they were after collecting the information.
- Record start and stop time at each location and circle the am or pm next to the blank boxes.
- Check to be sure survey is complete before leaving each location.

Note: All the photos included in this manual are borrowed from the Google Images.

Cover Page

- Record your Rater ID number, the Store ID number, Date, Start time and End time on the cover page.
- Record number of cash registers, WIC store Certification, Food Stamp store certification on the cover page.
- Check the types of food stores.

Supermarket

- A self-service shop that offers a wide variety of food and household products. It is larger and has a wider selection than earlier grocery stores.
- Comprises meat, fresh product, dairy, baked goods aisles, and so on. Sell a variety of products that are consumed regularly.

Convenience Store

- A small store that offers a limited selection of staple and fresh/raw groceries, non-food and other convenience items, like ready-to-heat and ready-to-eat foods.
- Includes food marts attached to gas stations.

Grocery Store

- Retail food outlet with items from all food categories, like fresh fruit/vegetables, raw meat and other items that need preparation or cooking, and convenience items like chips, canned goods and soda.
- Typically have service deli, service bakery & possibly pharmacy.

Supercenter

- A very large supermarket that sells food and also a wide range of other products.
- Offers a full line of groceries, an optical department, a tire and oil-change department, a grill, portrait studio, hair salon, photo-processing lab, and so on.

Common Columns

Common Columns are those found throughout the instrument. For the sake of brevity, they are discussed here and not repeated in the individual pages of instruction, which are reserved for specific instructions for each category of product. The Common Columns include:

Item: this column indicates the product of interest.

Availability: this column indicates whether items are carried by the store and if so, if they are currently in stock. Check the **Yes** when an item is present in the store. Otherwise, check the **No**. In some occasions, the item is not available because of out-of-stock, please write "**out-of-stock**" on the comment line. If you are not clear why it is not available, just check the "No" and no need to write anything on the comment line.

Lowest Price: this column is used to determine the lowest price of products regardless of whether it is a name brand, store brand, or generic brand product.

Price columns will vary depending on the product. Some columns ask for lowest **price per pound** and some for lowest **price per ounce**. Others ask for lowest **price per package or unit** (e.g. box, bag, can, pkg.). Please make sure to read this column carefully, and record prices accordingly. If you cannot find a price for an item, ask store personnel for help.

Labelled: this column is used to determine whether a food item is labelled or not. If it is labelled, please circle **Yes**; otherwise, circle **No**.

Item Specific Instruction

Measure #1A: Fresh Fruits

- Do not count the fresh fruits in bagged package, and only count the ones in loose package.
- **# of Varieties:** A variety is considered the same type of fruit (e.g. apple) that has a distinct name (e.g. Gala, Granny Smith, Golden Delicious. Here the # of varieties = 3). A single variety may have differences in size (e.g. small and large watermelons). However, these are not counted as two separate varieties.
- **Total # of Varieties:** After finishing counting the varieties of all fruits, write the sum of these totals where indicated. For example, total all varieties of fruits and write number in box to the left. (e.g., 5 varieties apples + 2 varieties avocados + 1 variety banana = 8)
- **Total # of Types:** count all different types of available fruits (i.e., apples, bananas, and oranges = 3). It is worth to note that a fruit having different varieties should be counted as one type. For example, Gala, Granny Smith, Golden Delicious apples, the number of types of apple = 1. Please fill this blank line with the sum of total number of "Yes". Note: Only include types that are listed.
- **Lowest Price:**

  (a) Most of fresh fruits in supermarket and grocery stores is priced by weight (LB). In this case, just write down the price for each pound. Write "NA" on Comments line if price is not available.

(b) In convenience stores fruits are usually sold by piece. Some items such as pears might be priced by piece as well. If you encounter a fruit that is priced individually or by piece (e.g. one pear for $0.99), please write down the price and circle per piece.

Strawberries – determine cheapest packaging option per pound

Example: 1 lb container $1.99; 2 lb container $3.00

2 lb container cheaper at $1.50/lb.

- Quality: This column is distinct for Fruit &Vegetable section. Circle the "A" box if more than 50% of the fruit item you are rating is found to be of acceptable quality (see the photos in the left column); Otherwise, circle the "U" (see the photos in the right column).

Left: good color, fresh looking, firm, and clean (**Acceptable Quality)**.

Right: bruised, old looking, mushy, over-ripe, dark sunken spots in irregular patches, cracked or broken surfaces, signs of shriveling, mold, or excessive softening **(Unacceptable Quality)**.

Measure #1B: Fresh Vegetables

- **# of Varieties**: A variety is considered the same type of vegetables (e.g. corn) that has a distinct name (e.g. white and yellow corns. Here the # of varieties = 2). A single variety may have differences in size (e.g. small and large Haas Avocados). However, these are not counted as two separate varieties.
- **Total # of Varieties**: After finishing counting the varieties of all vegetables, write the sum of these totals where indicated.
- **Total # of Types:** count all different types of available vegetables (i.e. Onion, corn and tomatoes = 3). It is worth to note that a fruit having different varieties should be counted as one type. For example, seedless and seeded watermelon, the number of types = 1. **Note:** Only include types that are listed.
- **Lowest Price:**

  (a) Most of fresh fruits in supermarket and grocery stores is priced by weight (LB). In this case, just write down the price for each pound. Write "N/A" on Comments line if price is not available.

  (b) In convenience stores fruits are usually sold by piece. Some items such as lettuce in grocery stores and supermarkets might be priced by piece as well. If you encounter a fruit that is priced individually or by piece (e.g. one head of lettuce for $1.49), please write down the price and circle per piece.

- Bell peppers: cheapest variety, regardless of color (green, red, yellow, orange). Green usually least expensive.

- Summer Squash (Yellow): there might be other types of summer squash (zucchinic and calabacita). But we only count the yellow one.
- Potatoes: record any types of potatoes but exclude sweet potatoes and yam.

Special Instruction for Greens and Lettuce

- **Greens:** Six greens (spinach, kale, turnip, mustard green, and collard green) are counted. Please circle **Yes** when one of them is available. If none of them are available, please circle **No**. At the bottom of page 2, circle any type of Greens if it is available. Then sum up these types and fill the number in the parenthesis () next to the “No” in the main table. For example, when you only circle spinach and kale, please write 2 in the parenthesis. Then count the number of varieties for each type and write the number in the line at the bottom box. In the end, sum up the total varieties of these types and fill this number in the line in the main table on page 2. Also, do not include pre-packaged ready-to-eat varieties. (washed and bagged greens are not included).
- **Leaf Lettuce**: three types of Leaf Lettuce (romaine, read, and green leaf) are counted. Do not count the iceberg lettuce. Please circle Yes when one of them is available. If none of them are available, please circle No. At the bottom of page 2, circle any type of Leaf Lettuce if it is available. Then sum up these types and fill the number in the parenthesis () next to the “No” in the main table. For example, when you only circle Romaine, please write 1 in the parenthesis. Then count the number of varieties for each type and write the number in the line at the bottom box. In the end, sum up the total varieties of

these types and fill this number in the line in the main table on page2. Please do not include pre-packaged ready-to-eat varieties.

- Quality: This column is distinct for Fruit &Vegetable section. Circle the “A” box if more than 50% of the fruit item you are rating is found to be of acceptable quality (see the photos in the left column); Otherwise, circle the “U” (see the photos in the right column).

- **Left:** good color, fresh looking, firm, and clean (**Acceptable Quality)**.
- **Right:** bruised, old looking, mushy, over-ripe, dark sunken spots in irregular patches, cracked or broken surfaces, signs of shriveling, mold, or excessive softening **(Unacceptable Quality)**.

Measure #2A: Frozen Fruits & Vegetables

- Exclude those containing fruits with added sugars and/or sauces; see the photo below.

- For mixed frozen fruits, it could contain any combination of 2 or more fruits packaged together.
- For frozen vegetables, do not include items that contain sauces (e.g., butter or cheese).
- For frozen Corn, note that packages may contain yellow, white, or a combination of both to be counted. Remember to only record the lowest price per ounce available.
- For mixed Vegetables, note that packages may contain any combination of 2 or more vegetables together. (e.g., peas and carrots, or corn, peas, carrots and lima beans). This does not include packages that may contain sauces such as Stir Fry Mix containing Teriyaki Sauce.

Measure #2B: Canned Fruits & Vegetables

- Find the lowest price per ounce for each type of fruits
- Do not include fruit gels or fruit suspended in or mixed with other ingredients
- 

Correct Wrong

Measure # 3: Milk & Cheese

- Light Yogurt and light cottage cheese should have “light”, “nonfat” or “fat free” on it.
- Regarrding the shelf space, it does not need to be accurate. An estimation is okay.

Measure # 4: Grain

- **NOTE:** To determine if an item is whole grain/whole wheat, count only items with one of these phrases on the package:

    - 100% whole grain
    - whole grain
    - 100% whole wheat
    - whole wheat
    - Exception: Corn tortillas

To verify if an item qualifies, check the ingredient list: If the first item includes the word "enriched," it is **NOT** a whole grain product.

- Lowest Price per ounce: please find the lowest price per ounce.

  NOTE: Each brand item has a size range (e.g. 16-24 oz.). To determine which item has the lowest price **per ounce**, simply divide the price by ounces:

Price	**Calculation**	**Result**
$2.89 (16 oz.)	$2.89/16 oz. = $0.18 per oz.	Since $0.149 is cheaper per ounce than $0.18, record $0.149/oz.
$3.59 (24 oz.)	$3.59/24 oz. = $0.149 per oz.	

- For cheerios, the healthy option has less than7g sugar per serving and whole grain. Whereas the regular one should have more than 7g sugar per serving and not whole grain

Measure # 5: Meats and Alternatives
• For ground beef, the percentage of fat should be more than 10%. For lean beef, the fat should be less than 10%. Please read the label carefully before you fill in the blanks.

Measure # 6: Beverages
• Always record the lowest price per ounce regardless of the size.

Measure # 7: Snacks
• For chips, the healthier option has less than 3g fat/ per 1 oz. serving. Whereas the regular option has more than 3g fat/ per 1 oz. serving. • For hard pretzels, the healthier option has less than 3g fat/ per 1 oz. serving. Whereas the regular option has more than 3g fat/ per 1 oz. serving.

Appendix D

IRB Approval Document

TEXAS STATE UNIVERSITY
SAN MARCOS
The rising STAR of Texas

In future correspondence please refer to 2018534

May 24, 2018

He Jin
Texas State University
601 University Drive.
San Marcos, TX 78666

Dear He:

Your IRB application 2018534 titled "Obesity and Food Insecurity: Food Deserts, Food Swamps, and Middle-aged Adults' Obesity in Austin, Texas" was reviewed and approved by the Texas State University IRB. It has been determined that risks to subjects are: (1) minimized and reasonable; and that (2) research procedures are consistent with a sound research design and do not expose the subjects to unnecessary risk. Reviewers determined that: (1) benefits to subjects are considered along with the importance of the topic and that outcomes are reasonable; (2) selection of subjects is equitable; and (3) the purposes of the research and the research setting is amenable to subjects' welfare and producing desired outcomes; that indications of coercion or prejudice are absent, and that participation is clearly voluntary.

1. In addition, the IRB found that you need to orient participants as follows: (1) signed informed consent is not required as participation implies consent; (2) Provision is made for collecting, using and storing data in a manner that protects the safety and privacy of the subjects and the confidentiality of the data; (3) Appropriate safeguards are included to protect the rights and welfare of the subjects. (4) Compensation will not be provided for participation.

This project is therefore approved at the Exempt Review Level

2. Please note that the institution is not responsible for any actions regarding this protocol before approval. If you expand the project at a later date to use other instruments, please re-apply. Copies of your request for human subjects review, your application, and this approval, are maintained in the Office of Research Integrity and Compliance.

Report any changes to this approved protocol to this office. All **unanticipated events and adverse events are to be reported to the IRB** within 3 days.

Sincerely,

Monica Gonzales

Monica Gonzales
IRB Regulatory Manager
Office of Research Integrity and Compliance

CC: Dr. Yongmei Lu

OFFICE OF THE ASSOCIATE VICE PRESIDENT FOR RESEARCH
601 University Drive | JCK #489 | San Marcos, Texas 78666-4616
Phone: 512.245.2314 | *fax:* 512.245.3847 | WWW.TXSTATE.EDU

This letter is an electronic communication from Texas State University-San Marcos, a member of The Texas State University System.

Appendix E

Verbal Consent Document

TEXAS STATE

VERBAL CONSENT

Study Title: Food Deserts, Food Swamps, and Middle-aged Adults' Prevalence of Obesity in Austin, TX

Principal Investigator: H. Hannah Jin **Faculty Advisor: Dr. Yongmei Lu**

My name is Hannah Jin and I am a doctoral student at Texas State University. I am doing this study to understand the consumer nutrition environment in food stores in the different neighborhoods in Austin, particularly in Food deserts and food swamps. I am asking your store to participate in this research because it is in a neighborhood of my interests. I'm going to tell you a little bit about the study so you can decide if your store can be included in this study.

The consumer nutrition environment includes in-store characteristics such as food availability, food price, and food quality. I intend to compare food environment in three neighborhoods in Austin, TX. The three neighborhoods have different food characteristics, and are selected from the neighborhoods that are classified as food deserts (defined as low access to healthy foods), food swamps (defined as high access to high calorie unhealthy foods), and food oasis (defined as high access to healthy food in affluent neighborhoods, opposite to the concept of food deserts). Your store is located in one of such neighborhoods. I am asking your support to allow me and my research assistants to conduct an in-store food audit.

To protect sensitive information and to minimize any possible risk, I provide the following assurances: (1) No identifiable information about your store will be divulged by the research team. (2) The original store survey data will not be disclosed to any parties other than the research team. (3) The publication of any maps or tables in the dissertation or later research articles will ensure that the particular stores are not identifiable. (4) The survey will not interrupt your normal operations of business; there will be minimum involvement of your employees or staff during the survey.

Your participation is voluntary. If your stores opt to participate in this study, the two surveyors will enter the store and conduct the audit. The food survey will take approximately 20 minutes or less to complete. You can stop the survey at any time.

Please let me know if you have questions for me or would like me to clarify any issues related to the survey.

Do we have your verbal consent to conduct the survey in this store?

RESPECT FOR PERSONS ★ BENEFICENCE ★ JUSTICE
TEXAS STATE
Exempt Review
05/24/2018
INSTITUTIONAL REVIEW BOARD

Appendix F

Letter to Store Managers

To Store Manager/owner,

My name is H, Hannah Jin, a doctoral student at Texas State University. I am conducting a research to examine the consumer nutrition environment in the different neighborhoods in Austin, particularly in Food deserts and food swamps. Your store is invited to participate in this research because it is in a neighborhood of my interests.

The consumer nutrition environment includes in-store characteristics such as food availability, food price, and food quality. I intend to examine food environment in three neighborhoods in Austin, TX. The three neighborhoods should have different food characteristics, and they are selected from the neighborhoods that are classified as food deserts (defined as low access to healthy foods), food swamps (defined as high access to high calorie unhealthy foods), and food oasis (defined as high access to healthy food in affluent neighborhoods, opposite to the concept of food deserts). I will conduct an in-store food audit in selected stores. The findings will provide specific knowledge about the connection between food environment at neighborhood level and at store level. It will help us recommend strategies to improve food environment aiming at benefit population health. Particularly, I am hoping that this research will provide some guidance for us to tackle food insecurity problem in Austin.

This study involves no foreseeable concerns or risks for your store. However, in order to provide the best privacy protection for the participating stores, my research team and I will take every necessary measure to minimizing any possible risks. Particularly, we will ensure the following:

(1) The purpose of the data collection is to conduct a scientific research. The research team and I have no intention or plan to evaluate the store or your policies.
(2) The survey will not interrupt your normal operations of business; the involvement of your employees or staff will be minimum in the survey.
(3) Your store name, address, policies, and food prices will not be published or publicized.

TEXAS STATE

(4) Findings from individual stores will be completely confidential; the information gathered from individual stores will be combined with that from many other stores and that the results will be reported in statistical form only (i.e., percentages and totals);
(5) The team members must never divulge store names or factual information about any store surveyed;
(6) The researchers will secure data documents within locked locations in Department of Geography at Texas State University. Security codes will be assigned to digital records in computer. The access to identifying information will be limited to the research team;
(7) We will keep all data securely three years in Department of Geography at Texas State University. All data will be properly disposed and destroyed three years later.

If you have any questions or concerns, please feel free to contact me or my research advisor, Dr. Yongmei Lu. Our contact information is included below.

He Jin, Doctoral student
Department of Geography
570-213-2815
h_j29@txstate.edu

Dr. Yongmei Lu, Professor
Department of Geography
512-245-1337
yl10@txstate.edu

Thank you so much for your participation. Your support is important to the research.

Sincerely,

He Jin

H. Hannah Jin

Appendix G

Recruiting Flyer

Appendix H

Estimation of the Number of Pieces of F&V Per Pound

Item	Size	Approx. Pieces per Pound
Apple	Medium	3 ea
Banana	Medium	3 ea
Cantaloupe	Medium	1/3 cantaloupe
Grapefruit	3-3/4″dia (8 oz for each)	2 ea
Mango	Median (7 oz for each)	2 1/4 ea
Orange	2 5/8” diameter	3 ½ ea.
Papaya	Medium size	1/2 papaya
Peach	Medium: 2 5/8” dia.	4 ea.
Pear	Medium	2 ½ ea
Pineapple	Medium	1/2 Pineapple
Plum	Medium	7 each
Watermelon	15” long x 7 ½” dia.	1/10 of a melon
Avocado	medium	3 ea
Asparagus	one bunch	1 ea
Bell Pepper	Medium: 2 ½” dia.	4 ea
Broccoli	medium head (9 oz)	2 ea
Cucumber	medium size (7 oz)	2 1/4 ea
Green	medium size (8 oz)	2 ea
Leaf Lettuce	one bunch	1 ea
Onion	medium (110 g)	4 ea
Summer Squash	medium size	2 ea
Tomato	medium (4.3 oz)	4 ea
Potato	medium (5.3 oz)	3 ea
Carrot	medium (6 to 7 inches length) (4 oz)	4 ea

REFERENCES

Advocacy California Center for Public Health. 2007. Searching for Healthy Food: The Food Landscape in California Cities and Counties: California Center for Public Health Advocacy Davis (CA).

Algert, Susan J, Aditya Agrawal, and Douglas S Lewis. 2006. "Disparities in Access to Fresh Produce in Low-Income Neighborhoods in Los Angeles." American journal of preventive medicine 30, no. 5: 365-370.

Alnasrallah, Mohammad. 2015. Geographic Disparities of Obesity as a Public Health Issue in Summit County, Ohio: Kent State University.

Anderson, Beth, et al. 2011. "Peer Reviewed: Fast-Food Consumption and Obesity among Michigan Adults." Preventing chronic disease 8, no. 4.

Anselin, L. 1995. "(1995a)‘Local Indicators of Spatial Association-Lisa’geographical Analysis 27: 93-115."

Anselin, Luc. 2005. "Spatial Regression Analysis in R–a Workbook." Urbana 51: 61801.

Anselin, Luc, and Arthur Getis. 1992. "Spatial Statistical Analysis and Geographic Information Systems." The Annals of Regional Science 26, no. 1: 19-33.

Anselin, Luc, and Daniel A Griffith. 1988. "Do Spatial Effecfs Really Matter in Regression Analysis?" Papers in Regional Science 65, no. 1: 11-34.

Antrim, Aaron, and Sean J Barbeau. 2013. "The Many Uses of Gtfs Data–Opening the Door to Transit and Multimodal Applications." Location-Aware Information Systems Laboratory at the University of South Florida 4.

Apparicio, Philippe, Marie-Soleil Cloutier, and Richard Shearmur. 2007. "The Case of Montreal's Missing Food Deserts: Evaluation of Accessibility to Food Supermarkets." International journal of health geographics 6, no. 1: 4.

Austin, S Bryn, et al. 2005. "Clustering of Fast-Food Restaurants around Schools: A Novel Application of Spatial Statistics to the Study of Food Environments." American Journal of Public Health 95, no. 9: 1575-1581.

Baker, Elizabeth A, et al. 2006. "Peer Reviewed: The Role of Race and Poverty in Access to Foods That Enable Individuals to Adhere to Dietary Guidelines." Preventing chronic disease 3, no. 3.

Ball, Kylie, David Crawford, and Gita Mishra. 2006. "Socio-Economic Inequalities in Women's Fruit and Vegetable Intakes: A Multilevel Study of Individual, Social and Environmental Mediators." Public health nutrition 9, no. 5: 623-630.

Ball, Kylie, Anna F Timperio, and David A Crawford. 2006. "Understanding Environmental Influences on Nutrition and Physical Activity Behaviors: Where Should We Look and What Should We Count?" International Journal of Behavioral Nutrition and Physical Activity 3, no. 1: 33.

Balstrøm, Thomas. 2002. "On Identifying the Most Time-Saving Walking Route in a Trackless Mountainous Terrain." Geografisk Tidsskrift-Danish Journal of Geography 102, no. 1: 51-58.

Barker, M, et al. 2008. "Women of Lower Educational Attainment Have Lower Food Involvement and Eat Less Fruit and Vegetables." Appetite 50, no. 2-3: 464-468.

Beaulac, Julie, Elizabeth Kristjansson, and Steven Cummins. 2009. "Peer Reviewed: A Systematic Review of Food Deserts, 1966-2007." Preventing chronic disease 6, no. 3.

Beaumont, John, et al. 1995. "Report from the Policy Sub-Group to the Nutrition Task Force Low Income Project Team of the Department of Health." Radlett, Hertfordshire: Institute of Grocery Distribution.

Becker, ES, et al. 2001. "Obesity and Mental Illness in a Representative Sample of Young Women." International Journal of Obesity 25, no. S1: S5.

Behjat, Amirmohsen. 2016. "Exploring the Geography of Food Deserts and Potential Association with Obesity in Rural British Columbia."

Beydoun, May A, Lisa M Powell, and Youfa Wang. 2008. "The Association of Fast Food, Fruit and Vegetable Prices with Dietary Intakes among Us Adults: Is There Modification by Family Income?" Social science & medicine 66, no. 11: 2218-2229.

Bhattacharya, Jayanta, Janet Currie, and Steven Haider. 2004. "Poverty, Food Insecurity, and Nutritional Outcomes in Children and Adults." Journal of health economics 23, no. 4: 839-862.

Black, Jennifer L, et al. 2011. "Exploring the Distribution of Food Stores in British Columbia: Associations with Neighbourhood Socio-Demographic Factors and Urban Form." Health & Place 17, no. 4: 961-970.

Blanchard, Troy, and Thomas Lyson. 2006. "Food Availability and Food Deserts in the Nonmetropolitan South." Mississippi, MS: Southern Rural Development Center.

Block, Daniel, and Joanne Kouba. 2006. "A Comparison of the Availability and Affordability of a Market Basket in Two Communities in the Chicago Area." Public health nutrition 9, no. 7: 837-845.

Block, Jason P, Richard A Scribner, and Karen B DeSalvo. 2004. "Fast Food, Race/Ethnicity, and Income." American journal of preventive medicine 27, no. 3: 211-217.

Blok, Jason P, Richard A Scribner, and Karen B DeSalvo. 2004. "Fast Food, Race/Ethnicity, and Income. A Geographical Analysis." American Journal of Preventive Medicine 27, no. 3: 211.

Bodor, J Nicholas, et al. 2008. "Neighbourhood Fruit and Vegetable Availability and Consumption: The Role of Small Food Stores in an Urban Environment." Public health nutrition 11, no. 4: 413-420.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen. 2011. "Calorie Posting in Chain Restaurants." American Economic Journal: Economic Policy 3, no. 1: 91-128.

Booth, Katie M, Megan M Pinkston, and Walker S Carlos Poston. 2005. "Obesity and the Built Environment." Journal of the American Dietetic Association 105, no. 5: 110-117.

Bostick, Roberd M, et al. 1994. "Sugar, Meat, and Fat Intake, and Non-Dietary Risk Factors for Colon Cancer Incidence in Iowa Women (United States)." Cancer Causes & Control 5, no. 1: 38-52.

Bovell-Benjamin, AC, et al. 2009. "Healthy Food Choices and Physical Activity Opportunities in Two Contrasting Alabama Cities." Health & place 15, no. 2: 429-438.

Bowman, Shanthy A. 2006. "A Comparison of the Socioeconomic Characteristics, Dietary Practices, and Health Status of Women Food Shoppers with Different Food Price Attitudes." Nutrition research 26, no. 7: 318-324.

Briefel, Ronette R, Ander Wilson, and Philip M Gleason. 2009. "Consumption of Low-Nutrient, Energy-Dense Foods and Beverages at School, Home, and Other Locations among School Lunch Participants and Nonparticipants." Journal of the American Dietetic Association 109, no. 2: S79-S90.

Broda, Christian, Ephraim Leibtag, and David E Weinstein. 2009. "The Role of Prices in Measuring the Poor's Living Standards." Journal of economic Perspectives 23, no. 2: 77-97.

Brown, Barbara B, Douglas D Perkins, and Graham Brown. 2004. "Incivilities, Place Attachment and Crime: Block and Individual Effects." Journal of environmental psychology 24, no. 3: 359-371.

Brown, CD, KA Donato, and E Obarzanek. 1998. "Body Mass Index and Prevalence of Risk Factors for Cardiovascular Disease." Obes Res.

Brown, Cheryl, and Stacy Miller. 2008. "The Impacts of Local Markets: A Review of Research on Farmers Markets and Community Supported Agriculture (Csa)." American Journal of Agricultural Economics 90, no. 5: 1298-1302.

Brown, Susan L. 2011. Using a Social-Ecological Model to Examine Obesity Interventions: Iowa State University.

Brown, Timothy A. 2014. Confirmatory Factor Analysis for Applied Research: Guilford Publications.

Canto, Amber, Laura E Brown, and Steven C Deller. 2014. "Rural Poverty, Food Access, and Public Health Outcomes." Choices 29, no. 2: 1-5.

Carbonneau, Elise, et al. 2017. "Development and Validation of the Perceived Food Environment Questionnaire in a French-Canadian Population." Public health nutrition 20, no. 11: 1914-1920.

Carr, Deborah, and Michael A Friedman. 2005. "Is Obesity Stigmatizing? Body Weight, Perceived Discrimination, and Psychological Well-Being in the United States." Journal of health and social behavior 46, no. 3: 244-259.

Cassady, Diana, Karen M Jetter, and Jennifer Culp. 2007. "Is Price a Barrier to Eating More Fruits and Vegetables for Low-Income Families?" Journal of the American Dietetic Association 107, no. 11: 1909-1915.

Census, United States. 2010 https://www.census.gov/2010census/data/.

Sustainable Food Center. 1996. "Access Denied." Accessed October 20 2017 https://940026988b3db2fff063-73e10890c944ea622176b82969dac6c6.ssl.cf2.rackcdn.com/31f82d56f07244438033bd7325f25306.pdf.

---. 2011. "Central Texas Foodshed Assessment " Accessed October 20 2017. https://940026988b3db2fff063-73e10890c944ea622176b82969dac6c6.ssl.cf2.rackcdn.com/52eafe6c1adf40a1b6f0043a0190bce1.pdf.

Chan, June M, et al. 1994. "Obesity, Fat Distribution, and Weight Gain as Risk Factors for Clinical Diabetes in Men." Diabetes care 17, no. 9: 961-969.

Charreire, Hélene, et al. 2010. "Measuring the Food Environment Using Geographical Information Systems: A Methodological Review." Public health nutrition 13, no. 11: 1773-1785.

Chen, Xiang, and Jill Clark. 2013. "Interactive Three-Dimensional Geovisualization of Space–Time Access to Food." Applied Geography 43: 81-86.

---. 2016. "Measuring Space–Time Access to Food Retailers: A Case of Temporal Access Disparity in Franklin County, Ohio." The Professional Geographer 68, no. 2: 175-188.

Chung, Chanjin, and Samuel L Myers. 1999. "Do the Poor Pay More for Food? An Analysis of Grocery Store Availability and Food Price Disparities." Journal of consumer affairs 33, no. 2: 276-296.

Chute, Christopher G, et al. 1991. "A Prospective Study of Body Mass, Height, and Smoking on the Risk of Colorectal Cancer in Women." Cancer Causes & Control 2, no. 2: 117-124.

Clarke, Graham, Heather Eyre, and Cliff Guy. 2002. "Deriving Indicators of Access to Food Retail Provision in British Cities: Studies of Cardiff, Leeds and Bradford." Urban Studies 39, no. 11: 2041-2060.

Colditz, G. A. 1992. "Economic Costs of Obesity." Am J Clin Nutr 55, no. 2 Suppl (Feb): 503s-507s.

Centers for Disease Control and Prevention. 2009. "The Social-Ecological Model: A Framework for Prevention." Injury center: Violence prevention.

Cooke, Rachel, and Angeliki Papadaki. 2014. "Nutrition Label Use Mediates the Positive Relationship between Nutrition Knowledge and Attitudes Towards Healthy Eating with Dietary Quality among University Students in the Uk." Appetite 83: 297-303.

Cooksey-Stowers, Kristen, Marlene B Schwartz, and Kelly D Brownell. 2017. "Food Swamps Predict Obesity Rates Better Than Food Deserts in the United States." International journal of environmental research and public health 14, no. 11: 1366.

Coombs, Nick, et al. 2010. A Spatial and Social Analysis of Food Deserts and Community Gardens in Madison, Wisconsin: University of Wisconsin-Madison.

Coulter, Sara Denise. 2009. "Access to Healthy and Less Healthy Food Options in a Low-Income, Racially Diverse Seattle Neighborhood." University of Washington.

Council, National Research. 2006. Food Insecurity and Hunger in the United States: An Assessment of the Measure: National Academies Press.

Coveney, John, and Lisel A O'Dwyer. 2009. "Effects of Mobility and Location on Food Access." Health & place 15, no. 1: 45-55.

Cummins, Steven, and Sally Macintyre. 2002. "" Food Deserts"-Evidence and Assumption in Health Policy Making." British medical journal 325, no. 7361: 436-438.

Cummins, Steven, et al. 2009. "Variations in Fresh Fruit and Vegetable Quality by Store Type, Urban–Rural Setting and Neighbourhood Deprivation in Scotland." Public health nutrition 12, no. 11: 2044-2050.

D'Acosta, Jenora. 2015. "Finding Food Deserts: A Study of Food Access Measures in the Phoenix-Mesa Urban Area." University of Southern California.

Dai, Dajun. 2010. "Black Residential Segregation, Disparities in Spatial Access to Health Care Facilities, and Late-Stage Breast Cancer Diagnosis in Metropolitan Detroit." Health & place 16, no. 5: 1038-1052.

Dai, Dajun, and Fahui Wang. 2011. "Geographic Disparities in Accessibility to Food Stores in Southwest Mississippi." Environment and Planning B: Planning and Design 38, no. 4: 659-677.

Donkin, Angela JM, et al. 1999. "Mapping Access to Food at a Local Level." British Food Journal 101, no. 7: 554-564.

Dowler, Elizabeth, et al. 2001. Measuring Access to Healthy Food in Sandwell: Sandwell Health Action Zone.

Drewnowski, Adam. 2009. "Obesity, Diets, and Social Inequalities." Nutrition reviews 67, no. s1.

Drewnowski, Adam, and SE Specter. 2004. "Poverty and Obesity: The Role of Energy Density and Energy Costs." The American journal of clinical nutrition 79, no. 1: 6-16.

Farah, Hodan, and Jean Buzby. 2005. "Us Food Consumption up 16 Percent since 1970." Amber Waves 3, no. 5: 5.

Feng, Jing, et al. 2010. "The Built Environment and Obesity: A Systematic Review of the Epidemiologic Evidence." Health & place 16, no. 2: 175-190.

Flegal, Katherine M, et al. 2002. "Prevalence and Trends in Obesity among Us Adults, 1999-2000." Jama 288, no. 14: 1723-1727.

Fleischhacker, Sheila E, et al. 2011. "A Systematic Review of Fast Food Access Studies." Obesity reviews 12, no. 5: e460-e471.

Ford, Mary Margaret, and Linda D Highfield. 2016. "Exploring the Spatial Association between Social Deprivation and Cardiovascular Disease Mortality at the Neighborhood Level." PloS one 11, no. 1: e0146085.

Forsyth, Ann, Leslie Lytle, and David Van Riper. 2010. "Finding Food: Issues and Challenges in Using Geographic Information Systems to Measure Food Access." Journal of transport and land use 3, no. 1: 43.

Franco, Manuel, et al. 2008. "Neighborhood Characteristics and Availability of Healthy Foods in Baltimore." American journal of preventive medicine 35, no. 6: 561-567.

Frank, Lawrence D, Martin A Andresen, and Thomas L Schmid. 2004. "Obesity Relationships with Community Design, Physical Activity, and Time Spent in Cars." American journal of preventive medicine 27, no. 2: 87-96.

Frank, Lawrence, et al. 2009. "Food Outlet Visits, Physical Activity and Body Weight: Variations by Gender and Race–Ethnicity." British Journal of Sports Medicine 43, no. 2: 124-131.

Frayling, Timothy M, et al. 2007. "A Common Variant in the Fto Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity." Science 316, no. 5826: 889-894.

French, Simone A, Lisa Harnack, and Robert W Jeffery. 2000. "Fast Food Restaurant Use among Women in the Pound of Prevention Study: Dietary, Behavioral and Demographic Correlates." International journal of obesity 24, no. 10: 1353.

Furey, Sinéad, Christopher Strugnell, and Ms Heather McIlveen. 2001. "An Investigation of the Potential Existence of``Food Deserts"in Rural and Urban Areas of Northern Ireland." Agriculture and Human Values 18, no. 4: 447-457.

Gallagher, Mari. 2006. Examining the Impact of Food Deserts on Public Health in Chicago: Chicago: Mari Gallagher Research and Consulting Group.

---. 2007. "Examining the Impact of Food Deserts on Public Health in Detroit." Chicago, IL: Mari Gallagher Research & Consulting Group.

Galvez, Maida P, et al. 2008. "Race and Food Store Availability in an Inner-City Neighbourhood." Public health nutrition 11, no. 6: 624-631.

Gans, Kim M, et al. 2010. "Peer Reviewed: Availability, Affordability, and Accessibility of a Healthful Diet in a Low-Income Community, Central Falls, Rhode Island, 2007-2008." Preventing chronic disease 7, no. 2.

Glanz, Karen, et al. 2005. "Healthy Nutrition Environments: Concepts and Measures." American journal of health promotion 19, no. 5: 330-333.

---. 2007. "Nutrition Environment Measures Survey in Stores (Nems-S): Development and Evaluation." American journal of preventive medicine 32, no. 4: 282-289.

Gloria, Christian T, and Mary A Steinhardt. 2010. "Texas Nutrition Environment Assessment of Retail Food Stores (Txnea-S): Development and Evaluation." Public health nutrition 13, no. 11: 1764-1772.

Gordon, Cynthia, et al. 2011. "Measuring Food Deserts in New York City's Low-Income Neighborhoods." Health & place 17, no. 2: 696-700.

Greenfield, Emily A. 2012. "Using Ecological Frameworks to Advance a Field of Research, Practice, and Policy on Aging-in-Place Initiatives." The Gerontologist 52, no. 1: 1-12.

Grzywacz, Joseph G, and Nadine F Marks. 2001. "Social Inequalities and Exercise During Adulthood: Toward an Ecological Perspective." Journal of Health and Social Behavior: 202-220.

Guagliardo, Mark F. 2004. "Spatial Accessibility of Primary Care: Concepts, Methods and Challenges." International journal of health geographics 3, no. 1: 3.

Gustafson, Alison A, et al. 2012. "Validation of Food Store Environment Secondary Data Source and the Role of Neighborhood Deprivation in Appalachia, Kentucky." BMC public health 12, no. 1: 688.

Gustafson, Alison, Scott Hankins, and Stephanie Jilcott. 2012. "Measures of the Consumer Food Store Environment: A Systematic Review of the Evidence 2000–2011." Journal of community health 37, no. 4: 897-911.

Haffner, Steven M, et al. 1991. "Greater Influence of Central Distribution of Adipose Tissue on Incidence of Non-Insulin-Dependent Diabetes in Women Than Men." The American journal of clinical nutrition 53, no. 5: 1312-1317.

Hager, Erin R, et al. 2017. "Food Swamps and Food Deserts in Baltimore City, Md, USA: Associations with Dietary Behaviours among Urban Adolescent Girls." Public health nutrition 20, no. 14: 2598-2607.

Haining, Robert, Stephen Wise, and Jingsheng Ma. 1998. "Exploratory Spatial Data Analysis." Journal of the Royal Statistical Society: Series D (The Statistician) 47, no. 3: 457-469.

Hallett IV, Lucius F, and Dave McDermott. 2011. "Quantifying the Extent and Cost of Food Deserts in Lawrence, Kansas, USA." Applied Geography 31, no. 4: 1210-1215.

Hamre, R, et al. 2014. Improving Nutrition, Physical Activity and Obesity Prevention: Performance Report of the Nutrition and Physical Activity Program to Prevent Obesity and Other Chronic Diseases: July 1 through December 31, 2005. Centers for Disease Control and Prevention, Rti International. 2006.

Han, Joan C, Debbie A Lawlor, and Sue YS Kimm. 2010. "Childhood Obesity." The Lancet 375, no. 9727: 1737-1748.

Handy, Susan L, and Debbie A Niemeier. 1997. "Measuring Accessibility: An Exploration of Issues and Alternatives." Environment and planning A 29, no. 7: 1175-1194.

Hargreaves, Margaret K, David G Schlundt, and Maciej S Buchowski. 2002. "Contextual Factors Influencing the Eating Behaviours of African American Women: A Focus Group Investigation." Ethnicity and Health 7, no. 3: 133-147.

Harrington, Daniel W, and Susan J Elliott. 2009. "Weighing the Importance of Neighbourhood: A Multilevel Exploration of the Determinants of Overweight and Obesity." Social Science & Medicine 68, no. 4: 593-600.

Harrison, Michelle S, et al. 2007. "The Increasing Cost of the Basic Foods Required to Promote Health in Queensland." Medical Journal of Australia 186, no. 1: 9.

Helling, Amy, and David S Sawicki. 2003. "Race and Residential Accessibility to Shopping and Services." Housing Policy Debate 14, no. 1-2: 69-101.

Hendrickson, Deja, Chery Smith, and Nicole Eikenberry. 2006. "Fruit and Vegetable Access in Four Low-Income Food Deserts Communities in Minnesota." Agriculture and Human Values 23, no. 3: 371-383.

Hilmers, Angela, David C Hilmers, and Jayna Dave. 2012. "Neighborhood Disparities in Access to Healthy Foods and Their Effects on Environmental Justice." American Journal of Public Health 102, no. 9: 1644-1654.

Hodgson, Kimberley. 2012. "Planning for Food Access and Community-Based Food System." US: American Planning Association.

Hogg, Robert Vincent, and Elliot A Tanis. 2009. Probability and Statistical Inference: Pearson Educational International.

Horowitz, Carol R, et al. 2004. "Barriers to Buying Healthy Foods for People with Diabetes: Evidence of Environmental Disparities." American Journal of Public Health 94, no. 9: 1549-1554.

Hsieh, Stephanie, et al. 2015. "Built Environment Associations with Adiposity Parameters among Overweight and Obese Hispanic Youth." Preventive medicine reports 2: 406-412.

Huang, Terry T-K, and Thomas A Glass. 2008. "Transforming Research Strategies for Understanding and Preventing Obesity." Jama 300, no. 15: 1811-1813.

Iceland, John, and Erika Steinmetz. 2003. "The Effects of Using Census Block Groups Instead of Census Tracts When Examining Residential Housing Patterns." US Census Bureau, Washington, DC.

Ikram, Samina Z, Yujie Hu, and Fahui Wang. 2015. "Disparities in Spatial Accessibility of Pharmacies in Baton Rouge, Louisiana." Geographical Review 105, no. 4: 492-510.

James, Delores. 2004. "Factors Influencing Food Choices, Dietary Intake, and Nutrition-Related Attitudes among African Americans: Application of a Culturally Sensitive Model." Ethnicity and Health 9, no. 4: 349-367.

Jetter, Karen M, and Diana L Cassady. 2006. "The Availability and Cost of Healthier Food Alternatives." American journal of preventive medicine 30, no. 1: 38-44.

Jilcott, Stephanie B, et al. 2009. "Perceptions of the Community Food Environment and Related Influences on Food Choice among Midlife Women Residing in Rural and Urban Areas: A Qualitative Analysis." Women & health 49, no. 2-3: 164-180.

Joseph, Alun E, and Peter R Bantock. 1982. "Measuring Potential Physical Accessibility to General Practitioners in Rural Areas: A Method and Case Study." Social science & medicine 16, no. 1: 85-90.

Judge, Timothy A, and Daniel M Cable. 2011. "When It Comes to Pay, Do the Thin Win? The Effect of Weight on Pay for Men and Women." Journal of Applied Psychology 96, no. 1: 95.

Kaufman, Phil R. 1997. "Do the Poor Pay More for Food?: Item Selection and Price Differences Affect Low-Income Household Food Costs."

Kestens, Yan, and Mark Daniel. 2010. "Social Inequalities in Food Exposure around Schools in an Urban Area." American journal of preventive medicine 39, no. 1: 33-40.

Kestens, Yan, et al. 2012. "Association between Activity Space Exposure to Food Establishments and Individual Risk of Overweight." PloS one 7, no. 8: e41418.

Kestens, Yan, et al. 2010. "Using Experienced Activity Spaces to Measure Foodscape Exposure." Health & Place 16, no. 6: 1094-1103.

Kim, Hae-Young. 2013. "Statistical Notes for Clinical Researchers: Assessing Normal Distribution (2) Using Skewness and Kurtosis." Restorative dentistry & endodontics 38, no. 1: 52-54.

Kremers, Stef PJ, et al. 2006. "Environmental Influences on Energy Balance-Related Behaviors: A Dual-Process View." International Journal of Behavioral Nutrition and Physical Activity 3, no. 1: 9.

Krukowski, Rebecca A, et al. 2010. "Neighborhood Impact on Healthy Food Availability and Pricing in Food Stores." Journal of community health 35, no. 3: 315-320.

Kuai, Xuan, and Qunshan Zhao. 2017. "Examining Healthy Food Accessibility and Disparity in Baton Rouge, Louisiana." Annals of GIS 23, no. 2: 103-116.

Kumanyika, Shiriki, et al. 2007. "Expanding the Obesity Research Paradigm to Reach African American Communities." Preventing chronic disease 4, no. 4.

Kwate, Naa Oyo A, et al. 2009. "Inequality in Obesigenic Environments: Fast Food Density in New York City." Health & place 15, no. 1: 364-373.

Lamichhane, Archana P, et al. 2013. "Spatial Patterning of Supermarkets and Fast Food Outlets with Respect to Neighborhood Characteristics." Health & place 23: 157-164.

Langford, Mitchel, and Gary Higgs. 2006. "Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population." The Professional Geographer 58, no. 3: 294-306.

Laraia, Barbara A, et al. 2006. "Psychosocial Factors and Socioeconomic Indicators Are Associated with Household Food Insecurity among Pregnant Women." The Journal of nutrition 136, no. 1: 177-182.

Larsen, Kristian, and Jason Gilliland. 2008. "Mapping the Evolution Of'food Deserts' in a Canadian City: Supermarket Accessibility in London, Ontario, 1961–2005." International Journal of Health Geographics 7, no. 1: 16.

Larson, Nicole I, Mary T Story, and Melissa C Nelson. 2009. "Neighborhood Environments: Disparities in Access to Healthy Foods in the Us." American journal of preventive medicine 36, no. 1: 74-81. e10.

Lawrence, Wendy, et al. 2009. "Why Women of Lower Educational Attainment Struggle to Make Healthier Food Choices: The Importance of Psychological and Social Factors." Psychology and Health 24, no. 9: 1003-1020.

Leibtag, Ephraim S, and Phillip R Kaufman. 2003. Exploring Food Purchase Behavior of Low-Income Households: How Do They Economize?: United States Department of Agriculture, Economic Research Service.

Leone, Angela F, et al. 2011. "Peer Reviewed: Store Type and Demographic Influence on the Availability and Price of Healthful Foods, Leon County, Florida, 2008." Preventing chronic disease 8, no. 6.

Li, Jingjing, and Changjoo Kim. 2018. "Measuring Individuals' Spatial Access to Healthy Foods by Incorporating Mobility, Time, and Mode: Activity Space Measures." The Professional Geographer 70, no. 2: 198-208.

Ling, N. 2005. "A Comparison of Prices For'healthy'and'less Healthy'food Baskets in Contrasting Neighbourhoods." University of O tago.

Lisabeth, Lynda D, et al. 2010. "The Food Environment in an Urban Mexican American Community." Health & place 16, no. 3: 598-605.

Lopez, Russ P. 2007. "Neighborhood Risk Factors for Obesity." Obesity 15, no. 8: 2111-2119.

Lounsbury, David William, and Shannon Gwin Mitchell. 2009. "Introduction to Special Issue on Social Ecological Approaches to Community Health Research and Action." American journal of community psychology 44, no. 3-4: 213-220.

Luan, Hui. 2016. "Spatial and Spatio-Temporal Analyses of Neighborhood Retail Food Environments: Evidence for Food Planning and Interventions."

Lucan, Sean C, et al. 2015. "Urban Farmers' Markets: Accessibility, Offerings, and Produce Variety, Quality, and Price Compared to Nearby Stores." Appetite 90: 23-30.

Luo, Jun. 2014. "Integrating the Huff Model and Floating Catchment Area Methods to Analyze Spatial Access to Healthcare Services." Transactions in GIS 18, no. 3: 436-448.

Luo, Wei, and Yi Qi. 2009. "An Enhanced Two-Step Floating Catchment Area (E2sfca) Method for Measuring Spatial Accessibility to Primary Care Physicians." Health & place 15, no. 4: 1100-1107.

Luo, Wei, and Fahui Wang. 2003. "Measures of Spatial Accessibility to Health Care in a Gis Environment: Synthesis and a Case Study in the Chicago Region." Environment and Planning B: Planning and Design 30, no. 6: 865-884.

Luo, Wei, and Tara Whippo. 2012. "Variable Catchment Sizes for the Two-Step Floating Catchment Area (2sfca) Method." Health & place 18, no. 4: 789-795.

Lytle, Leslie A. 2009. "Measuring the Food Environment: State of the Science." American journal of preventive medicine 36, no. 4: S134-S144.

Ma, Ting, and Gerrit Jan-Knaap. 2014. Analyzing Employment Accessibility in a Multimodal Network Using Gtfs: A Demonstration of the Purple Line, Maryland. The Association of Collegiate Schools of Planning (acsp) Annual Conference, Philadelphia, Pennsylvania.

Macdonald, Laura, et al. 2011. "Is Proximity to a Food Retail Store Associated with Diet and Bmi in Glasgow, Scotland?" BMC public health 11, no. 1: 464.

Macintyre, Sally. 2007. "Deprivation Amplification Revisited; or, Is It Always True That Poorer Places Have Poorer Access to Resources for Healthy Diets and Physical Activity?" International Journal of Behavioral Nutrition and Physical Activity 4, no. 1: 32.

Macintyre, Sally, Laura Macdonald, and Anne Ellaway. 2008. "Do Poorer People Have Poorer Access to Local Resources and Facilities? The Distribution of Local Resources by Area Deprivation in Glasgow, Scotland." Social science & medicine 67, no. 6: 900-914.

Mao, Liang, and Dawn Nekorchuk. 2013. "Measuring Spatial Accessibility to Healthcare for Populations with Multiple Transportation Modes." Health & place 24: 115-122.

Matheson, Flora I, et al. 2012. "Development of the Canadian Marginalization Index: A New Tool for the Study of Inequality." Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique: S12-S16.

McGinnis, J Michael, and William H Foege. 1993. "Actual Causes of Death in the United States." Jama 270, no. 18: 2207-2212.

McGrail, Matthew R, and John S Humphreys. 2009. "The Index of Rural Access: An Innovative Integrated Approach for Measuring Primary Care Access." BMC Health Services Research 9, no. 1: 124.

---. 2014. "Measuring Spatial Accessibility to Primary Health Care Services: Utilising Dynamic Catchment Sizes." Applied Geography 54: 182-188.

McKenzie, Brian S. 2014. "Access to Supermarkets among Poorer Neighborhoods: A Comparison of Time and Distance Measures." Urban Geography 35, no. 1: 133-151.

Meyers, Lawrence S, Glenn C Gamst, and AJ Guarino. 2013. Performing Data Analysis Using Ibm Spss: John Wiley & Sons.

Michimi, Akihiko, and Michael C Wimberly. 2010. "Associations of Supermarket Accessibility with Obesity and Fruit and Vegetable Consumption in the Conterminous United States." International Journal of Health Geographics 9, no. 1: 49.

Milicic, Sandra, and Philip DeCicca. 2017. "The Impact of Economic Conditions on Healthy Dietary Intake: Evidence from Fluctuations in Canadian Unemployment Rates." Journal of nutrition education and behavior 49, no. 8: 632-638. e1.

Minaker, L. 2013. "Measuring the Food Environment in Canada." Ottawa, Ontario: Health Canada.

Minaker, Leia Michelle. 2013. Evaluating Food Environment Assessment Methodologies: A Multi-Level Examination of Associations between Food Environments and Individual Outcomes: University of Alberta (Canada).

Mobley, Lee R, et al. 2006. "Environment, Obesity, and Cardiovascular Disease Risk in Low-Income Women." American journal of preventive medicine 30, no. 4: 327-332. e1.

Moore, Latetia V, and Ana V Diez Roux. 2006. "Associations of Neighborhood Characteristics with the Location and Type of Food Stores." American journal of public health 96, no. 2: 325-331.

Moore, Latetia V, et al. 2008. "Associations of the Local Food Environment with Diet Quality—a Comparison of Assessments Based on Surveys and Geographic Information Systems: The Multi-Ethnic Study of Atherosclerosis." American journal of epidemiology 167, no. 8: 917-924.

Moore, Latetia V, Ana V Diez Roux, and Shannon Brines. 2008. "Comparing Perception-Based and Geographic Information System (Gis)-Based Characterizations of the Local Food Environment." Journal of Urban Health 85, no. 2: 206-216.

Morland, Kimberly B, and Kelly R Evenson. 2009. "Obesity Prevalence and the Local Food Environment." Health & place 15, no. 2: 491-495.

Morland, Kimberly, Ana V Diez Roux, and Steve Wing. 2006. "Supermarkets, Other Food Stores, and Obesity: The Atherosclerosis Risk in Communities Study." American journal of preventive medicine 30, no. 4: 333-339.

Morland, Kimberly, et al. 2002. "Neighborhood Characteristics Associated with the Location of Food Stores and Food Service Places." American journal of preventive medicine 22, no. 1: 23-29.

Morton, Lois Wright, and Troy C Blanchard. 2007. "Starved for Access: Life in Rural America's Food Deserts." Rural Realities 1, no. 4: 1-10.

Mulrooney, Timothy, et al. 2017. "A Comparison of Raster-Based Travel Time Surfaces against Vector-Based Network Calculations as Applied in the Study of Rural Food Deserts." Applied Geography 78: 12-21.

Mushi-Brunt, Christina, et al. 2007. "Fruit and Vegetable Intake and Obesity in Preadolescent Children: The Role of Neighborhood Poverty and Grocery Store Access." American Journal of Health Education 38, no. 5: 258-265.

Nakaya, T, et al. 2017. Gwr4 User Manual. 2014: GWR.

Neff, Roni A, et al. 2009. "Food Systems and Public Health Disparities." Journal of Hunger & Environmental Nutrition 4, no. 3-4: 282-314.

Neuhouser, Marian L, Alan R Kristal, and Ruth E Patterson. 1999. "Use of Food Nutrition Labels Is Associated with Lower Fat Intake." Journal of the American dietetic Association 99, no. 1: 45-53.

Ngui, André Ngamini, and Philippe Apparicio. 2011. "Optimizing the Two-Step Floating Catchment Area Method for Measuring Spatial Accessibility to Medical Clinics in Montreal." BMC health services research 11, no. 1: 166.

O'Dwyer, Lisel A, and Deborah L Burton. 1998. "Potential Meets Reality: Gis and Public Health Research in Australia." Australian and New Zealand Journal of Public Health 22, no. 7: 819-823.

O'Dwyer, Lisel A, and John Coveney. 2006. "Scoping Supermarket Availability and Accessibility by Socio-Economic Status in Adelaide." Health Promotion Journal of Australia 17, no. 3: 240-246.

Ollberding, Nicholas Jay, Randi L Wolf, and Isobel Contento. 2011. "Food Label Use and Its Relation to Dietary Intake among Us Adults." Journal of the American Dietetic Association 111, no. 5: S47-S51.

Opfer, Pamela R. 2010. "Using Gis Technology to Identify and Analyze 'Food Deserts' on the Southern Oregon Coast."

Food and Agriculture Organization. 2006. "Food Security." FAO Policy Brief Accessed October 25 2017. http://www.fao.org/forestry/13128-0e6f36f27e0091055bec28ebe830f46b3.pdf.

World Health Organization, 2017. "Information Sheet on Obesity and Overweight." Accessed October 1 2017. http://www.who.int/mediacentre/factsheets/fs311/en/.

Ossiander, Eric M, et al. 2004. "Driver's Licenses as a Source of Data on Height and Weight." Economics & Human Biology 2, no. 2: 219-227.

Palermo, CE, et al. 2008. "The Cost of Healthy Food in Rural Victoria." Rural Remote Health 8, no. 4: 1074.

Papas, Mia A, et al. 2007. "The Built Environment and Obesity." Epidemiologic reviews 29, no. 1: 129-143.

Pearce, Jamie, et al. 2007. "Neighborhood Deprivation and Access to Fast-Food Retailing: A National Study." American journal of preventive medicine 32, no. 5: 375-382.

Pearce, Jamie, Peter Day, and Karen Witten. 2008. "Neighbourhood Provision of Food and Alcohol Retailing and Social Deprivation in Urban New Zealand." Urban Policy and Research 26, no. 2: 213-227.

Pearce, Jamie, Karen Witten, and Phil Bartie. 2006. "Neighbourhoods and Health: A Gis Approach to Measuring Community Resource Accessibility." Journal of Epidemiology & Community Health 60, no. 5: 389-395.

Pedraza Sanchez, Lauramaria. 2015. "A Farmers' Market in a Food Desert: Evaluating Walkability and Streetscape as Factors of Farmers' Market Effectiveness in Food Accessibility: The Case of Farmers' Market East in Austin, Texas."

Peng, Zhong-Ren. 1997. "The Jobs-Housing Balance and Urban Commuting." Urban studies 34, no. 8: 1215-1235.

Piontak, Joy Rayanne. 2013. Childhood Obesity and Place: Poverty, Race, and Food Access: North Carolina State University.

Powell, Lisa M, Frank J Chaloupka, and Yanjun Bao. 2007. "The Availability of Fast-Food and Full-Service Restaurants in the United States: Associations with Neighborhood Characteristics." American journal of preventive medicine 33, no. 4: S240-S245.

Powell, Lisa M, and Binh T Nguyen. 2013. "Fast-Food and Full-Service Restaurant Consumption among Children and Adolescents: Effect on Energy, Beverage, and Nutrient Intake." JAMA pediatrics 167, no. 1: 14-20.

Powell, Lisa M, et al. 2007. "Food Store Availability and Neighborhood Characteristics in the United States." Preventive medicine 44, no. 3: 189-195.

Power, Michael L, and Jay Schulkin. 2008. "Sex Differences in Fat Storage, Fat Metabolism, and the Health Risks from Obesity: Possible Evolutionary Origins." British Journal of Nutrition 99, no. 5: 931-940.

Raja, Samina, Changxing Ma, and Pavan Yadav. 2008. "Beyond Food Deserts: Measuring and Mapping Racial Disparities in Neighborhood Food Environments." Journal of Planning Education and Research 27, no. 4: 469-482.

Richard, Lucie, Lise Gauvin, and Kim Raine. 2011. "Ecological Models Revisited: Their Uses and Evolution in Health Promotion over Two Decades." Annual review of public health 32: 307-326.

Robinson, Tanya. 2008. "Applying the Socio-Ecological Model to Improving Fruit and Vegetable Intake among Low-Income African Americans." Journal of community health 33, no. 6: 395-406.

Rose, Donald, et al. 2009. "Neighborhood Food Environments and Body Mass Index: The Importance of in-Store Contents." American journal of preventive medicine 37, no. 3: 214-219.

Rundle, Andrew, et al. 2009. "Neighborhood Food Environment and Walkability Predict Obesity in New York City." Environmental health perspectives 117, no. 3: 442.

Russell, Scott E, and C Patrick Heidkamp. 2011. "'Food Desertification': The Loss of a Major Supermarket in New Haven, Connecticut." Applied Geography 31, no. 4: 1197-1209.

Saelens, Brian E, et al. 2007. "Nutrition Environment Measures Study in Restaurants (Nems-R): Development and Evaluation." American journal of preventive medicine 32, no. 4: 273-281.

Sallis, James F, and Karen Glanz. 2009. "Physical Activity and Food Environments: Solutions to the Obesity Epidemic." The Milbank Quarterly 87, no. 1: 123-154.

Sarrafzadegan, Nizal, et al. 2013. "Parental Perceptions of Weight Status of Their Children." ARYA atherosclerosis 9, no. 1: 61.

Satia, Jessie A, Joseph A Galanko, and Marian L Neuhouser. 2005. "Food Nutrition Label Use Is Associated with Demographic, Behavioral, and Psychosocial Factors and Dietary Intake among African Americans in North Carolina." Journal of the American Dietetic Association 105, no. 3: 392-402.

Schlundt, D. 2014. Selection of Food Deserts for Nashville's Communities Putting Prevention to Work.

Shannon, Jerry. 2016. "Beyond the Supermarket Solution: Linking Food Deserts, Neighborhood Context, and Everyday Mobility." Annals of the American Association of Geographers 106, no. 1: 186-202.

Sharkey, Joseph R, Scott Horel, and Wesley R Dean. 2010. "Neighborhood Deprivation, Vehicle Ownership, and Potential Spatial Access to a Variety of Fruits and Vegetables in a Large Rural Area in Texas." International Journal of Health Geographics 9, no. 1: 26.

Shaw, Hillary J. 2006. "Food Deserts: Towards the Development of a Classification." Geografiska Annaler: Series B, Human Geography 88, no. 2: 231-247.

Sherrill, KR, B Frakes, and S Schupbach. 2010. "Travel Time Cost Surface Model: Standard Operating Procedure." Natural Resource Report. Nps/Nrpc/Imd/Nrr–2010/238. Natural Resources Program Center, Fort Collins, Colorado. Published Report-2164894.

Shi, Xun, et al. 2012. "Spatial Access and Local Demand for Major Cancer Care Facilities in the United States." Annals of the Association of American Geographers 102, no. 5: 1125-1134.

Slater, Joyce, et al. 2009. "The Growing Canadian Energy Gap: More the Can Than the Couch?" Public health nutrition 12, no. 11: 2216-2224.

Smed, Sinne, et al. 2018. "The Consequences of Unemployment on Diet Composition and Purchase Behaviour: A Longitudinal Study from Denmark." Public health nutrition 21, no. 3: 580-592.

Smith, Adam B, et al. 2002. "Factor Analysis of the Hospital Anxiety and Depression Scale from a Large Cancer Population." Psychology and Psychotherapy: Theory, Research and Practice 75, no. 2: 165-176.

Smoyer-Tomic, Karen E, et al. 2008. "The Association between Neighborhood Socioeconomic Status and Exposure to Supermarkets and Fast Food Outlets." Health & place 14, no. 4: 740-754.

Smoyer - Tomic, Karen E, John C Spence, and Carl Amrhein. 2006. "Food Deserts in the Prairies? Supermarket Accessibility and Neighborhood Need in Edmonton, Canada." The Professional Geographer 58, no. 3: 307-326.

Sparks, Andrea L, Neil Bania, and Laura Leete. 2011. "Comparative Approaches to Measuring Food Access in Urban Areas: The Case of Portland, Oregon." Urban Studies 48, no. 8: 1715-1737.

Spence, John C, et al. 2009. "Relation between Local Food Environments and Obesity among Adults." BMC public health 9, no. 1: 192.

Stamler, Rose, et al. 1978. "Weight and Blood Pressure: Findings in Hypertension Screening of 1 Million Americans." Jama 240, no. 15: 1607-1610.

Stein, Dana Oppenheim. 2011. 'Food Deserts' and 'Food Swamps' in Hillsborough County, Florida: Unequal Access to Supermarkets and Fast-Food Restaurants: University of South Florida.

Story, Mary, et al. 2008. "Creating Healthy Food and Eating Environments: Policy and Environmental Approaches." Annu. Rev. Public Health 29: 253-272.

Strauss, Richard. 2002. "Perspectives on Childhood Obesity." Current gastroenterology reports 4, no. 3: 244-250.

American Community Survey. 2016. Accessed November 11 2017. https://www.census.gov/programs-surveys/acs/.

Swinburn, Boyd, Garry Egger, and Fezeela Raza. 1999. "Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity." Preventive medicine 29, no. 6: 563-570.

Thornton, Rachel LJ, et al. 2016. "Evaluating Strategies for Reducing Health Disparities by Addressing the Social Determinants of Health." Health Affairs 35, no. 8: 1416-1423.

Thow, Anne Marie, et al. 2011. "Taxing Soft Drinks in the Pacific: Implementation Lessons for Improving Health." Health Promotion International 26, no. 1: 55-64.

Tobler, Waldo R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." Economic geography 46, no. sup1: 234-240.

Turrell, Gavin, et al. 2002. "Socioeconomic Differences in Food Purchasing Behaviour and Suggested Implications for Diet - Related Health Promotion." Journal of Human Nutrition and Dietetics 15, no. 5: 355-364.

Vallianatos, Mark, et al. 2010. "Peer Reviewed: Food Access, Availability, and Affordability in 3 Los Angeles Communities, Project Cafe, 2004-2006." Preventing chronic disease 7, no. 2.

Van Meter, Emily M, et al. 2010. "An Evaluation of Edge Effects in Nutritional Accessibility and Availability Measures: A Simulation Study." International journal of health geographics 9, no. 1: 40.

Variyam, Jayachandran N, and John Cawley. 2006. Nutrition Labels and Obesity: National Bureau of Economic Research.

Ver Ploeg, Michele. 2010. Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences: Report to Congress: Diane Publishing.

Vo, Au, Miloslava Plachkinova, and Rahul Bhaskar. 2015. "Assessing Healthcare Accessibility Algorithms: A Comprehensive Investigation of Two-Step Floating Catchment Methodologies Family."

Walker, Renee E, Jason Block, and Ichiro Kawachi. 2012. "Do Residents of Food Deserts Express Different Food Buying Preferences Compared to Residents of Food Oases? A Mixed-Methods Analysis." International Journal of Behavioral Nutrition and Physical Activity 9, no. 1: 41.

Walker, Renee E, Christopher R Keane, and Jessica G Burke. 2010. "Disparities and Access to Healthy Food in the United States: A Review of Food Deserts Literature." Health & place 16, no. 5: 876-884.

Walker, Susan P, et al. 1996. "Body Size and Fat Distribution as Predictors of Stroke among Us Men." American Journal of Epidemiology 144, no. 12: 1143-1150.

Wallace, Richard, et al. 2005. "Access to Health Care and Nonemergency Medical Transportation: Two Missing Links." Transportation Research Record: journal of the transportation research board, no. 1924: 76-84.

Wan, Neng, et al. 2012. "A Relative Spatial Access Assessment Approach for Analyzing Potential Spatial Access to Colorectal Cancer Services in Texas." Applied Geography 32, no. 2: 291-299.

Wan, Neng, et al. 2013. "Spatial Access to Health Care Services and Disparities in Colorectal Cancer Stage at Diagnosis in Texas." The Professional Geographer 65, no. 3: 527-541.

Wan, Neng, Bin Zou, and Troy Sternberg. 2012. "A Three-Step Floating Catchment Area Method for Analyzing Spatial Access to Health Services." International Journal of Geographical Information Science 26, no. 6: 1073-1089.

Wang, Fahui. 2000. "Modeling Commuting Patterns in Chicago in a Gis Environment: A Job Accessibility Perspective." The Professional Geographer 52, no. 1: 120-133.

---. 2014. Quantitative Methods and Socio-Economic Applications in Gis: CRC Press.

Wang, Fahui, and Wei Luo. 2005. "Assessing Spatial and Nonspatial Factors for Healthcare Access: Towards an Integrated Approach to Defining Health Professional Shortage Areas." Health & place 11, no. 2: 131-146.

Wang, Haoluan, et al. 2016. "The Role of Socio-Economic Status and Spatial Effects on Fresh Food Access: Two Case Studies in Canada." Applied Geography 67: 27-38.

Wang, May C, et al. 2007. "Socioeconomic and Food-Related Physical Characteristics of the Neighbourhood Environment Are Associated with Body Mass Index." Journal of Epidemiology & Community Health 61, no. 6: 491-498.

Weiss, Linda, et al. 2007. "Defining Neighborhood Boundaries for Urban Health Research." American journal of preventive medicine 32, no. 6: S154-S159.

Widener, Michael J, et al. 2015. "Spatiotemporal Accessibility to Supermarkets Using Public Transit: An Interaction Potential Approach in Cincinnati, Ohio." Journal of Transport Geography 42: 72-83.

Widener, Michael J, et al. 2013. "Using Urban Commuting Data to Calculate a Spatiotemporal Accessibility Measure for Food Environment Studies." Health & place 21: 1-9.

Widener, Michael J, Sara S Metcalf, and Yaneer Bar-Yam. 2011. "Dynamic Urban Food Environments: A Temporal Analysis of Access to Healthy Foods." American journal of preventive medicine 41, no. 4: 439-441.

Wiig, Kristen, and Chery Smith. 2009. "The Art of Grocery Shopping on a Food Stamp Budget: Factors Influencing the Food Choices of Low-Income Women as They Try to Make Ends Meet." Public health nutrition 12, no. 10: 1726-1734.

Wilcox, Sara, et al. 2013. "Frequency of Consumption at Fast-Food Restaurants Is Associated with Dietary Intake in Overweight and Obese Women Recruited from Financially Disadvantaged Neighborhoods." Nutrition research 33, no. 8: 636-646.

Williden, Micalla, et al. 2006. "The Apple Project: An Investigation of the Barriers and Promoters of Healthy Eating and Physical Activity in New Zealand Children Aged 5-12 Years." Health Education Journal 65, no. 2: 135-148.

Wilson, William Julius. 1996. "When Work Disappears." Political Science Quarterly 111, no. 4: 567-595.

Winkler, Elisabeth, Gavin Turrell, and Carla Patterson. 2006. "Does Living in a Disadvantaged Area Mean Fewer Opportunities to Purchase Fresh Fruit and Vegetables in the Area? Findings from the Brisbane Food Study." Health & place 12, no. 3: 306-319.

Witten, Karen. 2016. Geographies of Obesity: Environmental Understandings of the Obesity Epidemic: Routledge.

Wong, David WS. 2004. "The Modifiable Areal Unit Problem (Maup)." In Worldminds: Geographical Perspectives on 100 Problems, 571-575: Springer.

Woodham, Carly Louise. 2011. "Food Desert or Food Swamp? An in-Depth Exploration of Neighbourhood Food Environments in Eastern Porirua and Whitby." University of Otago.

Wrigley, Neil, et al. 2002. "Assessing the Impact of Improved Retail Access on Diet in A'food Desert': A Preliminary Report." Urban Studies 39, no. 11: 2061-2082.

Xu, Yanqing. 2014. "Built Environment and Risk of Obesity in the United States: A Multilevel Modeling Approach."

Yang, Yong, and Ana V Diez-Roux. 2013. "Using an Agent-Based Model to Simulate Children’s Active Travel to School." International journal of behavioral nutrition and physical activity 10, no. 1: 67.

Yu, Mandi, et al. 2014. "Using a Composite Index of Socioeconomic Status to Investigate Health Disparities While Protecting the Confidentiality of Cancer Registry Data." Cancer Causes & Control 25, no. 1: 81-92.

Zadnik, Vesna, and BJ Reich. 2006. "Analysis of the Relationship between Socioeconomic Factors and Stomach Cancer Incidence in Slovenia." Neoplasma 53, no. 2: 103-110.

Zenk, Shannon N, et al. 2005a. "Fruit and Vegetable Intake in African Americans: Income and Store Characteristics." American journal of preventive medicine 29, no. 1: 1-9.

Zenk, Shannon N, et al. 2005b. "Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit." American journal of public health 95, no. 4: 660-667.

Zenk, Shannon N, et al. 2011. "Activity Space Environment and Dietary and Physical Activity Behaviors: A Pilot Study." Health & place 17, no. 5: 1150-1161.

Zenk, Shannon N, Amy J Schulz, and Angela Odoms-Young. 2009. "How Neighborhood Environments Contribute to Obesity." The American journal of nursing 109, no. 7: 61.

Zhang, Lianjun, Zhihai Ma, and Luo Guo. 2009. "An Evaluation of Spatial Autocorrelation and Heterogeneity in the Residuals of Six Regression Models." Forest Science 55, no. 6: 533-548.