SUPERMARKET ACCESS, CONSUMPTION OF FRESH PRODUCE, AND WEIGHT STATUS FOR U.S. ADULTS

By

KAITLIN DIPAOLA

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Food and Business Economics

Written under the direction of

Dr. Yanhong Jin

And approved by

_____

_____

_____

New Brunswick, New Jersey

May 2016

ABSTRACT OF THE THESIS

Supermarket Access, Consumption of Fresh Produce, and Weight Status for U.S. Adults

By KAITLIN DIPAOLA

Thesis Director:
Dr. Yanhong Jin

As overweight and obesity have reached pandemic levels in the United States, the local food environment has been found to play an integral role. The relationship between supermarket access and weight status has gained traction from both Academia and government sectors, in a hope to lead Americans toward a healthier lifestyle. The objective of this thesis is to examine the impact of supermarket access, in conjunction with important individual socioeconomic and neighborhood attributes, on a) the risk of being overweight or obese in American adults; and b) the probability of meeting the federally recommended level of fruits and vegetables.

Using BRFSS survey data, The Reinvestment Fund's Limited Supermarket Access (LSA) study, as well as data from the Census, the American Community Survey, and ESRI's Consumer Expenditures Survey, I find that supermarket access does have an impact on weight status and the probability of meeting recommended levels of fruits and vegetables. In particular, when controlling for the same percentage of land classified as limited supermarket access areas, individuals living in an area with one LSA area are more likely to be overweight or obese than when compared with multiple disjointed LSA areas. From an intervention standpoint, when considering resources, it can be incredibly resource-

intensive to eradicate all LSAs from neighborhoods. Perhaps though, if supermarkets are more strategically placed in neighborhoods, which may potentially provide more access to supermarkets to a greater number of individuals, and thus enhance the likelihood of meeting the recommended consumption of fruits and vegetables. On the other hand, as the number of LSAs increase, the risk of being overweight or obese also increases. Thus, decreasing the number of LSAs in a neighborhood will help to fight against the risk of overweight/obesity.

ACKNOWLEDGEMENTS

First and foremost I would like to express my sincere gratitude to my advisor, Dr. Yanhong Jin, who has tirelessly supported and worked with me to complete this thesis. I would further like to thank Dr. Carl Pray and Dr. Paul Gottlieb, both of whose guidance and expertise were critical to both my time as a student at Rutgers as well as for the completion of this thesis; this would not have been possible without their support and knowledge. An immense amount of gratitude must also be given to Mr. Jim Trimble of the Center for Remote Sensing and Spatial Analysis. Without his patience and ArcGIS knowledge, again, this thesis would not have been possible!

Finally, I would like to thank both my parents, and my future husband, Tim, whose love, support, patience, and encouragement was most instrumental in the finishing my master's degree.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter I
Introduction

1.1 A Brief Background

The severity of overweight and obesity as serious health issues has been increasing for the past few decades; one may even assess the situation as a worldwide pandemic. Both overweight and obesity do not discriminate against race, gender, or age — men, women, and children, regardless of race, are increasingly affected (Olshansky et al. 2005). Research on this topic, the association, correlation, and causation between the environment and diet and physical activity, and the subsequent health outcomes has been going on for decades now.

As early as 1999, The United States Surgeon General noted that an estimated 61 percent of adults were overweight or obese in the United States (U.S. Department of Health and Human Services, 2001). In 2013, the World Health Organization (WHO) listed obesity as one of twelve preventable risk factors that are associated with increased mortality and morbidity (WHO 2013). The latest obesity statistics from The Centers for Disease Control and Prevention (CDC) indicate that approximately one-third of adults (34.9%) and about 17 percent of children and adolescents are considered to be obese in the United States in 2012 (CDC: Division of Nutrition 2014).

Healthy People, released by the United States Department of Health and Human Services (USDHHS) each decade, has recently noted the goals and objectives of *Healthy People 2020* — the 4th iteration of this initiative. With regard to nutrition and weight status, the main goal is to promote health and reduce chronic disease risk through the consumption of healthful diets and achievement and maintenance of healthy body weights (U.S.

Department of Health and Human Services 2014). A healthful diet, as defined by the USDHHS, means consuming a variety of nutrient-dense foods within and across the food groups, especially whole grains, fruits, vegetables, low-fat or fat-free milk or milk products, and lean meats and other protein sources. The Dietary Guidelines for Americans, updated every five years, with the most recent one published in 2010, noted poor diet as one of the most important factors contributing to overweight and obesity in the United States (U.S. Department of Health and Human Services Services 2010). Individuals with healthy weight are found to be less likely to develop chronic diseases or complications during pregnancy, and are less likely to die at an earlier age (U.S. Department of Health and Human Services Services 2014).

Of economic importance is the financial burden that obesity places on individuals, as well as our nation. In 2006 dollars, the per capita annual medical cost on obese individuals was $1,429 higher than for a normal weight individual, a 42 percent increase over costs for normal weight individuals (Finkelstein et al. 2009). Furthermore, increased spending to combat obesity is applicable to children, as well, and this burden has the potential to carry on through adulthood, compounding over time.

Time and time again, obesity is regarded as something that is preventable. The CDC notes that healthy lifestyle habits, including healthy eating and physical activity, can lower the risk of becoming obese and developing related diseases (CDC 2015; WHO 2015, 2015). For this reason, continued research of the association with the local food environment is imperative. This thesis hopes to contribute to the ongoing body of

research with a focus on the neighborhood food environment and how it relates to obesity.

1.2 Purpose and Justification of Thesis

The overall objective of this thesis is to examine the impact of supermarket access, in conjunction with important individual socioeconomic and neighborhood attributes, on the risk of being overweight or obese in American adults, as well as how supermarket access affects the consumption level of both fruits and vegetables. As overweight and obesity has reached pandemic levels, continued research to pinpoint a direct cause is essential. Though there is ample research already conducted on this topic, a large majority of the literature is regional, or conducted within one state or even one county. Conversely, this thesis will use datasets that are nationally represented, an aspect that has not been explored quite as deeply yet.

1.3 Research Objectives

First and foremost, this thesis will seek to determine if there is a statistically significant association between living in an area with low access to supermarkets and the risk of being overweight or obese, and the likelihood of meeting federal recommendations for fruit and vegetable intake levels. To do this, The Reinvestment Fund's (TRF) Limited Supermarket Access data will be used in conjunction with the CDC's Behavioral Risk Factor Surveillance Systems (BRFSS) data, both nationally representative datasets. Second, this thesis will further seek to control for both individual and neighborhood socioeconomic variables by including the latest census data, American Consumer Survey data, and Consumer Spending data.

1.4 Organization of Thesis

This thesis will first review the literature with regard to the prevalence of overweight and obesity, and discuss their subsequent health consequences on millions of Americans (Chapter 2). It will further assess empirical analyses related to the local environment and overweight and obesity prevalence and the individual and neighborhood socioeconomic and demographic characteristics that may have an effect on weight status and eating habits. These characteristics include things such as education, income, race and ethnicity, and finally, the availability of access to fresh, healthy food by way of full service supermarkets and its impact on an individual's weight and eating habits.

Chapter 3 will discuss the data sets used for empirical analysis as well as key variables, and chapter 4 will discuss the methodology and hypotheses. Chapter 5 will discuss the empirical results, and the final chapter, chapter 6, will conclude the thesis with the main findings, policy implications, and limitations.

Chapter II
Literature Review

This thesis draws on a variety of literature and empirical analyses in an attempt to better understand the relationship between the local food environment and overweight and obesity. The literature review consists of three sections. The first section of the literature review will discuss the trends, prevalence of, and costs associated with overweight and obesity, and the increasingly serious health issue it has become. The second section will explore weight status as it relates to various socio-economic factors, including race, age, education, and income. Behavioral factors will then be discussed, as they have been found to have a subsequent impact on individual weight status. Perhaps most importantly, and most integral to this thesis, the final section of the literature review will focus on obesity and the local food environment. This will include supermarket access, various socioeconomic variables and their association with supermarket access and the subsequent prevalence of overweight and obesity, and the various measures with which prior research has assessed supermarket access.

As it has been documented that obesity is preventable with the incorporation of healthy eating habits and healthy lifestyle habits, it follows that an individual's environment may be problematic in terms of being conducive to healthy choices (Poston and Foreyt 1999; Lopez 2007; French, Story, and Jeffery 2001). While the Surgeon General calls for individuals to be responsible for healthy choices, at a broader level, they call for communities to support healthy lifestyles as well (U.S. Department of Health and Human Services 2001). The term "obesogenic" environment has been coined to include the environmental factors within neighborhoods that may lead to the increase in weight status

(Swinburn, Egger, and Raza 1999). The food environment is where the purchase and subsequent consumption of fresh, healthy, high-quality food at affordable prices occurs for the individuals who reside there (Glanz et al. 2005; Ford and Dzewaltowski 2008). Several studies have found a negative relationship between the neighborhood food environment and individual weight status—one or more aspects of the neighborhood food environment have been found to potentially cause an unintended increase in an individual's weight--which indicates that more research is required to better understand the neighborhood food environment (Papas et al. 2007). This thesis focuses on the local food environment, more specifically, whether the lack of supermarkets in a neighborhood is making individuals less fit.

2.1 Obesity in the United States: Trend, Prevalence and Costs

Overweight and obesity prevalence has been increasing steadily among children, adolescents, and adults since at least the 1960s (Bundred, Kitchiner, and Buchan 2001; Ogden et al. 2002; CDC 2012; Cynthia L. Ogden 2010). Quite alarmingly, the prevalence of obesity among adults aged 20 to 74 has doubled between 1971-1974 and 2011-2012, as noted in the table below.

Table 1. Obesity Prevalence, 1971 - 2012

Population Cohort	**1971-1974**	**2011-2012**
Preschoolers (2-5 years)	5.0[1]	8.4[2]
Children (6-11 years)[2]	4.0	17.7
Adolescents (12-19 years)[3]	6.1	20.5
Adults (20 years +)	12.1[4]	35.1[5]

1 (American HeartAssociation 2014)

2 (CDC 2015)

3 (Cheryl D. Fryar 2014)

4 This percentage is for adults aged 20 through 74 only,

http://www.cdc.gov/nchs/data/hestat/obesity_adult_11_12/obesity_adult_11_12.htm.

5 (CDC 2015)

Prevalence of obesity in the United States

Ogden et al. (2014) found that while there were no significant increases in obesity prevalence between 2003-2004 and 2011-2012, obesity prevalence remains high — more than one third of adults and 17% of youth in the Unites States are obese (Ogden et al. 2014).

Overweight and obesity, and the subsequent increase in related health risks — including but not limited to, hypertension, heart disease, premature mortality, and morbidity — have been well documented (Theodore B. Van Itallie 1985; Pi-Sunyer 1992; S Kumanyika 2002; Mokdad et al. 2003). Perhaps equally as alarming, as the severity of overweight and obese increases, so does the prevalence of obesity-related comorbidities (Aviva Must et al. 1999), regardless of gender, race, and differing socioeconomic groups (Paeratakul et al. 2002). In addition to the consequences obesity places on individual

health, there are economic consequences that obesity places on the healthcare system. Women who are obese have been estimated to cost the healthcare system just over $8,300 more than a normal weight woman, while obese men cost the system about $6,500 more (Dor et al. 2010). In general, our country is looking at healthcare costs of about $142 billion dollars a year, just from obesity (Rahman, Cushing, and Jackson 2011). In 2008 dollars, per capita annual medical expense on obese individuals was $1,429 higher relative to a normal weight individual (Finkelstein et al. 2009). It further must be mentioned that because obesity tends to cause additional health related diseases such as type 2 diabetes, the long-term effects on individuals will continue to be a burden (An 2015).

While this thesis focuses on adults — individuals aged 18 and up — many studies have been previously devoted to analyzing overweight and obesity in childhood. Because research has found that overweight and obesity in childhood is associated with health problems and obesity in adulthood (Rahman, Cushing, and Jackson 2011; Guo and Chumlea 1999; Serdula et al. 1993; Baker, Olsen, and Sørensen 2007), briefly discussing the prevalence and severity of childhood obesity is necessary. Serdula et al. (1993), offers a systematic review of the literature related to the link between obesity in childhood and overweight and obesity in adulthood. Specifically, about a third of obese preschool children were obese as adults, and about half of obese school-age children were obese as adults. The authors note that for all studies reviewed, and across all ages, the risk of adult obesity was at least twice as high for obese children as for non-obese children.

Cowley and Meyerhoefer (2012) noted that previous research, including Finkelstein (2009), may have actually underestimated the individual cost, with data indicating that obesity may raise annual medical costs by $2,741 when compared with normal weight individuals. A research report out of George Washington University found that the annual cost of obesity, including loss of life, was $8,365 for obese women and $6,518 for obese men (Dor et al. 2010). Rahman et al. (2011) cited a report by the American Public Health Association, noting that health care costs associated with overweight and obesity were estimated to be approximately $142 billion in 2010. Between 1998 and 2011, health care costs associated with obesity noticeably increased, and because obesity and associated type 2 diabetes primarily leads to life-long disability, obesity has the potential to be quite taxing to our health care system (An 2015)

This trend of increased spending on obesity is also applicable to overweight children. A frequently cited report by Marder and Chang (2006) found that under private insurance, the average total health expenses for an obese child are $3,743, a value that is more than triple the average expense for normal weight children ($1,108). Among children covered by Medicaid, the average healthcare costs increase greatly for obese children than for children on average ($6,730 vs. $2,446) (Jeffrey Levi et al. 2015; Marder and Chang 2006). A systematic review of the literature conducted in 2010 found excess healthcare medical spending to be as much as $14.3 billion annually for obese children (Hammond and Levine 2010), and a later study in 2012 found similar costs, as an elevated BMI in childhood was associated with $14.1 billion in additional prescription drug, emergency room, and outpatient visit costs annually (Chatterjee 2009).

2.2. Obesity and Socio-demographic Factors

Age, Education, Race, Income, and Marital Status

Obesity does not discriminate by age, race, or socioeconomic level, as people of all categories have had instances of obesity. However, the prevalence of obesity is found to be higher among middle-aged adults, less educated individuals, and low-income individuals (Sobal and Rauschenbach 2003; Charlotte A. Schoenborn 2002). First, let's address overweight and obesity in relation to age. More specifically, according to the CDC, obesity is highest among middle age adults aged 40-59 (39.5%) when compared with other adult age cohorts; obesity rates are lower for those aged 20-39 (30.3%) and adults over 60 years of age (35.4%) (CDC: Division of Nutrition 2015). A report published by the CDC in 2002 indicates that overweight and obesity varies significantly by age, gender, marital status, and a combination of each. For example, adults aged 18 to 24 years of age were found to be significantly less likely than adults of any other age to be overweight, and women were significantly more likely than men to be a healthy weight (Charlotte A. Schoenborn 2002).

Next, let's add gender to the age discussion, as weight status varies between men and women. Looking at adults aged between 45 years and 64 years of age, more men were found to be overweight (7 in 10 men, versus 6 in 10 women). Men aged 45 to 64 years of age were found to be twice as likely as the younger men and the older men to obese. For women, the prevalence of obesity was highest amongst those aged 45 to 64 years of age, and lowest amongst those aged 18 to 24 years of age—which is consistent with the notion

that the youngest adults are the least likely to be overweight (Charlotte A. Schoenborn 2002).

Minorities have shown a higher likelihood to be overweight or obese (Charlotte A. Schoenborn 2002). According to the CDC estimates, non-Hispanic blacks have the highest age-adjusted rates of obesity (47.8%), followed by Hispanics (42.5%), non-Hispanic whites (32.6%), and non-Hispanic Asians (10.8%) in 2012 (CDC: Division of Nutrition 2015).

Next, let's add in marital status. Continuing to cite the NCHS report, married men were found to be more likely than never married men to be overweight, yet married women were more likely than never married women *and* divorced or separated women to be in the healthy weight range. In general, the prevalence of obesity was found to be the lowest among cohabiting adults when compared with adults in other marital status groups (Charlotte A. Schoenborn 2002).

Ogden (2010) indicated that particularly among women, the prevalence of obesity is significantly lower among those who have a college education than those that have lower than a high school education, but the association between education level and the obesity risk is not statistically different for men. A separate, earlier report produced by the CDC indicated that men were more likely than women at each education level to be overweight, and that, in general, the prevalence of obesity decreases as education increased (Charlotte A. Schoenborn 2002). In other words, a healthy weight status is

positively associated with level of education for both men and women, although the association is more striking for women than men. Similarly, Giles-Corti, et al. (2003) found the odds of being obese were nearly three times higher in those with only a secondary education compared with respondents who indicated they had at least a college degree.

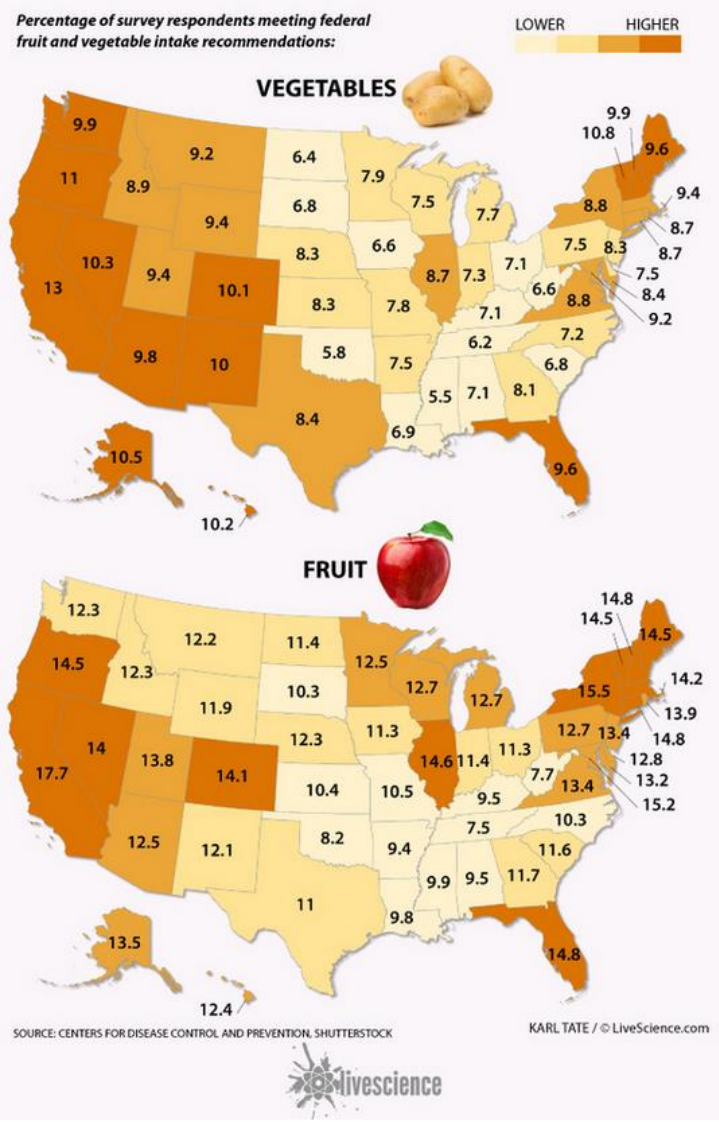
Income is an important factor that affects eating and exercise behaviors. Lopez (2007) finds that median household income is associated with the risk of obesity. Chang and Lauderdale (2005) find that when income is considered, both white and black women show an inverse relationship between income and weight status regardless of the beginning level of income, as does a 2010 NCHS data brief previously mentioned (Ogden CL 2010). White men tend to show the same relationship, though not as strong, while conversely, it appears that black men have an increase in weight status as income increases. Further, it was that found generally speaking, obesity prevalence is quite similar regardless of income level — though it does tend to be slightly higher at higher income levels. However, when taking race and ethnicity into consideration, the relationship between obesity and income level does vary. For example, among non-Hispanic black men, obesity prevalence decreases as income decreases, but there is no significant difference in obesity prevalence by poverty level among non-Hispanic white men (Ogden CL 2010). Low-income households, in general, were found to have a negative association with the availability of fruits and vegetables in the home (Edmonds et al. 2001). While there are different relationships between income, race, and weight status, Diez-Roux et al. (Diez-Roux et al. 1999) found that as individual income

increases, so does the consumption of healthier food items. Consistently, as neighborhood median income increases, as does the intake of healthier food items. It is unclear why these relationships occur — perhaps highly educated, high-income individuals simply have more choice as to food and are able to make more informed decisions regarding their healthcare, eating and exercise habits, and general environments in which they live.

2.3 Obesity and Behavioral Factors

Eating healthy is an integral part of maintaining a healthy weight and living a healthy lifestyle (CDC 2015). The *2010 Dietary Guidelines* were designed specifically to help people consume a healthy diet, and recommend the consumption of more foods such as fruits and vegetables, and less foods that contain sodium and saturated fats (U.S. Department of Health and Human Services 2010), as doing so helps manage body weight in addition to reducing the risk for heart disease and other ailments (Moore and Thompson 2015). In 2013, the CDC used BRFSS data to determine what percentage of American adults are meeting the federally recommended consumption of fruit (1.5 - 2 cups) and vegetables (3 cups)(Moore and Thompson 2015). Results indicated that less than 18% of adults in each state consumed the recommended amount of fruit, and less than 14% consumed the recommended amount of vegetables. Further, Fan and Jin (2014) find that relative to their normal-weight counterparts, overweight/obese individuals are characterized by poor eating habits as they are less likely to meet the recommended levels of fruit and vegetable consumption and nutrient and energy intakes.

Figure 1



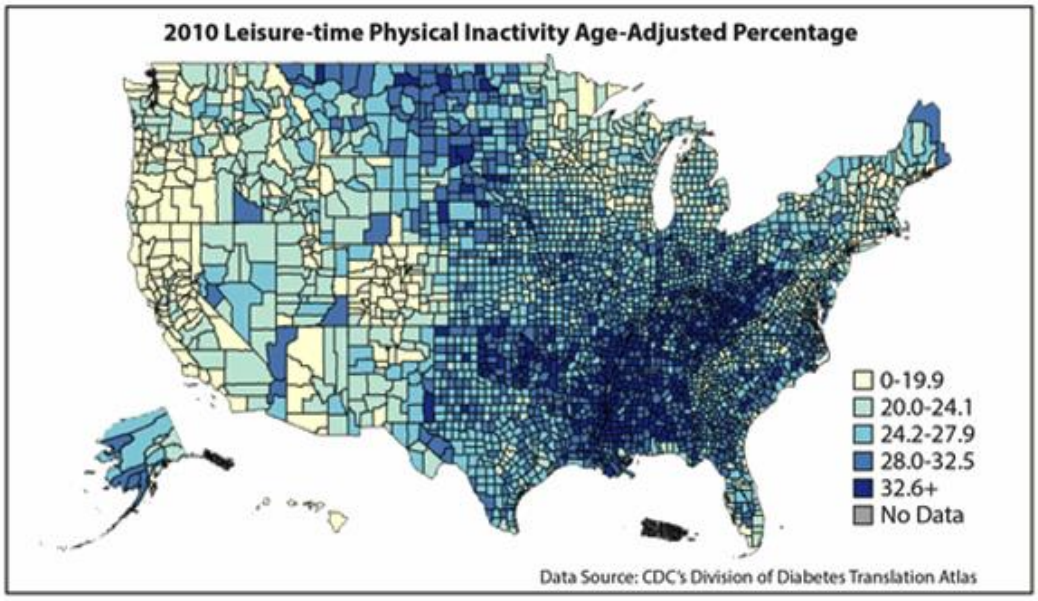
Source: (Tate 2015)

The CDC states that living a healthy lifestyle, particularly one that includes physical activity, is crucial. As of 2010, 50 percent of adults aged 18 and older did not meet the federal guidelines for aerobic activity or muscle strengthening physical activity (CDC 2012), and that statistic continues into 2014, as 48% of all adults still did not meet requirements (CDC 2014). The physical activity level demonstrates some regional,

gender, and racial patterns. American adults living in the South are less physically active than those living in the West, Northeast and Midwest; more non-Hispanic white adults (22.8 percent) meet the physical activity guidelines than non-Hispanic black adults and Hispanic adults (17.3 and 14.4 percent, respectively), and men typically exercise more than women. Perhaps as expected, more younger adults versus older adults meet the requirements, as do those with more education and those with family income above the poverty level (CDC 2014). Figure 1 illustrates age-adjusted estimates of the percentage of adults who are physically **inactive** by county. It visually indicates that the highest percentages of inactive adults are located in the Southeast, while much of the West coast contains more adults who are considered to be active.

Figure 2

2010 Age-Adjusted Estimates of the Percentage of Adults Who Are Physically Inactive



Source: (CDC 2014)

Smoking is another behavior factor that is used to analyze the obesity risk. Previous studies find that smoking may reduce total weight due to reduced calorie consumption (Lopez 2007). Patel et al. (2011) found that among low-income white and African American women, women who were identified as current smokers had a significantly lower average BMI when compared to former smokers and never smokers, but no significant differences were found between former and never smokers. The results are robust across different race groups (Patel et al. 2011). A study conducted in the United Kingdom with middle-aged adults also found that current smokers were less likely to be obese than never smokers, but age may play a role, as there was no significant relationship found for those less than 40 years of age. Further, former smokers were more likely to be obese than both current and never smokers (Dare, Mackay, and Pell 2015).

The literature offers mixed results on the relationship between binge drinking and overweight and obesity. Suter (2005) performed a systematic review of the literature and found the calories consumed from alcohol are more detrimental among moderate, nondaily consumers of alcohol than in daily (heavy) consumers; they are also more detrimental when consumed in combination with a high-fat diet, and among individuals who are already considered to be overweight or obese. The review, in general, supports the notion that while there are positive associations between higher amounts of alcohol and weight status, additional factors must be considered such as lifestyle, genetics, and social factors as there is potential for high variability from one consumer to the other (Suter and Tremblay 2005). For example, if a person is already overweight or obese, the

effects of alcohol may be more pronounced, and heavy drinkers are at a higher risk of being obese than moderate drinkers. With regard to binge drinking, researchers have found that the odds of overweight and obesity were significantly higher among this population — as well as heavy drinkers — in comparison to individuals who consume the same amount of alcohol, though spread out over multiple sessions (Mary Gatineau 2012). Arif and Rohrer (2005), analyzing data from just over 8,000 respondents in the third NHANES, found that the odds of overweight and obesity were significantly greater among binge drinkers and those consuming four or more drinks a day.

2.4. Obesity and Local Environment

Researchers Egger and Swinburn (1997) noted the importance of environmental influences on obesity. If an environment is particularly conducive to promoting obesity — an obesogenic environment — then any type of intervention or preventative programs are much less likely to be effective. Environmental influences have the potential to be controlled by public health agencies and continued research to pinpoint the cause of overweight and obesity is imperative to curb the pandemic (French, Story, and Jeffery 2001; Swinburn 1997). Because it is unlikely for the human gene pool to have changed drastically enough to cause the increased prevalence of overweight and obesity, it is likely that the cause is actually due to our environment (Poston and Foreyt 1999). The Surgeon General, in their *Call to Action to Prevent and Decrease Overweight and Obesity*, reaffirms that both behavioral and environmental factors are large contributors to overweight and obesity (Satcher 2010). It must be noted that behavior is not independent

of the environment, and that the environment has the potential to help determine behavior.

Our environment is very broadly defined as all that is external to an individual — the air we breathe, the water we drink, and the land and built structures that surround us. This includes the way in which we acquire healthy foods (Caballero 2007; Prevention 2009), as well as the safety of public places for recreational activities (Caballero 2007). This thesis will only focus on the local food environment.

While there is research related to neighborhood safety and its impact on the prevalence of overweight or obesity, it is quite scarce. Generally, results seem to indicate that as the perceived or actual rate of crime increases in a neighborhood, the prevalence of overweight and obesity will subsequently increase, an example of an environment influencing behavior.

Obesity and the Local Food Environment

Several studies have sought to measure the effects of the neighborhood food environment on individual weight status. This has become an increasingly important topic as researchers have accepted the notion that obesity may not just be a single-factor issue, but may be related to a host of additional factors — mainly, the construct of the neighborhood food environment. The CDC highlights the importance of consuming nutrient-dense foods including, but not limited to, fresh fruits and vegetables (Services 2010), and the World Health Organization (WHO) documents the importance of

increased consumption of fruits and vegetables and the importance of these items being available to all, at affordable prices (WHO 2015).

In its most basic form, the supermarket is how many Americans today, and for the past few decades, purchase and obtain their food. As such, it follows that access to one or more supermarkets within a reasonable proximity to a person's home is an integral part of how a person's diet is constructed — the foods that are available to purchase and consume, and at what prices, and the convenience of access to these foods. As early as 1999, access to supermarkets was suggested as an important factor in the physical environment and its subsequent importance in reducing pandemic levels of obesity (Swinburn, Egger, and Raza 1999), and, in their 2010 Dietary Guidelines, the CDC confirmed that incorporating improved eating patterns will lead to better weight management and has the potential to prevent or reduce overweight and obesity. An integral part of this improvement is done by increasing the intake of fruits and vegetables, and research has shown that it is easiest to purchase at full-service supermarkets.

The literature discusses large, full-service supermarkets or chain supermarkets and why it is that these stores are better for consumers to purchase a variety of healthy foods at lower prices. A multitude of prior research indicates that prices and availability, particularly the availability of fresh produce, tend to be much better at chain supermarkets versus those at smaller grocery stores or convenience stores (Chung and Myers 1999; Morland and Filomena 2007; Bodor et al. 2008; Andreyeva et al. 2008; ERS 2009). In 2007, Papas et al. (2007) conducted a systematic review of the literature that

studied associations between obesity, or body mass index (BMI), and some aspect of the built environment. Of the 20 articles reviewed, 17 reported a statistically significant relationship between BMI and an aspect of the built environment. Perhaps most importantly, two longitudinal studies found a significant association between lower fruit and vegetable prices and lesser gains in BMI over a 3-year period.

In terms of availability, shelving space devoted to fruits and vegetables was found to be considerably larger for supermarkets than for small food stores, and this discrepancy in space was especially apparent for fresh produce items, in which supermarkets had greater amounts of shelf space than small food stores (Bodor et al. 2008). The same study found that supermarkets also tend to offer a considerably larger number of varieties of fresh produce than do small food stores. Horowitz et al. (Horowitz et al. 2004) had consistent findings — availability of healthy foods tends to vary by store size and prices tend to be higher in small bodegas rather than large supermarkets. Though this particular study examined young children, Sturm and Datar (2005) indicated that lower prices for fresh fruits and vegetables were found to predict a significantly lower BMI. In addition to trends of larger selection and more affordable prices, grocery stores were also found to have higher quality (Cerin et al. 2011), particularly in higher-income areas (Andreyeva et al. 2008).

As noted, the literature on the effect of local food environment on obesity offers mixed results (Cannuscio, Weiss, and Asch 2010; Koplan, Liverman, and Kraak 2005). This

following section will explore research related to supermarket access to fresh and healthy food and its subsequent effect on diet and weight status.

In addition to supermarkets, we must address that individuals may also eat food away from the home, at restaurants or fast-food chains. Larson et al. (2009) reviewed the literature and identified associations between restaurant access and dietary intake; greater availability of chain fast food restaurants may promote greater fast-food intake in low-income groups (Janne Boone-Heinonen et al. 2011). Further, there tend to be more fast-food chain restaurants in low-income areas and minority areas — neighborhoods that are already at risk for obesity as previously established. This is of concern as data indicates that people who regularly consume foods away from the home, particularly at fast-food restaurants, are heavier than those who do not; this is still the case after controlling for various sociodemographic variables (Macintyre 2006).

Supermarket Access

The relationship between having access to a supermarket and decreased prevalence of obesity is perhaps linked directly to consumption of healthy food items purchased at those stores. As early as 1991, researchers found statistically significant relationships between measure of availability of healthful products and the subsequent reported consumption of those products by individuals living near the stores (Cheadle et al. 1991). These relationships seem to be most relevant for two units of geographic aggregation — the community and zip code. Cheadle defines a community in his research as generally an entire county or city, and then these communities are further subdivided by zip code

boundaries. Another study conducted in 2002 indicated that when at least one supermarket was located in the census tract, black Americans reported a higher consumption of fruits and vegetables (Morland, Wing, and Roux 2002). Zenk et al. (2005) further confirms the relationship of supermarkets and increased intake of fruits and vegetables; the women studied were found to consume 1.2 and 2.37 more fruits and vegetables per day than women who shopped at independent stores. Lopez (2007) also found that having one or more supermarkets in a zip code tabulation area decreased the risk of obesity by nearly 11%, and the CDC has noted that supermarket access is associated with a reduced risk for obesity (CDC 2013, 2011).

It must be noted though, that a limitation to many of the studies conducted on prices and availability of healthy foods is that the research is done in a geographically small area rather than on a national level (Andreyeva et al. 2008).

Supermarket Access and Socioeconomic Factors

As the link between healthy food consumption and supermarkets has been established, we must now discuss the difference in access and availability of supermarkets with relation to socioeconomic factors.

Morland et al. (2002) and Powel et al. (2007) found that there were more than three times as many supermarkets in wealthier neighborhoods and that these neighborhoods also contain fewer small grocery and convenience stores when compared with lower income neighborhoods. This is important because the research has upheld the notion that smaller grocery stores and convenience stores tend to have a limited selection of items —

especially fresh fruits and vegetables — and tend to have higher prices compared to supermarkets. Furthermore, it has been found that these chain stores, with the better prices and increased availability, tend to not locate in lower income areas which leads to low-income households paying higher prices for fresh, healthy food (Chung and Myers 1999).

Access to grocery stores is not as simple as living in a low- versus high-income area; race and ethnicity play an important factor. Morland et al. (2002) found that supermarket prevalence was four times as high in predominantly white neighborhoods when compared to predominantly black neighborhoods. Morland et al. (2002) confirmed these findings — more supermarkets were located in census tracts where white residents resided when compared to census tracts in which black residents resided (Morland, Wing, and Roux 2002). Furthermore, Morland and Filomena (2007), and numerous others (Moore and Diez Roux 2006; Powell et al. 2007)(Moore and Diez Roux 2006; Powell et al. 2007) had consistent findings — supermarkets were more prevalent in predominantly white areas, while smaller grocery stores and bodegas were found in predominantly African American areas. Of the stores located in predominantly African American areas, they had lower varieties of fresh produce when compared with predominantly white areas (Morland and Filomena 2007; Moore and Diez Roux 2006). Powell et al. (2007) found that Hispanic neighborhoods have only 32% as many chain supermarkets compared to non-Hispanic neighborhoods. The results remain the same even after controlling for income for both Hispanic and predominantly African American neighborhoods. These results may imply racism by retailers, but these particular studies do not address the issue.

Morre et al. (2008) assessed both the perception-based and geographic information system (GIS)-based food environment. With regard to perception, in general, minorities and lower income respondents reported lower perceived availability of health foods than whites and higher income respondents, respectively. For the GIS-based environment, supermarket densities were lower for blacks than other race/ethnic groups, and were higher for participants with annual household incomes over $50,000. Non-white and lower income participants were generally reported living in areas with higher densities of smaller stores than white and higher-income participants. Perceived availability of healthy foods was positively associated with the densities of supermarkets (Moore, Roux, and Brines 2008).

Conversely, and inconsistent with a majority of the data, some researchers have found that high poverty neighborhoods actually have *more* supermarkets per land area when compared with more affluent or wealthy neighborhoods (Lee 2012). In addition to differing results possibly caused by the use of different datasets, the findings suggest that it may not be a result of lack of access, but rather an issue of ease of access (Lee 2012). A multi-study report published by the U.S. Department of Agriculture (USDA) highlights this finding (ERS 2009).

Food insecurity must also be addressed, as prior research seems to indicate that individuals who do not have the means to acquire sufficient food for themselves and their families may have an increased risk of diet-related diseases, particularly obesity (Brandi Franklin 2012). Food insecurity is generally defined by the USDA as limited or uncertain

availability of, or ability to acquire, nutritionally adequate food by socially acceptable means (i.e. without stealing, accessing emergency food supplies, or relying on other coping strategies) (Brandi Franklin 2012). In a review of the research regarding the insecurity-obesity paradox, Dinour et al. (2007)summarized fourteen studies, published from 1999 to 2006, and found that there are positive links between food insecurity and obesity, particularly among women (Dinour). Additionally, the feast-famine cycle has been linked with obesity, as individuals who use food stamps often "feast" for the 3 weeks that food stamps and money are available, and then go through a 1 week period of "famine" until the following month (MS Townsend 2001; PE Wilde 2000).

Measure of Supermarket Access

The literature offers different measures of supermarket access and how it affects eating behaviors and the consequent health outcomes. Some studies found that proximity to supermarkets has been positively associated with consumption of a healthy diet and negatively associated with overweight or obesity (Laraia et al. 2004; Morland, Roux, and Wing 2006; Morland, Wing, and Roux 2002), particularly when compared with smaller grocery stores or fast food restaurants (Morland and Evenson 2009), and convenience stores (Morland, Roux, and Wing 2006). Bodor et al. (2008) found that respondents who had a small food store within 100m of their home had a significantly higher average intake of vegetables and a marginally significant average intake of fruits. Consistently, respondents with no fresh vegetable space available within a block of their residence had the lowest average intake of vegetables (Bodor et al. 2008). Taking store classification a bit further, Rundle et al. (2009) classified food establishments into BMI-healthy and

BMI-unhealthy, where a BMI-healthy food outlet is defined as a supermarket or fruit and vegetable market, they found that where there is a higher density of BMI-healthy food outlets, there is a lower mean BMI and lower prevalence of overweight and obesity in the neighborhood. Boone-Heinonen et al. (2013) also found that increasing supermarket density over time leads to a healthy BMI. Liu et al. (2007) found that an increased distance between a subject's residence and the nearest large brand name supermarket was associated with increased risk of overweight; however, this relationship was only significant for subjects living in lower population density regions (Liu et al. 2007). On a much smaller scale and with a sample sized that focused on pregnant women, Laraia et al. (Laraia et al. 2004) find that if a supermarket was located more than 4 miles from the home, there was a significant negative association with diet quality.

Other studies use the presence of supermarkets to measure availability. Using BRFSS data from 1998 to 2002 for Eastern Massachusetts and after controlling for individual-level factors, median household income, population density, employment density, establishment density, and the presence of a supermarket, Lopez (2007) found that the presence of a supermarket decreased the risk of being obese. This study faced numerous limitations, as does this thesis; for example, height and weight is self-reported, and the data is used geographically by zip code which tend to be large areas that could differ quite greatly in terms of socioeconomic factors. Dissimilar to this thesis though, Lopez only included zip codes in Eastern Massachusetts-- quite a small sample size compared to the entire United States, which this thesis utilizes.

Though much of the data indicates that there is a relationship between fruit and vegetable consumption and distance to and availability of supermarkets in neighborhoods, Pearson et al. (2005) finds no significant relationship. It must be noted, though, the study was very limited in that it had a very small sample population.

Chapter III Data and Key Variables

3.1 Data Sources

The main data set used for this thesis is the 2007 Behavioral Risk Factor Surveillance System (BRFSS) data collected by the CDC. The BRFSS survey is a cross-sectional telephone survey conducted by state health departments on a yearly basis. The purpose of the BRFSS survey is to collect information regarding individual risk behaviors and preventative health practices that can affect their health status among adult residents living in households in the United States (CDC 2014). Once the CDC receives the data from the state health departments, it is edited, processed, weighted, and analyzed (CDC 2007). Perhaps most relevant to this thesis, the BRFSS collects information of self-reported height and weight, as measured without shoes on, socio-demographic status (e.g., age, marital status, race, income, education, and employment status), health conditions (e.g., self-reported health status), risk behaviors (e.g., smoking and drinking history), fruits and vegetable consumption, and physical activities. Based on the 2007 BRFSS data, I created dependent variables, including the risk of being overweight or obese and the probability of meeting the federal recommended level of fruit and vegetable consumption. A set of independent variables at the individual level were also created to measure socio-demographic status, health conditions, and risk behaviors.

The second main data source is The Reinvestment Fund's (TRF) Limited Supermarket Access (LSA) data of 2010. Classified as a community development financial institution (CDFI), TRF is a national leader in the financing of neighborhood revitalization and has provided the policy expertise and financing for supermarkets, grocery stores, and other healthy food projects that plan to operate in underserved communities. TRF's Policy

Solutions group uses spatial analysis to define an area that has limited supermarket access, which is conceptually identified as an LSA area--areas where residents travel longer distances to reach supermarkets when compared to the average distance traveled by residents of non-low/moderate income areas (Policy Solutions 2011). First, all block groups within the continental United States were categorized using census data for population density and car ownership. Census block groups were used as the geographic unit of analysis. TRF further calculates the distance traveled from the population center of every census block to the nearest full-service store. Once this was completed, benchmark distances were calculated. The benchmark distance represents a "comparatively acceptable" distance for households to travel to a supermarket, and TRF has defined "comparatively acceptable" as the distance that residents of well-served areas (with incomes greater than 120% of the area's median income) travel to the nearest supermarket, compared to other residents within the same density category. Each block group was then assigned an access score, which represents the percentage that an LSA block group's distance would have to be lowered in order to equal the reference group's distance. As such, a high access score indicates a more pronounced problem of lacking supermarket access, and clusters of block groups with high access scores are finally identified as LSA areas. Generally speaking, TRF seeks to define areas with inadequate supermarket access as areas in which residents must travel significantly farther to the nearest full-service grocery store than residents of areas showing similar population density and car-ownership characteristics as well as median household incomes greater than 120% of the area median. TRF uses supermarket data from Trade Dimensions, and only includes full-service supermarkets (Catherine Califano 2012).

The original LSA data is at the block group level. I was able to aggregate to the zip code level with the help of Jim Trimble, a Database Administrator for the Center for Remote Sensing and Spatial Analysis at Rutgers University. For more information regarding this process, please refer to the data appendix.

I use zip codes to match the LSA data and the 2007 BRFSS data, as the 2007 BRFSS data is the most recently available year that contained zip code information. In particular, I create two variables of interest at the zip code level, namely, the percent of land that is classified as an LSA area and the number of disjointed LSA within each zip code. Disjointed LSAs are defined as more than one land area, within in a zip code, that are separate and do not touch. Disjointed LSAs may be located geographically opposite of each other within a neighborhood, or not, but most importantly they are two separate areas.

The supplementary data sets used for this thesis are the 2010 Census, the 2010 American Community Survey, and the 2012 Esri Consumer Expenditures Survey data to gather community level variations at the zip code level.

3.2 Key Variables

Dependent variables

I have created two dependent variables: the risk of being overweight or obese, and the probability of meeting the federal recommended consumption levels of fruits and vegetables. This thesis utilizes an incredibly rich set of both individual- and neighborhood-level characteristics that when controlled for, will help to determine the

main drivers of overweight and obesity, and the failure to meet federal recommended consumption levels of fruits and vegetables.

Using the self-reported weight and height data, as measured without shoes on, that are reported by BRFSS respondents, I calculate the body mass index (BMI). Two issues must be noted. First, self-reported weight and height are found to have significant measurement errors. The nationally representative data set that I am aware of that has actual measured height and weight data for individuals is the National Health and Nutrition Examination Survey (NHANES); however, the NHANES does not have zip code information. Second, although a person's BMI correlates to the amount of body fat, it may not be an accurate measure for some people, such as athletes as BMI has been shown to overestimate fatness in those who are muscular (Richard V. Burkhauser 2008). Yet, BMI is a generally accepted measurement for both academia, particularly the social sciences, and public health. An adult is considered normal weight if their BMI is between 18.5 and 24.9, overweight if their BMI is between 25 and 29.9, and obese if their BMI is 30 or higher (CDC 2012, 2015). As such, I created three binary variables for weight outcomes: overweight only, obese, and overweight and obese. Overweight contains only individuals who are overweight coded as one, and those are who are normal weight are coded as zero. Obese contains individuals who considered obese, but not overweight, coded as one, and those who are normal weight coded as zero. Lastly, overweight/obese contains individuals who are both overweight *and* obese coded as one, and those who are normal weight coded as zero. Of the 310,400 total observations, 36.9% of individuals

self-reported a height and weight that indicates they are overweight, and 26.9% are classified as obese.

The second, separate, dependent variable, measures whether an individual does or does not meet the federal recommended consumption levels of fruits and vegetables. The USDA, in their *MyPlate* guidelines, recommends 1.5 cups of fruit for women age 31 years or older, and 2 cups for women between the ages of 19 and 30. Men, regardless of age, are recommended to consume 2 cups of fruit (USDA). Additionally, women between the ages of 19 and 50 years of age are recommended to consume 2.5 cups of vegetables, while women 51 years of age or older only need 2 cups. Men between the ages of 19 and 50 require 3 cups of vegetables, while men 51 years of age or older only require 2.5 cups (USDA). Note that each of these requirements are meant to be fulfilled each day. As such, the BRFSS survey asks respondents the number of servings of fruits and vegetables consumed each day (CDC 2007). For that reason, this thesis equates a cup to a "serving", and 2 cups of fruits and 3 cups of vegetables were the benchmark used to define the federally recommended level for daily fruit and vegetable consumption. That is, if an individual consumed 2 (3) servings of fruits (vegetables) a day, he/she meets the federal recommended consumption level. Similar to the first set of independent, binary variables, if the recommended level of consumption is met, the response is coded as one, and zero if it is not. As prior researches have indicated, adequate consumption of fruits and vegetables is essential to both attain and maintain a healthy body weight. Not even a quarter of respondents indicated meeting the fruit and vegetable consumption; just over

18% reported consuming at least 2 servings of fruit per day, and only 11.5% reported consuming at least 3 servings of vegetables per day.

Independent Variables

I have two sets of independent variables, as briefly mentioned above; individual- and neighborhood-level characteristics.

Perhaps most integral to this thesis, let's first discuss the two LSA variables that I have created for each zip code area based on TRF's LSA data: percentage of zip code land area considered to be an LSA and number of disjoined LSA areas within a zip code. If two zip code areas have the same percentage of land classified as LSA areas, we expect that individuals living in a zip code area with disjoined LSAs may have better access to supermarkets compared with those living in a zip code area with only one LSA area. Of the 24,475 unique zip codes identified for this thesis, the average percentage of land classified as an LSA within each zip code is 7.32%.[1] For the purpose of this thesis, I am able to look only at zip codes in which BRFSS survey respondents currently reside, and as such, it must be noted that I am not able to assess the entire country

Individual level characteristics are from the BRFSS data. Physical activity is another side of the energy balance equation; just as the consumption of fruits and vegetables is crucial to attaining and maintaining a healthy body weight, so is regular physical activity. As recently as 2013, only 20 percent of Americans were meeting the federal guidelines for

[1] Though the number of zip codes in the United States tends to fluctuate month by month, there are currently approximately 43,000 zip codes (*Zip Code FAQs* 2014)

aerobic activity or muscle strengthening physical activity (CDC 2013). For the best health results, an individual should use a mix of both aerobic and muscle-strengthening activities. According to the CDC, in the least, adults need 2 hours and 30 minutes of moderate-intensity aerobic activity a week in conjunction with muscle-strengthening activities on 2 or more days a week that will work major muscle groups. It is possible to substitute the 2 hours and 30 minutes of moderate-intensity aerobic activity each week with 1 hour and 15 minutes of vigorous-intensity aerobic activity. Lastly, it is also possible to do an equivalent mix of moderate- and vigorous-intensity aerobic activity. To gain the full benefit, muscle-strength training must be performed at least twice a week, regardless of your method of aerobic exercise (CDC 2014).

For this thesis, meeting physical activity requirements is defined as meeting moderate physical activity requirements, meeting vigorous physical activity requirements, or both; responses are self-reported by the individual, and subsequently re-calculated by the CDC to determine if requirements are met. The BRFSS data defines meeting moderate physical activity requirements as doing moderate activity such as brisk walking, bicycling, vacuuming, gardening, or anything else that causes some increase in breathing or heart rate, for at least 30 minutes a day, five days a week. For vigorous physical activity, the requirement is at least 20 minutes a day, for three days a week, and vigorous activity includes running, aerobics, heavy yard work, or anything else that causes large increases in breathing or heart rate. An individual is reported to meet the federal recommended requirement for moderate or rigorous physical activities if the above requirement is fulfilled. Of the survey respondents, 20.5% met the federally recommended level of

moderate physical activity, while just 11.7% met the recommended level of vigorous physical activity.

Other individual level variables that prior research has deemed to have an association to obesity risk are education, age, sex, employment status, marital status, race/ethnicity, income level, health conditions, and risk behaviors (smoking and drinking).

Neighborhood characteristic data was extracted from Esri Demographics, and by zip code, included 2010 Census information (population density, age, average household size), 2006 - 2010 American Community Survey (ACS) information (poverty ratio, median household income, households with public assistance, average number of vehicles available), and finally, 2012 consumer spending information (amount of money spent on fast food, dining out, and overall food expenditures, as well as money spent on exercise equipment and various food items--fruits and vegetables).

Together, these variables provide an incredibly rich description of each zip code or neighborhood where the BRFSS survey respondents currently reside.

Chapter IV Estimation Methodology

4.1 Conceptual Framework

The literature suggests that overweight and obesity are the result of a long-term balance between calories consumed and those expended, where both calorie intake and expenditure are affected by individual characteristics, the built environment, and their personal behavioral choices (Swinburn, Egger, and Raza 1999). An individual is classified as either overweight/obese or normal weight, which is used as a dependent variable for this thesis. The theoretical basis for the use of discrete choice models is the random utility model. Individuals are assumed to derive their utility from individual-level characteristics and neighborhood attributes. We denote individual-level socio-demographic characteristics by X_i and neighborhood-level characteristics by C_j, where the subscriptions *i* indicate individual characteristics and *j* indicate neighborhood characteristics. The random utility of an individual *i* living in neighborhood *j* is given below:

$$(1)\ U_{ij} = V_{ij}(X_i, C_j) + \varepsilon_{ij} = \alpha' X_i + \beta' C_j + \varepsilon_{ij}$$

Where V_{ij} represents the deterministic component, ε_{ij} is the error term, and the set of parameters, α*'s* and β*'s*, are to be estimated and they reflect the impact of changes in the independent variable on the obesity risk.

Specifically, individual characteristics X_i are compiled based on the BRFSS data. These characteristics include, but are not limited to, gender, age, race, education level, health risk factors, marital status, employment status, and income. Health risk factors include an

individual's smoking and binge drinking habits, as well as their self-reported general health condition. Neighborhood attributes include LSA-related variables, indicating the percentage of land classified as an LSA and the number of disjointed LSA areas within a given zip code, which are the variables of interest for this thesis. Other neighborhood attributes include median household income, average spending habits related to food and exercise equipment purchases, and the poverty ratio.

Let Y_{ij} indicate whether an individual *i* living in area with zip code *j* is obese ($Y_{ij} = 1$) or not ($Y_{ij} = 0$). A binomial logit model assumes the error term is independent and identically distributed to reflect all that cannot be controlled for. Following (Green 2011), the probability of being obese can be derived based on equation (1a and 1b):

$$\text{(1a) Prob } (Y_{ij} = 1) = F(\alpha' X_i + \beta' C_j) = \frac{e^{\alpha' X_i + \beta' C_j}}{1 + e^{(\alpha' X_i + \beta' C_j)}}$$

$$\text{(1b) Prob } (Y_{ij} = 0) = 1 - F(\alpha' X_i + \beta' C_j) = \frac{1}{1 + e^{(\alpha' X_i + \beta' C_j)}}$$

where F(•) is the logistic cumulative distribution function. The expected obesity risk can be written as:

$$\text{(2) } E[y] = 0\,[1 - F(\alpha' X_i + \beta' C_j)] + 1\,[F(\alpha' X_i + \beta' C_j)] = F(\alpha' X_i + \beta' C_j)$$

The estimates of α's and β's do not fully capture the marginal effect of the independent variables on the obesity risk. Instead, the marginal effect of each independent variable on the obesity risk is

$$\text{(3a) } \frac{\partial E[y]}{\partial x_i} = F(\alpha' X_i + \beta' C_j)\,(1 - F(\alpha' X_i + \beta' C_j))\ \alpha_i$$

(3b) $\frac{\partial E[y]}{\partial c_j}$ = F (α' X_i + β'C_j) (1 – F (α' X_i + β'C_j)) β_i

Equations (3a) and (3b) allow us to quantify the effect that individual characteristics and neighborhood attributes have on the obesity risk.

The same model can be used to analyze the effect of individual and neighborhood characteristics on the likelihood of meeting the federal recommended intake levels of fruits and vegetables.

Stata/IC 12.1 was used to complete the data analysis for this thesis.

4.2 Hypotheses

This thesis proposes two testable hypotheses relating to (1) the relationship between the weight measures and whether an individual resides in a limited supermarket access area, and (2) between the weight measures and whether an individual meets the federally recommended consumption level of fruits and vegetables.

Listed below is each hypothesis that will assist in fulfilling the objective of this thesis.

H1: ***The local food environment, such as limited access to supermarkets, will increase obesity risk and decrease the probability that an individual will meet federally recommended intake levels of fruits and vegetables.***

Many studies have been conducted in an attempt to measure the food environment and its subsequent impact on the consumption of fruits and vegetables, as well as weight status. Hypothesis H1 hypothesizes that lacking access to a supermarket will increase the risk of being overweight or obese. This hypothesis is supported by previous findings indicating

that proximity to supermarkets is positively associated with consumption of a healthful diet and negatively associated with overweight or obesity (Morland et al. 2002; Laraia et al. 2004; Morland, Roux, and Wing 2006). It is further possible that these two ideas go hand in hand, as the increased consumption of fruits and vegetables is made possible by access to supermarkets (Cheadle et al. 1991; Morland, Wing, and Roux 2002; Zenk et al. 2005; Lopez 2007).

More specifically, let *P_LSA* indicate the percentage of land in a zip code that is classified as LSA areas and *N_LSA* counts the number of disjointed LSAs in a zip code. The impact of these two variables on the weight measures can be formulized below:

(4a) $ME_{P\_LSA} = \frac{\partial E[obesity]}{\partial_{P\_LSA}} = F(\alpha' X_i + \beta' C_j)(1 - F(\alpha' X_i + \beta' C_j)\beta_{P\_LSA}$

(4b) $ME_{N\_LSA} = \frac{\partial E[obesity]}{\partial_{N\_LSA}} = F(\alpha' X_i + \beta' C_j)(1 - F(\alpha' X_i + \beta' C_j)\ \beta_{N\_LSA}$

Where $\partial_{P\_LSA\_Pop}$ and ∂_{N_LSA} are the estimated coefficients.

Equation (4a) formulates the marginal effect of *P_LSA* on the obesity risk, while keeping everything else constant, including *N_LSA*. That is, given that the number of disjointed LSA areas remains the same, a marginal increase in the percentage of the LSA area will change the obesity risk by ME_{P_LSA}. We expect that ME_{P_LSA} is positive — as the percentage of zip code considered an LSA increases, the risk for obesity also increases.

Equation (4b) formulates the marginal effect of the number of disjointed LSA areas in a zip code on the obesity risk, while keeping everything else constant, including the percentage of LSA areas in the zip code area.

Assume two zip code areas have the same percentage of land classified as LSA areas. One zip code area has only one big LSA, and the other zip code area has two disjointed

LSA areas. In general, individuals who reside in the second zip code area have better access to supermarkets compared with their counterparts in the first zip code area. Therefore, it is possible that the ME_{N_LSA} may be negative, indicating a decrease in the obesity risk if the number of disjointed LSAs increases but the total LSA land proportion remains the same. Similarly, I expect the percentage of LSA areas in the zip code to decrease the likelihood of meeting the federal recommended intake level of fruits and vegetables, while an increased number of disjointed LSAs, keeping the same LSA land proportion, will increase the likelihood of meeting the recommended fruit and vegetable consumption.

H2: ***Individual demographic characteristics, such as higher levels of income and education, increase the probability that an individual will meet federally recommended intake levels of fruits and vegetables, as well as reduce obesity risk.***

Individual demographic characteristics that have been previously found to be or hypothesized to be associated with obesity risk include gender, age, income, and race. In general, older individuals, less educated individuals, low-income individuals, and women are more likely to be obese (Charlotte A. Schoenborn 2002, {Sobal, 2003 #66; Sobal and Rauschenbach 2003). Conversely, Giles-Corti et al. (2003) found that men were more likely than women to be overweight, and this was positively correlated with age (Giles-Corti et al. 2003). The same study found similar results with education—the odds of being obese were nearly three times higher in those with only a high school diploma versus those with at least a college education. The National Center for Health Statistics indicated consistent findings in their 2010 data brief (Ogden CL 2010). Quite consistent throughout the research, income is found to have an inverse relationship with obesity risk

(Chang and Lauderdale 2005; Lopez 2007), and increased income is found to have an association with increased consumption of healthier food items (Diez-Roux et al. 1999). As various individual-level demographic characteristics have been found to be associated with obesity risk, this thesis also hopes to determine if these same characteristics have an impact on the consumption of healthy fruits and vegetables. For example, Edmonds et al. found that low-income households tend to have a negative association with the availability of fruits and vegetables (Edmonds et al. 2001).

Behavioral characteristics must also be included in this hypothesis, particularly an individual's smoking habits and tendency towards binge drinking. This thesis hypothesizes that overweight and obese status will be negatively associated in individuals who are current smokers, and in general, current smokers tend to have a lower BMI than former and never smokers (Patel et al. 2011; Dare, Mackay, and Pell 2015). While the relationship between weight status and binge drinking has been documented to be quite complex, this thesis hypothesizes that binge drinkers are more likely to be overweight or obese (Arif and Rohrer 2005). Lastly, a person's physical activity must be analyzed, as well, as the CDC states that living a healthy lifestyle, particularly one that includes physical activity, is crucial in maintaining a healthy weight. This thesis hypothesizes that those who meet the federal guidelines for at least moderate physical activity will have a lower weight status and have a higher likelihood of meeting recommended fruit and vegetable intake levels.

Chapter 5 Empirical Results

5.1 Descriptive Statistics

BRFSS Survey responses that were missing zip codes were immediately removed, as were responses with missing demographic-level characteristics, or neighborhood-level characteristics. A total of 310,400 survey responses remained for this thesis.

The data was then identified as survey design within Stata and sample weights were assigned.

Referring to Table 2 below, of the 310,400 total observations, 36.2% were classified as having a normal weight, while nearly 37% were classified as overweight and 26.9% as obese. Only 18% of respondents meet the federally recommended level of fruit consumption (2 cups), and even less (11.5%) meet the recommended level of vegetable consumption (3 cups). Just under a third (32.2%) meet at least one of the federally recommended levels of physical activity (moderate or vigorous).

Table 2. Weight status, recommended fruit/vegetable consumption level, and physical activity requirements of the sample

	Mean (%)	**Std. Dev.**
Normal Weight	36.15	0.0020
Overweight Only	36.95	0.0019
Obese	26.90	0.0018
Meeting federally recommended fruit intake level	18.30	0.0016
Meeting federally recommended vegetable intake level	11.50	0.0012
Meeting federally recommended level of moderate physical activity	20.54	0.0015
Meeting federally recommended level of vigorous physical activity	11.71	0.0014
Meeting *at least* one of the federally recommended level of physical activity (moderate *or* vigorous)	32.25	0.0019
Total Observations		**310,400**

5.1.1 Individual-level characteristics

Of the 310,400 survey respondents used for this thesis, exactly half identified as male (50.3%). The mean age of all respondents was 45.9 years. Almost two thirds indicated that they were married (62.8%), more than a third had acquired a college degree or higher (36.4%), and the majority of respondents identified as being currently employed (64.4%). Further, while 14.9% indicated that they were retired, nearly 21% are not currently employed. Individuals were overwhelmingly non-Hispanic white (71.1%), followed by Hispanic (12.6%), followed by non-Hispanic black (9.2%) and other 7.0%, which includes all other races. The mean household income is $50,900.

In terms of self-reported health, the vast majority have reported good, very good, or excellent health (84.8%). The remaining individuals have self-reported their health as either fair, or poor. Nearly 56% reported never having smoked, while about a quarter of respondents (24.6%) were current smokers and 19.6% had quit; just under 17% report consuming enough alcoholic beverage to be considered a "binge" drinker. Barely a quarter of the respondents reported meeting moderate physical activity requirements (20.5%), and even less reported meeting vigorous activity requirements (11.7%). Similarly, only 18.3% of respondents indicated they consumed enough fruit to meet the federally recommended level, and even less (11.5%) indicated they consumed enough vegetables to meet the federal recommended level. Please refer to Table 3 below for listed results.

Table 3. Individual-level Characteristics

	Mean (%)	Std. Dev.
Male (%)	50.33	0.0020
Age (years)	45.86	0.0702
Marital Status: Married (%)	62.84	0.0021
Education: Less than high school (%)	9.33	0.0015
High School (%)	54.30	0.0020
College Graduate (Bachelor's degree or higher) (%)	36.37	0.0019
Employment: Currently Employed (%)	64.42	0.0019
Retired (%)	14.85	0.0011
Unemployed (%)	20.72	0.0017
Household Income ($1,000)	50.90	.1077
Race: Non-Hispanic White (%)	71.15	0.0022
Hispanic (%)	12.60	0.0020
Non-Hispanic Black (%)	9.23	0.0012
Other (%)	7.02	0.0013
Self-reported General health: Good (%)	84.76	0.0015
Poor (%)	15.24	0.0015
Smoke Status: Former Smoker (%)	19.58	0.0016
Current Smoker (%)	24.60	0.0020
Never Smoked (%)	55.58	0.0012
Binge Drinker (%)	16.67	0.0017

5.1.2 Zip Code-level characteristics

The largest percentage of survey respondents lived in the South Atlantic region of the country, followed by the Pacific region, and the East North Central region. Using data from the Census, as well as the American Community Survey, we are able to find out a bit more information about the zip codes, or neighborhoods, in which survey participants resided.

First, let's briefly discuss the spending habits. The average share of household income, by zip code, spent on food is 11.3%. Of the total spent on food, the average share spent on food eaten away from the home is nearly 39%, which includes fast food spending (15.4%). Interestingly, exercise spending, which includes the purchase and participation in sports, recreation, exercise equipment, and bicycles, comes in at 25%. It must be noted that this value includes spending on motorized sporting equipment and recreational vehicles.

Of the zip codes analyzed, the mean population density per square mile is 3,294. The zip codes are overwhelming non-Hispanic white (74.4%) as also supported from the individual-level characteristics. The median household income is $56,810 which is above the US average of $53,046 according to American Survey Community data from 2009-2013 (Census). Average household income is also quite high ($66,760), more than three quarters of participants are above 125% of the federal poverty level, and households average less than one vehicle. Looking across all zip codes, the mean percentage of each that is considered to be an LSA is 7.3%. It must be remembered though, that some zip codes have zero percent of land considered to be an LSA, while others have much more. The percentage of zip codes with zero LSAs is 60.4%, while 32% had one disjointed LSA. Just fewer than 8% of zip codes analyzed had two or more disjointed LSAs.

Table 4. Zip code-level characteristics

	Mean	Std. Dev.
Food Budget Share (%)	11.26	0.0047
Percentage of food away from home (%)	38.62	0.0052
Fast Food Percentage (%)	15.43	0.0033
Exercise Spending Share (%) [1]	25.00	0.0001
Population Density per square mile, 2010	3,294.48	43.1118
Percentage of each race cohorts (%): Non-Hispanic White	74.39	0.0010
Hispanic	14.77	0.0011
Non-Hispanic Black	11.36	0.0007
Median Household Income ($1,000)	56.81	0.0966
Average Household Income ($1,000)	66.76	0.1493
Per Capita Income ($1,000)	26.95	0.0574
Under 125% of Federal Poverty Line (%)	17.40	0.0450
Average number of vehicles per household	0.33	0.0026
LSA measures: Percent of land classified as LSA within zip code (%)	7.30	0.0609
Percentage of zip codes with zero disjointed LSAs	60.35	0.0020
Percentage of zip codes with one disjointed LSAs	31.98	0.0019
Percentage of zip codes with two or more disjointed LSAs	7.67	0.0011
Region: Northeast, New England (%)	4.76	0.0004
Middle Atlantic (%)	12.84	0.0014
Midwest, East North Central (%)	16.07	0.0013
West North Central (%)	7.27	0.0007
South, South Atlantic (%)	19.06	0.0013
East South Central (%)	5.54	0.0007
West South Central (%)	10.33	0.0010
West, Mountain (%)	6.67	0.0008
Pacific (%)	17.50	0.0023

1 - this includes average household spending on sports, recreation, exercise equipment, and bicycles

2 - average crime is measured at 100, indicating that the majority of residents live in zip codes where total crime levels are less than the US average

5.1.3 Analysis of differences between normal versus overweight and obese individuals

Using two-sample t-tests, I have determined the following:

Between normal weight individuals and those who are overweight or obese, a higher mean percentage of normal weight individuals meet the fruit and vegetable intake requirements (20.9% vs. 16.9% for fruits, and 12.42% vs. 10.9% for vegetables), and the difference is very significant. The average percentage of land classified as an LSA is higher among overweight and obese individuals, rather than those who are normal weight, and this is also significant. Similarly, significant relationships are found with individuals who are currently *not* employed, those who identify as non-Hispanic white, and individuals with at least a college degree. Perhaps unsurprisingly, of those who self-report their health as good, a higher percentage are considered normal weight rather than overweight or obese, as are individuals who meet federally recommended levels of physical activity, both moderate and vigorous. Average and median household incomes are higher among normal weight individuals, and they are more likely to live in zip codes with a higher density of non-Hispanic white residents. A higher percentage of overweight or obese individuals live in zip codes where a larger percentage of overall income is spent on food, though this may also be a byproduct of living in a poorer neighborhood, as poorer individuals tend to spend a higher percentage of their income on necessities such as food.

Table 5. Normal weight vs. overweight and obese individuals, by individual and neighborhood-level characteristics

	Normal Weight	Overweight / Obese	Difference
Meeting federally recommended fruit intake level (%)	20.8514	16.8576	-0.0399***
	(0.0028)	(0.0020)	
Meeting federally recommended vegetable intake level (%)	12.4273	10.9721	-0.0146***
	(0.0021)	-(0.0015)	
LSA measures by zip code: Percent of land classified as LSA in each zip code (%)	6.9625	7.4941	0.5316***
	(0.1049)	(0.0746)	
Percentage of zip codes with zero disjointed LSAs (%)	61.7062	59.5887	-.0212***
	(0.0033)	(0.0024)	
Percentage of zip codes with one disjointed LSA (%)	31.3186	32.3501	.0103***
	(0.0032)	(0.0023)	
Percentage of zip codes with two or more disjointed LSAs (%)	6.9752	8.0611	.0109***
	(0.0017)	(0.0014)	
Age (years)	43.8249	47.0160	3.1911***
	(0.1252)	(0.0830)	
Male (%)	39.7593	56.3106	0.1655***
	(0.0036)	(0.0024)	
Marital Status: Married (%)	59.258	64.8697	0.0561***
	(0.0035)	(0.0025)	
Employment: Currently employed (%)	62.3206	65.6152	0.0329***
	(0.0033)	(0.0023)	
Retired (%)	13.8445	15.4213	0.0158***
	(0.0018)	(0.0014)	
Unemployed (%)	23.8349	18.9634	-0.0487***
	(0.0032)	(0.0020)	
Household Income ($1,000)	52,217	50,150	-2,067***
	(0.0002)	(0.0001)	
Race: Non-Hispanic White (%)	73.3001	69.9266	-0.0337***
	(0.0038)	(0.0027)	
Hispanic (%)	10.8924	13.5642	0.0267***
	(0.0033)	(0.0025)	
Non-Hispanic Black (%)	6.5223	10.7651	0.0424***
	(0.0018)	(0.0016)	
Other (%)	9.2852	5.7440	-0.0354***

continued

	(0.0026)	(0.0014)	
Education: Less than high school (%)	8.0449	10.0557	0.0201***
	(0.0024)	(0.0019)	
High school (%)	50.5782	56.4124	0.0583***
	(0.0035)	(0.0025)	
College graduate (bachelor's degree or higher) (%)	41.3769	33.532	-0.0784***
	(0.0033)	(0.0023)	
Households that include children less than 18 years of age (%)	45.6407	44.1816	-0.0146***
	(0.0035)	(0.0025)	
Self-reported general health: Good (%)	88.7261	82.5124	-0.0621***
	(0.0022)	(0.0020)	
Binge Drinker (%)	16.5323	16.7523	0.0022
	(0.0028)	(0.0021)	
Smoke Status: Never smoked (%)	58.4465	53.9594	-0.0449***
	(0.0033)	(0.0025)	
Former smoker (%)	20.6568	18.9723	-0.0168***
	(0.0027)	(0.0020)	
Current smoker (%)	20.6392	26.8471	0.0621***
	(0.0025)	(0.0020)	
Meeting federally recommended level of moderate physical activity (%)	21.4199	20.0381	-0.0138***
	(0.0025)	(0.0019)	
Meeting federally recommended level of vigorous physical activity (%)	12.5054	11.2583	-0.0125***
	(0.0025)	(0.0018)	
Neighborhood-level Characteristics by zip code			
Food Budget Share (%)	11.1935	11.3012	0.1077***
	(0.0083)	(0.0056)	
Percentage of food away from home (%)	38.8174	38.5901	-0.2274***
	(0.0081)	(0.0062)	
Fast Food Percentage (%)	15.4818	15.4528	-0.0290***
	(0.0055)	(0.0042)	
Exercise Spending Share (%)	24.9482	25.0198	0.0007***
	(0.0002)	(0.0001)	
Population density per square mile, 2010	3593.4088	3125.2407	-468.168***
	(75.0489)	(52.4787)	
Population density by race: non-Hispanic white (%)	74.7743	74.18	-0.0059***
	(0.0017)	(0.0012)	
Hispanic (%)	14.2900	15.0367	0.0075***
	(0.0019)	(0.0014)	

continued

Non-Hispanic black (%)	10.325	11.9502	0.0163***
	(0.0011)	(0.0009)	
Median Household Income ($1,000)	59,426	55,333	-0.0041***
	(0.0002)	(0.0001)	
Average Household Income ($1,000)	70,112	64,865	-0.0052***
	(0.0003)	(0.0002)	
Per Capita Income ($1,000)	28,489	26,084	-0.0024***
	(0.0001)	(0.0001)	
Under 125% of Federal Poverty Line (%)	16.5400	17.8908	-1.3508***
	(0.0767)	(0.0554)	

5.1.4 Limited Supermarket Access areas

This section will attempt to describe the characteristics that are more likely to define a neighborhood in which there is an LSA, and individuals who are more or less likely to live in them.

As shown in Table 6, the average percentage of land classified as an LSA is just under 7% for normal weight individuals, while the average percentage of land classified as an LSA for overweight and obese individuals is 7.5%. This difference is also statistically significant.

Table 6. Average percentage of land classified as LSA by weight classifications

	Mean	Std. Error
Normal Weight	6.96	0.1049
Overweight / Obese	7.49	0.075
Difference	0.53	0.0166 (***)

Furthermore, Table 7 shows the obesity prevalence by the number of disjointed LSAs. As expected, if there are zero LSAs within a zip code, the prevalence of overweight or obese individuals remains at 63%. Of zip codes with one LSA, the prevalence of overweight and obese increases to 64.6%, and in zip codes with two or more disjointed LSAs, the prevalence increases to 67.1%. When comparing between zero LSAs and one LSA in a zip code, and one LSA vs. two ore more LSAs, the differences are statistically significant.

Table 7. Prevalence of overweight/obese individuals by number of LSAs within zip code

	Zero LSAs within zip code	One LSA within zip code	Two or more LSAs within zip code
Mean	63.04	64.6	67.12
Std. Error	0.0026	0.0034	0.0068
	N_LSA=1 vs N_LSA=0	N_LSA≥2 vs N_LSA=1	N_LSA=0 vs N_LSA≥2
t-test	-0.0158***	0.0264***	-0.0422***

Referencing Table 8, the data indicates that increasing the number of LSAs in a zip code from zero to one makes it more difficult for individuals to meet the federally recommended level of fruit consumption, but the difference is not statistically significant. As the number of LSAs increases, the change in mean percentage of individuals who meet the vegetable intake requirements is negligible. Additionally, a higher percentage of land classified as an LSA has a negative relationship with compliance in meeting fruit recommendations. There is no difference between compliance and non-compliance in terms of vegetable consumption when looking at percentage of land classified as an LSA. Please refer to the results in Table 9.

Table 8. Fruit and vegetable intake by number of LSAs within zip code

Fruit			
	Zero LSAs	One LSA	Two or more LSAs
Mean	18.52	17.96	18.01
Std. Error	0.0022	0.0026	0.0062
	N_LSA=1 vs N_LSA=0	N_LSA≥2 vs N_LSA=1	N_LSA=0 vs N_LSA≥2
t-test	0.0055	0.0004	0.0051*
Vegetables			
	Zero LSAs	One LSA	Two or more LSAs
Mean	11.56	11.66	10.33
Std. Error	0.0016	0.0022	0.0041
	N_LSA=1 vs N_LSA=0	N_LSA≥2 vs N_LSA=1	N_LSA=0 vs N_LSA≥2
t-test	-0.001	-0.0132***	0.0122***

Table 9. Average percentage of land classified as LSA by compliance with federally recommended levels of fruits and vegetables

	Fruit Intake		**Vegetable Intake**	
	Mean	**St. Error**	**Mean**	**St. Error**
Not compliant	7.40	0.0670	7.30	0.0646
Compliant	6.85	0.1453	7.30	0.1823
Difference	-0.5589	16.00 (***)	-.0041	19.37

In terms of individual-level characteristics, those most likely to live in a neighborhood with an LSA appear to be male, slightly older than those who live in neighborhoods *without* an LSA, married, currently employed or retired, and have at least a college degree; they are also mainly non-Hispanic white. Household income also tends to be higher, as does individuals who report themselves to be in good health.

In neighborhoods with at least one LSA, the share of household income spent on food is lower than neighborhoods with zero LSAs, as is percentage spent on fast food. Neighborhoods with at least one LSA tend to be overwhelmingly non-Hispanic white and to have a lower population density.

Table 10. Limited Supermarket Access area, by individual- and neighborhood-level characteristics

Limited supermarket access area, individual-level characteristics			
	N_LSA = 0	N_LSA ≥ 1	Difference
Male (%)	49.64	50.78	0.0113***
	(0.0031)	(0.0026)	
Age (years)	45.66	46.00	0.3349***
	(0.1084)	(0.0921)	
Marital Status: Married (%)	59.96	64.73	0.0478***
	(0.0032)	(0.0027)	
Education: Less than high school (%)	10.35	8.66	-0.0169***
	(0.0024)	(0.0019)	
High school (%)	55.3114	53.6412	-0.0167
	(0.0031)	(0.0026)	
College graduate (bachelor's degree or higher) (%)	34.34	37.70	0.0336***
	(0.0028)	(0.0025)	
Employment: Currently employed (%)	63.53	65.01	0.0149***
	(0.0029)	(0.0024)	
Retired (%)	14.85	14.85	-0.0000
	(0.0017)	(0.0014)	
Unemployed (%)	21.62	20.14	-0.0150***
	(0.0027)	(0.0022)	
			continued
Household Income ($1,000)	48,750	52,307	3,557***
	(0.1627)	(0.1424)	
Race: Non-Hispanic white (%)	66.43	74.24	0.0781***
	(0.0034)	(0.0030)	

Hispanic (%)	14.39	11.42	-0.0296***
	(0.0032)	(0.0026)	
Non-Hispanic black (%)	12.93	6.80	-0.0613***
	(0.0022)	(0.0014)	
Other race (%)	6.25	7.53	0.0129***
	(0.0018)	(0.0018)	
Self-reported general health: Good (%)	83.42	85.64	0.0221***
	(0.0025)	(0.0019)	
Poor (%)	16.58	14.36	-0.0221***
	(0.0025)	(0.0019)	
Smoke status: Former smoker (%)	20.6	18.92	-0.0168***
	(0.0026)	(0.0021)	
Current smoker (%)	24.1	24.93	0.0084***
	(0.0024)	(0.0021)	
Never smoked (%)	55.03	55.95	0.0092***
	(0.0031)	(0.0026)	
Binge Drinker (%)	16.96	16.48	-0.0048
	(0.0027)	(0.0021)	
Zip code-level characteristics			
	N_LSA = 0	N_LSA ≥ 1	Difference
Food Budget Share (%)	11.41	11.17	-0.2340***
	(0.0076)	(0.0058)	
Percentage of food away from home (%)	38.66	38.68	0.0120***
	(0.0065)	(0.0070)	
Fast food percentage (%)	15.57	15.39	-0.1832***
	(0.0051)	(0.0043)	
Exercise spending share (%)	24.92	25.04	0.0012***
	(0.0001)	(0.0001)	
Population density per square mile, 2010	3406.68	3220.77	-185.9006**
	(61.5802)	(58.8915)	
			continued
Population density by race: Non-Hispanic white (%)	70.32	77.07	0.0675***
	(0.0016)	(0.0013)	
Hispanic (%)	16.93	13.35	-0.0358***

	(0.0019)	(0.0014)	
Non-Hispanic black (%)	15.44	8.69	-0.0675***
	(0.0013)	(0.0007)	
Median household income ($1000)	51,975	59,991	8.0156***
	(0.1294)	(0.1303)	
Average household income ($1000)	60,276	71,023	10.7471***
	0.1960	0.2065	
Per capita income ($1000)	24,779	28,371	3.592***
	0.0749	0.0798	
Below 125% of Federal Poverty Line (%)	20.1945	15.5686	-4.6259***
	0.0801	0.0488	

5.2 Regression Results

Weight status as dependent variable

Four models were created to determine the effect that both individual-level characteristics and neighborhood-level characteristics have on overweight/obesity, the main dependent variable. The estimation results are presented in Appendix #3. Table 11 summarizes the marginal effects of independent variables on the risk of being overweight or obese after the discussion below.

Model A1 includes P_LSA, the percentage of land classified as an LSA in each zip code, as well as individual-level demographic characteristics; the regional difference is represented by the regional dummies, and no zip code-level variables were included. Many of the individual-level variables behaved as expected, though the percentage of land classified as an LSA was not found to be significant in any of the four models. Individuals who identify as male, as well as those who are married, are more likely to be

overweight or obese; this relationship is very significant. With each one year increase in age, the propensity to become overweight or obese increases by 0.4 percentage points. Those with only a high school education are more likely to become overweight or obese in comparison to those with at least a bachelor's degree; the coefficient for having at least a bachelor's degree is negative, indicating that as education level increases, the risk of becoming overweight or obese decreases. Employed individuals are more likely to be obese than their unemployed counterparts, though retirees are less likely than those who are unemployed to be overweight or obese. Both Hispanics and non-Hispanic blacks are more likely than non-Hispanic whites to be overweight (6.4 and 12.6 percentage points, respectively). Household income appears to have an inverse relationship with weight status; as income increases, the risk of being overweight or obese decreases.

Quite expectedly, individuals who self-reported their health as "good" have a decreased probability of becoming overweight or obese, as do individuals who meet at least the moderate level of federally recommended physical activity. Similarly, those who self-reported meeting the federally recommended daily consumption levels of fruits and vegetables were less likely than those who did not to be overweight or obese. Both former smokers and non-smokers are less likely than their smoking peers to be overweight or obese, and those who consume enough alcohol to be considered a binge drinker have a higher probability than their normal drinking companions, though the relationship is only significant at the .05% level. Unless other noted, all of the variables mentioned in relation to model one were found to very significant, with the exception of high school.

Model B1 includes the same variables as model one, and additionally includes N_LSA, the number of disjointed LSAs in a zip code, where N_LSA_2 is the base. The second model measures the effects of the number of LSAs, conditional on the same percentage of land classified as LSAs. This model produced very similar results as the first. When comparing neighborhoods with zero LSAs to those with two are more, individuals are 2.1 percentage points less likely to be obese and this is very significant; however, if you increase the number of LSAs to one and compare to two or more, it becomes 1.3 percentage points, and this variable is not significant. Again, this is holding the percentage of land classified as an LSA constant.

Model C1 includes the percentage of land classified as an LSA, however, we have added neighborhood-level characteristics in additional to the regional dummies. The relationships that have been previously mentioned have remained the same, though the risks have altered slightly which I will discuss here. Once neighborhood-level characteristics are added in, individuals with at least a bachelor's degree are less likely than those with no education to be overweight or obese; however, the risk has slightly lessened and is now only 5.0 percentage points less rather than 6.2 percentage points less. Retirees are even less likely than their unemployed counterparts to be overweight or obese (6.1% vs 5.8%). Household income, though still showing an inverse relationship, is no longer significant. Those in good health, those who exercise frequently, those who meet the federally recommended fruit and vegetable consumption levels, and those who are former or non-smokers are still less likely to overweight or obese, and the risk is the

same as models one and two. Once neighborhood characteristics are added in, binge drinkers are slightly more likely to be overweight or obese (14 percentage points vs. 12 percentage points).

For the neighborhood average, as the share of food budgets spent on fast food increases, the likelihood of becoming obese does, too. Unexpectedly, as the share of food budgets spent on food away from the home (not including fast food) increases, the likelihood of being overweight or obese decreases. This may be associated with increased income, and the increased probability of having luxury to exercise. Perhaps wealthier individuals who eat out do so purely for their entertainment and luxury, at finer and perhaps healthier establishments, and may also still practice healthy lifestyles otherwise. Neighborhoods that have a higher density of non-Hispanic black and Hispanic individuals are more likely to lead to increased weight status when compared with neighborhoods that are predominantly non-Hispanic white, though these relationships are not significant. As median household income increases, the probability of being overweight or obese decreases, similar to the relationship of individual household income. Finally, when compared to individuals who live above 125 percent of the poverty line, those who live below are less likely to be overweight or obese. This holds true for Model D1, as well.

Model D1, similarly, includes all of the variables from model three, however the number of disjointed LSAs has been added to measure the effect of an increase in disjointed LSAs, while holding the percentage of land classified as an LSA constant. Individual-level variable results remain the same, as do the noted neighborhood-level variables.

Keeping the percentage of land classified as an LSA constant, individuals who reside in a neighborhood with zero LSAs are 1.7% less likely to be overweight or obese when compared to those living in neighborhoods with two or more disjointed LSAs. If the number of disjointed LSAs is increased to one, individuals are still less likely than those living with two or more disjointed LSAs, though the likelihood is a bit less (1.2%). The relationship between zero and two or more disjointed LSAs is significant, though the relationship between one and two or more is not. Again, this is assuming the percentage of land classified as an LSA remains constant. It also is worth mentioning that individuals who reported living in the center of a metropolitan statistical area (MSA) are less likely than those *not* living in an MSA or those residing outside the center city of an MSA but inside the county containing the center city, as well as those residing inside a suburban county of the MSA or an MSA that has no center city. This relationship is significant. Interestingly, after controlling for all other variables, those that live in east south central region of the United States have the highest probability of being overweight or obese when compared with all of the other regions of the country. This finding holds true in each of the four models discussed here.

Table 11

Marginal Effects of Overweight/Obese

	A1	B1	C1	D1
Gender: Male	0.159***	0.159***	0.160***	0.160***
	0.004	0.004	0.004	0.004
Age (years)	0.004***	0.004***	0.004***	0.004***
	0	0	0	0
Married	0.055***	0.056***	0.049***	0.049***
	0.005	0.005	0.005	0.005
Variables for education (base = education less than high school graduate)				
High School Graduate	0.010	0.011	0.013	0.013
	0.009	0.009	0.01	0.01
College Graduate	-0.063***	-0.063***	-0.050***	-0.050***
	0.01	0.01	0.01	0.01
Variables for employment (base = currently unemployed)				
Currently Employed	0.057***	0.057***	0.052***	0.052***
	0.006	0.006	0.006	0.006
Retired	-0.058***	-0.058***	-0.061***	-0.061***
	0.007	0.007	0.007	0.007
Variables for race (base = non-Hispanic white)				
Hispanic	0.064***	0.063***	0.069***	0.069***
	0.008	0.008	0.009	0.009
Non-Hispanic Black	0.126***	0.125***	0.128***	0.128***
	0.007	0.007	0.008	0.008
Other Race	-0.090***	-0.090***	-0.072***	-0.072***
	0.01	0.01	0.011	0.011
With Children in House (Yes = 1; No = 0)	0.023***	0.023***	0.024***	0.023***
	0.005	0.005	0.005	0.005
Household Income ($1,000)	-0.001***	-0.001***	0	0
	0	0	0	0
Self-reported Good Health (Good=1; Poor=0)	-0.085***	-0.085***	-0.084***	-0.084***
	0.006	0.006	0.006	0.006
Meets the requirement for moderate physical activity	-0.012***	-0.012***	-0.013***	-0.013***
	0.005	0.005	0.005	0.005

continued

Meets the requirements for vigorous physical activity	-0.028***	-0.028***	-0.026***	-0.026***
	0.007	0.007	0.007	0.007
Meets fruit/vegetable requirements	-0.025**	-0.025**	-0.023**	-0.023**
	0.010	0.010	0.011	0.011
Variables for Smokers (base = current smoker)				
Former Smoker	-0.075***	-0.076***	-0.079***	-0.079***
	0.006	0.006	0.006	0.006
Non-Smoker	-0.037***	-0.037***	-0.038***	-0.038***
	0.005	0.005	0.005	0.005
Binge Drinker	0.012**	0.012**	0.014**	0.014**
	0.006	0.006	0.006	0.006
Variables for region (base = Pacific Census region)				
Northeast	0.020**	0.018**	0.016**	0.015**
	0.008	0.008	0.009	0.009
Middle Atlantic	0.033***	0.031***	0.032***	0.031***
	0.009	0.009	0.01	0.01
East North Central	0.043***	0.042***	0.019**	0.018**
	0.008	0.008	0.009	0.009
West North Central	0.049***	0.049***	0.020**	0.019**
	0.008	0.008	0.009	0.009
South Atlantic	0.035***	0.035***	0.011	0.011
	0.008	0.008	0.009	0.009
East South Central	0.078***	0.078***	0.041***	0.041***
	0.009	0.009	0.01	0.01
West South Central	0.054***	0.052***	0.024**	0.023**
	0.008	0.008	0.01	0.01
Mountain	0.01	0.009	-0.007	-0.008
	0.009	0.009	0.009	0.009
Variables for MSA (Metropolitan Statistical Area) (base = living in a non-MSA area)				
Living in a MSA city			-0.017***	-0.018***
			0.006	0.006
Other			0.007	0.006
			0.005	0.005
Variables for number of LSAs in the zip code (base = two disjointed LSAs)				
Zero LSAs		-0.021***		-0.017**

continued

		0.008		0.008
One LSA		-0.013		-0.012
		0.008		0.008
Percentage of land classified as LSA	0	0	0	0
	0	0	0	0
Fast Food Percentage			0.021***	0.022***
			0.007	0.007
Dining Out Percentage			-0.019***	-0.020***
			0.004	0.004
Variables for race prevalence within zip code (base = predominantly non-Hispanic white zip code)				
Predominantly Hispanic			0.007	0.006
			0.029	0.029
Predominantly non-Hispanic Black			0.019	0.015
			0.016	0.017
Predominantly Other Race			-0.039	-0.039
			0.027	0.027
Median Household Income			-0.001***	-0.001***
			0.216	0.217
Population Density			-0.000***	-0.000***
			0	0
Below 125% of Poverty Line			-0.001***	-0.001***
			0	0
* p<0.10		** p<0.05	*** p<0.01	

Next we look how individual-level and neighborhood-level characteristics affect the ability to meet the federally recommended level of fruit and vegetable consumption.

Fruit and vegetable consumption requirement as dependent variable

The dependent variable is equal to one if an individual meets the recommended consumption level of both fruit and vegetables. Otherwise, it equals zero. The estimation

results are summarized in the Appendix. Table 11 shows the marginal effects of independent variables on the probability of meeting the recommended level of fruits and vegetables.

Four models with the same specifications were estimated to measure the effect of both individual-level and neighborhood-level characteristics on the probability of meeting the federally recommended level of fruit and vegetable intake levels.

Model A2 includes the percentage of land classified as an LSA and individual-level variables; no zip code-level variables were included. Many of the variables behaved as expected. While an increase in the percentage of land classified as an LSA causes a decrease in the likelihood of meeting the fruit and vegetable consumption requirements, the probability is negligible and the relationship is not significant. However, those that identify as male and those who are married are less likely to meet requirements as the percentage of LSA land increases. These relationships are significant at the 0.051% and 0.05% level, respectively. Controlling for education and employment status, those who have at least a bachelor's degree are still more likely to meet the intake requirements than their lesser-educated counterparts, as are retirees when compared with those who are unemployed. Individuals who are non-Hispanic black or Hispanic will find it more difficult than non-Hispanic whites to meet the fruit and vegetable intake requirement as the percentage of LSA land increases, and these relationships are significant at the .01% and 0.05% levels, respectively. Oddly, those who self-report themselves in good health will find it more difficult than those who self-report poor health to meet the intake levels;

however, this relationship is not significant. Finally, those who meet the recommended levels of physical activity, both moderate and vigorous, have a higher probability of meeting the fruit and vegetable intake requirement despite the increase in percentage of LSA land, while former smokers and binge drinkers are less likely.

Model B2 includes the same variables as model one, and additionally includes the number of disjointed LSAs in a zip code (N_LSA), where having two or more LSAs is the base. This model measures the effects of an increase in the number of LSAs, conditional on the percentage of land classified as an LSA remaining constant. Both models include regional dummy variables. Comparing one disjointed LSA to two or more, those living in a neighborhood with only one are more likely to meet the intake requirements, and this relationship is significant at the 0.1% level. The relationship between two disjointed LSAs and zero, holding the percentage of LSA constant, is not significant, and inconclusive as the results indicate that when comparing each to two or more disjointed LSAs, those who reside in a neighborhood with one LSA have a higher probability of meeting requirements that those residing in a neighborhood with zero LSAs. This alludes to several limitations that occurred with the data; further research is necessary to look at much smaller land areas as zip codes are quite large, and additionally, further research is necessary to address the possibility of zip code residents traveling across zip code borders for better food access. These limitations will further be discussed in Chapter 6. These results indicate that if you hold the percentage of land classified as an LSA constant, the number of disjointed LSAs does not affect the probability of meeting fruit and vegetable intake requirements when controlling for

individual-level and neighborhood-level characteristics, unless you are comparing an neighborhood with one disjointed LSA to a neighborhood with two or more disjointed LSAs.

Model C2 is similar to model A2, as it only includes the percentage of land classified as an LSA, however, we have added neighborhood-level characteristics. The relationships that have been previously mentioned have remained the same. Interestingly, after controlling for all other variables, those that live in east south central part of the United States have the highest probability of not meeting intake requirements, when compared with all of the other regions of the country. This finding holds true in each of the four models discussed here. Once neighborhood-level characteristics are added in, it appears that neighborhoods with a higher Hispanic population will have a more difficult time meeting the requirements, as will neighborhoods that are not predominantly non-Hispanic white, Hispanic, or non-Hispanic black. In terms of education, those with at least a bachelor's degree are still more likely to meet the intake requirement levels of fruits and vegetables, but the probability is smaller than the previous model (1.2% vs. 1.3%). Finally, neighborhoods with a higher median household income will have a higher probability of meeting the intake requirements.

Model D2, again, similarly, includes all of the variables from model three, however the number of disjointed LSAs has been added to measure the effect of an increase in disjointed LSAs, while holding the percentage of land classified as an LSA constant. Individual-level variable marginal effect results remain the same, as do the noted

neighborhood-level variables, with the exception of median household income, which increased the probability of meeting the fruit and vegetable intake requirements slightly. Keeping the percentage of land classified as an LSA constant, the relationship with the number of disjointed LSAs remains unchanged from model three. In both the third and fourth models, an increase in the percentage of land classified as an LSA causes a decrease in the likelihood of meeting the fruit and vegetable consumption requirements for those residing in the center of an MSA, but the probability is negligible and the relationship is not significant.

Finally, the aggregate of these results present a clear picture of who is more likely to become overweight or obese, and who is more or less likely to meet the federally recommended fruit and vegetable intake levels when controlling for individual-level and neighborhood-level characteristics. The results will be further discussed in the conclusion, as will potential policy implications, and general limitations with the data and study.

Table 12 – Marginal Effects

Marginal Effects of Meeting Fruit and Vegetable Requirements				
	A2	**B2**	**C2**	**D2**
Male	-0.028***	-0.028***	-0.028***	-0.028***
	0.002	0.002	0.002	0.002
Age (years)	-0.000**	-0.000**	-0.000**	-0.000**
	0	0	0	0
Married	-0.002	-0.002	-0.002	-0.002
	0.002	0.002	0.002	0.002
Variables for education (base = education less than high school graduate)				
High School	-0.001	-0.001	-0.002	-0.002
	0.004	0.004	0.004	0.004
College Graduate	0.013***	0.013***	0.012***	0.012***
	0.004	0.004	0.004	0.004
Variables for employment (base = currently unemployed)				
Currently Employed	-0.001	-0.001	-0.001	-0.001
	0.002	0.002	0.002	0.002
Retired	0.001	0.001	0.002	0.002
	0.002	0.002	0.002	0.002
Variables for race (base = non-Hispanic white)				
Hispanic	-0.006**	-0.006**	-0.006*	-0.006*
	0.003	0.003	0.003	0.003
Non-Hispanic Black	-0.006***	-0.006***	-0.006**	-0.006**
	0.002	0.002	0.003	0.003
Other Race	-0.003	-0.003	-0.001	-0.001
	0.003	0.003	0.003	0.003
With Children in House (Yes = 1; No = 0)	-0.002	-0.002	-0.002	-0.002
	0.002	0.002	0.002	0.002
Household Income ($1,000)	0.000***	0.000***	0.000***	0.000***
	0	0	0	0
Self-reported Good Health (Good=1; Poor=0)	-0.001	-0.001	-0.001	-0.001
	0.002	0.002	0.003	0.002

continued

Meets the requirement for moderate physical activity	0.003*	0.003*	0.003*	0.003*
	0.001	0.001	0.001	0.001
Meets the requirements for vigorous physical activity	0.010***	0.010***	0.010***	0.010***
	0.003	0.003	0.003	0.003
Variables for Smokers (base = current smoker)				
Former Smoker	-0.011***	-0.011***	-0.011***	-0.011***
	0.002	0.002	0.002	0.002
Non-Smoker	0	0	0	0
	0.002	0.002	0.002	0.002
Binge Drinker	-0.005**	-0.005**	-0.005**	-0.005**
	0.002	0.002	0.002	0.002
Variables for region (base = Pacific Census region)				
Northeast	-0.008***	-0.008***	-0.010***	-0.010***
	0.002	0.002	0.002	0.002
Middle Atlantic	-0.007***	-0.007***	-0.010***	-0.010***
	0.003	0.002	0.003	0.003
East North Central	-0.011***	-0.011***	-0.013***	-0.013***
	0.002	0.002	0.002	0.002
West North Central	-0.015***	-0.015***	-0.016***	-0.016***
	0.002	0.002	0.002	0.002
South Atlantic	-0.011***	-0.011***	-0.013***	-0.013***
	0.002	0.002	0.002	0.002
East South Central	-0.020***	-0.020***	-0.021***	-0.021***
	0.002	0.002	0.002	0.002
West South Central	-0.010***	-0.010***	-0.012***	-0.012***
	0.002	0.002	0.002	0.002
Mountain	-0.008***	-0.008***	-0.009***	-0.009***
	0.002	0.002	0.002	0.002

continued

Variables for MSA (Metropolitan Statistical Area) (base = living in a non-MSA area)				
Living in a MSA city			0	0
			0.002	0.002
Other			0	0
			0.002	0.002
Variables for number of LSAs (base = two disjointed LSAs)				
Zero LSAs		0.003		0.003
		0.003		0.003
One LSA		0.005*		0.005*
		0.003		0.003
Percentage of land classified as an LSA	0	0	0	0
	0	0	0	0
Fast food percentage			0.003	0.003
			0.003	0.003
Dining out percentage			-0.001	0
			0.002	0.002
Variables for race prevalence within zip code (base = predominantly non-Hispanic white zip code)				
Predominantly Hispanic			-0.021**	-0.021**
			0.01	0.01
Predominantly non-Hispanic Black			-0.003	-0.002
			0.005	0.006
Predominantly Other Race			-0.020**	-0.020**
			0.009	0.009
Median Household Income			0.000*	0.000*
			0	0
Population Density, 2010			0	0
			0	0
Below 125% of poverty Line			0	0
			0	0

* p<0.10, ** p<0.05, *** p<0.01

In all of the models estimated, the percentage of land classified as an LSA was not found to be significant once individual- and neighborhood-level variables were controlled for. However, the number of disjointed LSAs within a zip code, holding the percentage of land classified as an LSA constant, was found to be significant. In general, as the number of disjointed LSAs increases from zero to one, there is a higher likelihood of overweight/obesity status. Conversely, when comparing zero disjointed LSAs and two or more disjointed LSAs to one disjointed LSA and two or more disjointed LSAs, those who live in the latter are more likely to meet the federally recommended levels of fruit and vegetables. This relationship is significant, though zero disjointed LSAs to two or more is not.

Chapter 6 Conclusion

The main research question this thesis hoped to be answered by the data is how living in a limited supermarket access affects weight status and the ability to meet federal recommendations on fruit and vegetable requirements. Using BRFSS survey data, as well as TRF's Limited Supermarket Access study, and data from the latest census data, American Community Survey, and Esri's Consumer Expenditures Survey, we find that an increase in the percentage of land classified as an LSA is shown to have a negative effect on overweight/obese risk with and without controlling for regional differences indicated by regional dummies and/or neighborhood-level characteristics, however, this relationship is not significant in any of the models. We also find, though, that conditional on the same percentage of land classified as an LSA, the number of disjointed LSAs within a neighborhood does have significant effects on overweight/obesity status, as well as the probability of being overweight/obese, and the probability of meeting the recommendations for fruit and vegetable intake levels. When comparing zero disjointed LSAs to two or more within a neighborhood, individuals are 2.1 percentage points less likely to be overweight/obese than those living in a neighborhood with two or more disjointed LSAs. Further, when comparing one disjointed LSA to a neighborhood with two or more disjointed LSAs, the likelihood of being overweight/obese is 1.3 percentage points lower. While living in a neighborhood with zero LSAs is still ideal for a lower weight status, having one disjointed LSA is more conducive to a favorable weight status than two or more disjointed LSAs.

Interestingly, after controlling for all other variables, those that live in east south central part of the United States have the highest risk of being overweight or obese, while individuals residing in the center city of an MSA have the lowest risk of being overweight or obese. If you refer back to Figure 1, you will notice that a lower percentage of individuals living in states in the east south central region of the Unites States consume the recommended intake for fruits and vegetables individuals than when compared with states in the rest of the country.

Healthy eating habits and the maintenance of a healthy lifestyle have been established as essential to obesity prevention, and previous research has shown that one's environment must therefore be conducive to maintaining these things. In some cases, a person's environment may be quite detrimental in terms of being conducive to making healthy choices (Poston and Foreyt 1999; French, Story, and Jeffery 2001; Lopez 2007). Since the term "obesogenic" environment had been coined by researchers to discuss the environmental factors within neighborhoods that may lead to an increase in weight status (Glanz et al. 2005; Ford and Dzewaltowski 2008), researchers have been clamoring to better understand the neighborhood food environment.

This thesis further found that both individual-level characteristics, as well as neighborhood-level characteristics have an effect on both overweight/obesity risk and the likelihood of meeting fruit and vegetable consumption requirements, which has been determined to be an essential piece of maintaining a healthy weight. Of the individual-level characteristics that have a relationship to weight status, the most significant variables include gender, age, marital status, education, race, household income, self-

reported health, physical activity level, smoker and binge drinker status, as well as the region of the country in which individuals reside. Perhaps intuitively, individuals who self-reported meeting the fruit and vegetable requirements were significantly less likely to be overweight or obese.

Further, in models were meeting fruit and vegetable requirements is the dependent variable, these same variables have a significant impact on the likelihood of individuals meeting fruit and vegetable consumption requirements.

When only controlling for individual-level characteristics and determining impact on weight status, it is apparent that men and those who are married are more likely to become overweight or obese (Charlotte A. Schoenborn 2002). Further, as age increases, so does the likelihood overweight/obesity risk, which supports findings from Sobal and Rauschenbach (2003)and Schoenborn (2002). The CDC notes that obesity rates are lower for adults over 60 years of age (CDC: Division of Nutrition 2015), and though this thesis does not calculate obesity statistics by age cohort, findings do indicate that retirees have a lower risk for obesity; retirees tend to be individuals over the age of 60.

Increased education, the data indicates, appears to have an inverse relationship with overweight/obese, meaning that the likelihood of overweight and obesity decreases as education level increases. This is similar to findings from Schoenborn (2002)and Ogden (Ogden CL 2010), whose research showed that in general, the prevalence of obesity decreases as education increases, and Giles-Corti et al. (Giles-Corti et al. 2003)who

found that the odds of being obese were nearly three times higher in those with only a secondary education compared with respondents who indicated they had at least a college degree.

Regression results indicate that there is an inverse relationship between income and weight, but this is only thinly supported by prior research. Once neighborhood-level characteristics were added into the model, income was no longer found to be significant. Chang and Lauderdale (Chang and Lauderdale 2005) found that non-Hispanic white and non-Hispanic black women show inverse relationships between income and weight status, regardless of beginning level of income, as does Ogden (Ogden CL 2010). Additionally, results do indicate that when controlling for neighborhood characteristics, individuals who reside in lower-income neighborhoods are more likely to be overweight/obese, and less likely to meet fruit and vegetable consumption requirements. This agrees with Edmonds et al. (Edmonds et al. 2001), who found a negative association with availability of fruits and vegetables in the home in low-income households. When controlling for race, being a race other than non-Hispanic white increases the likelihood of being overweight/obese, and further decreases the likelihood of meeting fruit and vegetable requirements.

While prior research has provided mixed results for the relationship between smoking and obesity, this thesis finds that non-smokers and former smokers are less likely to be obese when compared with those who currently smoke; this finding differs from prior research from Patel et al. (2011) and Dare, Mackay, and Pell (2015). Additionally,

individuals who are binge drinkers are more likely to be obese, however, this is unclear if it is simply a result of extra caloric intake, or other genetic and environmental and lifestyle factors.

Using fruit and vegetable consumption as the dependent variable produced incredibly similar relationships to the models that used overweight/obese as the dependent variable. As expected, men and older individuals are less likely to meet the fruit and vegetable requirements. The coefficient still remains the same for marital status, indicating that married individuals are less likely to meet the requirements, however, the relationship is no longer significant.

Education is still a significant variable, firmly establishing that those with higher levels of education are more likely to meet fruit and vegetable requirements. Employment status is no longer a significant characteristic in these models. Race, as expected is still very significant, as those who do not identify as non-Hispanic white have a much more difficult time of consuming fruits and vegetables.

Though not significant throughout all of the overweight/obese models, income is significant through each of the fruit and vegetable models. As expected, as income increases, so does the likelihood of meeting fruit and vegetable requirements.

Median household income, calculated as the median for an entire neighborhood, is associated with a risk of obesity. This supports research conducted by Lopez (2007), who

also used BRFSS data. Further, as neighborhood median income increases, so does the intake of healthier food items (Diez-Roux et al. 1999), a finding this thesis also supports. Diez-Roux et al. found that as individual income increases, so does the consumption of healthy food items (1999), .

Policy Implications

The research findings indicate that individuals are more likely to meet the federally recommended consumption of fruits and vegetables if they live in a neighborhood where zero percentage of land is classified as an LSA. It is obvious though, that that is not the case in many neighborhoods. The data further indicates that in the case of fruits and vegetable consumption, individuals are better off living in a neighborhood with only one disjointed LSA, versus two or more. From an intervention standpoint, when considering resources, it can be incredibly resource-intensive to eradicate all LSAs from neighborhoods. Perhaps though, if supermarkets are more strategically placed in neighborhoods, access can be granted to more individuals, thus enhancing the likelihood of meeting the recommended consumption of fruits and vegetables. By providing a variety of fresh fruits and vegetables, in conjunction with health and nutrition education, we may be able to lower overall weight status and increase individual's likelihood of meeting fruit and vegetable intake recommendations. On the other hand, as the number of LSAs increase, the risk of being overweight or obese also increases. Thus, decreasing the number of LSAs in a neighborhood will help to fight against the risk of overweight/obesity. The findings suggest the importance of providing supermarkets to the LSA areas as it sheds light on where to establish supermarkets if resources permit. That is, a supermarket added to an LSA neighborhood may break up LSA areas and could

potentially increase the likelihood of meeting the recommended consumption of fruits and vegetables.

The USDHHS notes that a main goal and objective is to promote health and reduce chronic disease through the consumption of healthful diets and achievement and maintenance of healthy body weights. It seems that the first place to start is with the food that people consume; food that is mainly acquired through going to full-service supermarkets. As poor diet has been established as an important factor contributing to overweight and obesity, it is imperative that the fundamental way in which the majority of people acquire food in the Unites States must be overhauled.

Without intervention of some sort, the obesity epidemic will continue, making people and sick and remaining a burden on our already over-burdened healthcare system. Overweight and obesity is preventable! Let's do something about it.

Education is also imperative to reducing the epidemic. What makes up a healthy diet? How often should one exercise? How does one go about exercising? How do you cook healthy, delicious tasting food that is, most importantly, healthy for your body?

Limitations

Using zip codes as a proxy for neighborhoods may be problematic because they can vary vastly in size; for example, the average size of a zip code is 90 square miles. Conditions and amenities may vary greatly within a single zip code, and the food environment may

not necessarily be reflected appropriately. Ideally, census block groups would be used; indeed, TRF provides LSA data by block group level, however, BRFSS data is by zip code only. In addition to the issue of not having a small enough blocks of land to adequately assess supermarket access, as well as neighborhood-level characteristics, this thesis did not address the possibility of individuals living in one zip code yet traveling to an adjacent or nearby zip code to purchase groceries or dine away from the home. If this type of data were available, results may be quite different.

Because of the geographic limitations noted above, it was essential to aggregate the LSA data, originally at the block group level, up to zip codes—a much larger land area. The percentages of LSAs were then added together for the zip code, and for this calculation it was necessary to assume a uniform distribution of the population. This is a limitation because this may not be the most accurate determinant of whether an individual lives in an LSA or does not. Further, the datasets used for this thesis are from different years, and this may cause issues in interpreting the results. For example, TRF's LSA data is from 2010, but we are using the 2007 BRFSS survey. Neighborhood attributes and amenities may have been quite different in 2007, yet TRF's data describes the conditions in 2010. The data also presented a mainly non-Hispanic white population, but the tendency is for Hispanic and non-Hispanic blacks to have a higher prevalence of overweight/obesity (Charlotte A. Schoenborn 2002; CDC: Division of Nutrition 2015)

It is also quite possible that there are inaccuracies with the use of self-reported height and weight (Lopez 2007), as it tends to be human nature to under- or over-estimate both.

Similar to Lopez, this thesis categorized overweight/obesity as a dependent variable rather than using BMI as a continuous variable. While federal guidelines were used to determine the BMI cutoffs for overweight and obesity, it could be argued that there is little difference in health and weight risks for people just below and just above the cutoff points (Lopez 2007). Additionally, already overweight and obese individuals may self-select themselves into neighborhoods that negatively impact their weight status, rather than the other way around. As such, it must be remembered that the outcomes stated in this thesis are solely statistical outcomes, and they do not necessarily imply causation. Results should be interpreted with caution.

Appendix

Appendix A
Data Descriptions

(1) Centers for Disease Control and Prevention Behavioral Risk Factor Surveillance System (BRFSS), 2007

The full questionnaire can be accessed here:
http://www.cdc.gov/brfss/questionnaires/pdf-ques/2007brfss.pdf

(2) The Reinvestment Fund's Limited Supermarket Access data, 2010
Limited Access Supermarkets

As noted, the original LSA dataset was available only at the block group level. In order to analyze the LSA data in conjunction with each of the additional datasets mentioned, it was necessary to aggregate the data to the zip code level. With the help of Jim Trimble, a Database Administrator for the Center for Remote Sensing and Spatial Analysis at Rutgers University, this was possible. Below is the methodology for which this process was completed.

Zip code layer in ArcGIS:
From Jim Trimble:
System Support Specialist, *Center for Remote Sensing and Spatial Analysis*
Dept. of Ecology, Evolution, & Natural Resources
ZIP CODE INFO
==========
Title: U.S. ZIP Code Areas (Five-Digit)

Publication date: 2014-07-01

Summary: U.S. ZIP Code Areas (Five-Digit) provides area, postal district name, and 2010 Census demographic information for the ZIP Code areas in the United States.

Description: U.S. ZIP Code Areas (Five-Digit) represents five-digit ZIP Code areas used by the U.S. Postal Service to deliver mail more effectively. The first digit of a five-digit ZIP Code divides the United States into 10 large groups of states numbered from 0 in the Northeast to 9 in the far West. Within these areas, each state is divided into an average of 10 smaller geographical areas, identified by the second and third digits. These digits, in conjunction with the first digit, represent a sectional center facility or a mail processing facility area. The fourth and fifth digits identify a post office, station, branch or local delivery area.

Processing: The following steps were performed by Esri: Extracted the data set from StreetMap Premium for ArcGIS North America TomTom 2014 Release 1. Removed any Canada and Mexico records. Attached the fields from Esri Data that includes the 2010 Census fields. Added the SQMI field and calculated its values. Added the field POP2013

and added and calculated the field POP13_SQMI. Removed fields POP2012 and POP12_SQMI. Put "-99" values into all number type fields (except SQMI) for records where there were no data.

Each of the following data sets were extracted at the residential zip code level. No user processing was required.

U.S. Standard Geographies: *Residential zip codes*
Created by the U.S. Postal Service to deliver the mail, ZIP codes do not represent standard census geographic areas for data reporting. Because ZIP Code boundaries are not contiguous with census geographic areas or stable over time, data estimated for ZIP Codes is also subject to change. Residential ZIP Code data is estimated from block group data, using a correspondence file created by assigned Census 2010 block points to Zip Code boundaries.
Source: http://doc.arcgis.com/en/esri-demographics/reference/census-geography.htm

(3) Census, 2010

(4) American Community Survey, 2006-2010

(5) ESRI's Consumer Expenditures Survey, 2012

This data is based on a combination of the latest Consumer Expenditure Surveys (CEX) from the Bureau of Labor Statistics. More information can be found here: https://doc.arcgis.com/en/esri-demographics/data/consumer-spending.htm#ESRI_SECTION1_825737AF565F4E6CAF9793185EA589C8

Appendix B
Definition of Variables

Variable	Description	Source
NormalWeight	1 if respondent is considered to be a normal weight; 0 = otherwise	BRFSS
OverweightObese	1 if respondent is considered to be overweight or obese; 0 = otherwise	BRFSS
Male	1 if respondent is male; 0 = otherwise	BRFSS
Age	Age, years	BRFSS
Married	1 if respondent is married; 0 = otherwise	BRFSS

LessHighSchool	1 if respondent has less than a high school degree; 0 = otherwise	BRFSS
HighSchoolGrad	1 if respondent has a high school diploma, but not a college degree or higher; 0 = otherwise	BRFSS
College Grad	1 if respondent has a college degree or higher; 0 = otherwise	BRFSS
EmployedNOW	1 if respondent is currently employed; 0 = otherwise	BRFSS
Retired	1 if respondent is retired; 0 = otherwise	BRFSS
NOTemployed	1 if respondent is unemployed; 0 = otherwise	BRFSS
White	1 if respondent is non-Hispanic white; 0 = otherwise	BRFSS
Hispanic	1 if respondent is Hispanic; 0 = otherwise	BRFSS
Black	1 if respondent is non-Hispanic black; 0 = otherwise	BRFSS
OtherRace	1 if respondent is another race other than non-Hispanic white or black, or Hispanic; 0 = otherwise	BRFSS
WithChildren	1 if respondent has children in the house under the age of 18; 0 = otherwise	BRFSS
IncomeCon	Annual household income (interval mean of income category)	BRFSS
GoodHealth	1 if respondent self-reported their health as excellent, very good, or good; 0 = otherwise	BRFSS
Moderate	1 if respondent meets recommendation for moderate level of physical	BRFSS

	activity; 0 = otherwise	
Vigorous	1 if respondent meets recommendation for vigorous level of physical activity; 0 = otherwise	BRFSS
Fruit2	1 if respondent meets level of federally recommended fruit (at least 2 cups daily); 0 = otherwise	BRFSS
Veg3	1 if respondent meets level of federally recommended vegetables (at least 3 cups); 0 = otherwise	BRFSS
FormerSmoker	1 if respondent is a former smoker; 0 = otherwise	BRFSS
NotSmoker	1 if respondent has never smoked; 0 = otherwise	BRFSS
Smoker	1 if respondent currently smokes; 0 = otherwise	BRFSS
BingeDrinker	1 if respondent consumes enough alcoholic beverages to be considered a binge drinker (5 or drinks on one occasion for males, 4 or more drinks on one occasion for females); 0 = otherwise	BRFSS
P_LSA	Percentage of land considered to be an LSA	TRF's LSA study
N_LSA_0	Zero LSAs within zip code	TRF's LSA study
N_LSA_1	One disjointed LSA within zip code	TRF's LSA study
N_LSA_2above	Two or more disjointed LSAs within zip code	TRF's LSA study
Region1	1 if respondent resides in the New England census region; 0 = otherwise	BRFSS
Region2	1 if respondent resides in the Middle Atlantic census region;	BRFSS

	0 = otherwise	
Region3	1 if respondent resides in the East North Central census region; 0 = otherwise	BRFSS
Region4	1 if respondent resides in the West North Central census region; 0 = otherwise	BRFSS
Region5	1 if respondent resides in the South Atlantic census region; 0 = otherwise	BRFSS
Region6	1 if respondent resides in the East South Central census region; 0 = otherwise	BRFSS
Region7	1 if respondent resides in the West South Central census region; 0 = otherwise	BRFSS
Region8	1 if respondent resides in the Mountain census region; 0 = otherwise	BRFSS
Region9	1 if respondent resides in the Pacific census region; 0 = otherwise	BRFSS
MSA_1	1 if respondent resides in center city of MSA; 0 = otherwise	BRFSS
MSA_2	1 if respondent resides outside the center city of an MSA but inside the county containing the center city, inside a suburban county of MSA, or in an MSA that has no center city; 0 = otherwise	BRFSS
MSA_3	1 if respondent resides not in an MSA; 0 = otherwise	BRFSS
FoodBudgetShare	Average share of household expenditures spent on food expenditures	Consumer Spending
FFPercentage	Of total food expenditures, percentage spent on fast	Consumer Spending

	food (includes breakfast, lunch, dinner, and snacks at fast food restaurants)	
DiningOutPercentage	Of total food expenditures, percentage spent on food away from the home (includes fast food)	Consumer Spending
Exercisespendingshare	Average share of household income spent on exercise & equipment (sports/rec/exercise equipment + bicycles)	Consumer Spending
Zip_White	Population density of race: White	Census
Zip_Hispanic	Population density of race (ethnicity): Hispanic	Census
Zip_Black	Population density of race: Black	Census
Zip_RaceOther	Population density of race/ethnicity other than white, black, or Hispanic	Census
ACSMEDHINC	Median household income	ACS
ACSAVGHINC	Average household income	ACS
ACSPCI	Per capita income	ACS
Poverty125Percent	1 if respondent is under 125% of the poverty line; 0= otherwise	ACS
POPDENS10	2010 population per square mile	Census

Appendix C.
Estimation for models

Estimation for Overweight/Obesity

	A1	B1	C1	D1
OverweightObese				
Male	0.701***	0.701***	0.706***	0.706***
	0.019	0.019	0.019	0.019
Age (years)	0.017***	0.017***	0.017***	0.017***
	0.001	0.001	0.001	0.001
Married	0.241***	0.242***	0.214***	0.215***
	0.021	0.021	0.022	0.022
Variables for education (base = education less than high school graduate)				
High School Graduate	0.046	0.046	0.058	0.058
	0.042	0.042	0.042	0.042
College Graduate	-0.274***	-0.273***	-0.217***	-0.216***
	0.044	0.044	0.044	0.044
Variables for employment (base = currently unemployed)				
Currently Employed	0.247***	0.247***	0.227**	0.228***
	0.026	0.026	0.026	0.026
Retired	-0.249***	-0.249***	-0.262***	-0.262***
	0.031	0.031	0.031	0.031
Variables for race (base = non-Hispanic white)				
Hispanic	0.292***	0.288***	0.315***	0.315***
	0.040	0.040	0.044	0.044
Black	0.608***	0.599***	0.619***	0.619***
	0.037	0.037	0.041	0.041
OtherRace	-0.379***	-0.379***	-0.306***	-0.305***
	0.043	0.043	0.044	0.044

With Children in House (Yes = 1; No = 0)	0.100***	0.100***	0.103***	0.103***
	0.022	0.022	0.022	0.022
Household Income ($1,000)	-0.002***	-0.002***	-0.001	-0.001
	0	0	0	0
Self-reported Good Health (Good=1; Poor=0)	-0.390***	-0.390***	-0.385***	-0.384***
	0.03	0.03	0.03	0.03
Meets the requirement for moderate physical activity	-0.053***	-0.053***	-0.055***	-0.055***
	0.02	0.02	0.02	0.02
Meets the requirements for vigorous physical activity	-0.122***	-0.122***	-0.114***	-0.114***
	0.03	0.03	0.031	0.031
Meets Fruit/Vegetable requirements	-0.106**	-0.106**	-0.100**	-0.100**
	0.045	0.045	0.046	0.046
Variables for Smokers (base = current smoker)				
Former Smoker	-0.323***	-0.323***	-0.339***	-0.339***
	0.026	0.026	0.026	0.026
Non-Smoker	-0.165***	-0.165***	-0.166***	-0.166***
	0.02	0.02	0.02	0.02
Binge Drinker	0.051**	0.051**	0.060**	0.060**
	0.027	0.027	0.028	0.028
Variables for region (base = Pacific Census region)				
Northeast	0.087***	0.082**	0.071**	0.065**
	0.035	0.035	0.038	0.038
Middle Atlantic	0.147***	0.140***	0.145***	0.138***
	0.041	0.041	0.044	0.044
East North Central	0.192***	0.187***	0.084**	0.080**
	0.037	0.037	0.04	0.04
West North Central	0.222***	0.221***	0.089**	0.086**
	0.038	0.038	0.042	0.042
South Atlantic	0.157***	0.155***	0.048	0.048
	0.036	0.036	0.04	0.04

East South Central	0.361***	0.361***	0.183***	0.184***
	0.042	0.042	0.047	0.047
West South Central	0.243***	0.236***	0.107**	0.103**
	0.039	0.039	0.043	0.043
Mountain	0.043	0.039	-0.033	-0.036
	0.039	0.039	0.041	0.041
Variables for MSA (Metropolitan Statistical Area) (base = living in a non-MSA area)				
Living in a MSA city			-0.75***	-0.080***
			0.026	0.026
Other			0.029	0.026
			0.024	0.024
Variables for number of LSAs (base = two disjointed LSAs)				
Zero LSAs		-0.091***		-0.075**
		0.035		0.037
One LSA		-0.055		-0.051
		0.035		0.036
Percentage of land classified as an LSA	0.001	0	0	-0.001
	0	0	0	0
Fast food percentage			0.093***	0.096***
			0.033	0.033
Dining Out Percentage			-0.083***	-0.086***
			0.020	0.020
Variables for race prevalence within zip code (base = predominantly non-Hispanic white zip code)				
Predominantly Hispanic			0.029	0.025
			0.126	0.126
Predominantly non-Hispanic Black			0.083	0.068
			0.072	0.073
Predominantly Other Race			-0.172	-0.169
			0.120	0.120
Median Household Income			-0.004***	-0.004***

			0.001	0.001
Population Density, 2010			-0.000***	-0.000***
			0	0
Below 125% of poverty line			0.005***	0.006***
			0.002	0.002
Constant	-0.352***	-0.273***	1.771***	1.900***
	0.076	0.082	0.320	0.375
R-squared				
N	310400	310400	310400	310400

* p<0.10, ** p<0.05, *** p<0.01

Estimation for Fruit/Vegetable Intake Requirements

	A1	B1	C1	D1
Male	-0.828***	-0.828***	-0.828***	-0.828***
	0.051	0.051	0.051	0.051
Age (years)	-0.004**	-0.004**	-0.004**	-0.004**
	0.002	0.002	0.002	0.002
Married	-0.055	-0.055	-0.047	-0.047
	0.054	0.054	0.054	0.054
Variables for education (base = education less than high school graduate)				
High School Graduate	-0.044	-0.045	-0.047	-0.047
	0.116	0.116	0.115	0.115
College Graduate	0.366***	0.366***	0.346***	0.346***
	0.118	0.118	0.117	0.117
Variables for employment (base = currently unemployed)				
Currently Employed	-0.043	-0.044	-0.036	-0.036
	0.058	0.058	0.057	0.057
Retired	0.042	0.042	0.049	0.049
	0.071	0.071	0.072	0.072
Variables for race (base = non-Hispanic white)				
Hispanic	-0.209**	-0.207**	-0.197*	-0.197*
	0.101	0.102	0.111	0.111

Black	-0.204***	-0.196**	-0.201**	-0.200**
	0.079	0.08	0.096	0.096
Other Race	-0.083	-0.082	-0.046	-0.046
	0.101	0.101	0.100	0.100
With Children in House (Yes = 1; No = 0)	-0.056	-0.056	-0.053	-0.053
	0.051	0.051	0.051	0.051
Household Income ($1,000)	0.005***	0.005***	0.004***	0.004***
	0.001	0.001	0.001	0.001
Self-reported Good Health (Good=1; Poor=0)	-0.034	-0.035	-0.035	-0.036
	0.074	0.074	0.074	0.074
Meets the requirement for moderate physical activity	0.079*	0.078*	0.079*	0.079*
	0.044	0.044	0.044	0.044
Meets the requirements for vigorous physical activity	0.286***	0.285***	0.284***	0.283***
	0.079	0.079	0.078	0.079
Variables for Smokers (base = current smoker)				
Former Smoker	-0.371***	-0.371***	-0.366***	-0.366***
	0.077	0.077	0.077	0.077
Non-Smoker	0.000	0.000	0.002	0.002
	0.046	0.046	0.047	0.047
BingeDrinker	-0.162**	-0.162**	-0.167**	-0.167**
	0.077	0.077	0.077	0.077
Variables for region (base = Pacific Census region)				
Northeast	-0.279***	-0.279***	-0.360***	-0.360***
	0.079	0.078	0.095	0.094
Middle Atlantic	-0.245***	-0.243***	-0.330***	-0.329***
	0.088	0.088	0.107	0.106
East North Central	-0.392***	-0.392***	-0.466***	-0.465***
	0.082	0.082	0.096	0.095
West North Central	-0.553***	-0.552***	-0.624***	-0.621***
	0.089	0.089	0.102	0.102
South Atlantic	-0.376***	-0.377***	-0.446***	-0.449***
	0.078	0.078	0.093	0.093
East South Central	-0.835***	-0.837***	-0.909***	-0.912***
	0.102	0.102	0.115	0.115
West South Central	-0.349***	-0.349***	-0.412***	-0.414***
	0.086	0.086	0.094	0.093
Mountain	-0.272***	-0.275***	-0.320***	-0.321***
	0.087	0.086	0.095	0.094
Variables for MSA (Metropolitan Statistical Area) (base = living in a non-MSA area)				
Living in a MSA city			-0.002	-0.001
			0.060	0.060

Other			0.002	0.003
			0.058	0.058
Variables for number of LSAs (base = two disjointed LSAs)				
Zero LSAs		0.94		0.92
		0.082		0.087
One LSA		0.142*		0.142*
		0.081		0.084
Percentage of land classified as an LSA	-0.001	-0.001	-0.001	-0.001
	0.001	0.001	0.001	0.001
Fast food percentage			0.081	0.077
			0.077	0.077
Dining out percentage			-0.016	-0.015
			0.052	0.052
Variables for race prevalence within zip code (base = predominantly non-Hispanic white zip code)				
Predominantly Hispanic			-0.649**	-0.644**
			0.294	0.293
Predominantly non-Hispanic Black			-0.081	-0.061
			0.167	0.171
Predominantly Other Race			-0.615**	-0.617**
			0.280	0.281
Median Household Income			0.004*	0.004*
			0.002	0.002
Population Density, 2010			0	0
			0	0
Below 125% of poverty Line			-0.004	-0.005
			0.004	0.004
Constant	-2.533***	-2.633***	-3.269***	-3.385***
	0.184	0.196	0.967	0.978
R-squared				
N	310400	310400	310400	310400

* p<0.10, ** p<0.05, *** p<0.01

An, R. 2015. Health care expenses in relation to obesity and smoking among US adults by gender, race/ethnicity, and age group: 1998–2011. *Public health* 129 (1):29-36.

Andreyeva, Tatiana, Daniel M Blumenthal, Marlene B Schwartz, Michael W Long, and Kelly D Brownell. 2008. Availability and prices of foods across stores and neighborhoods: the case of New Haven, Connecticut. *Health Affairs* 27 (5):1381-1388.

Arif, Ahmed A, and James E Rohrer. 2005. Patterns of alcohol drinking and its association with obesity: data from the third national health and nutrition examination survey, 1988–1994. *BMC Public Health* 5 (1):1-6.

Association, American Heart. *Overweight in Children* 2014. Available from http://www.heart.org/HEARTORG/GettingHealthy/Overweight-in-Children_UCM_304054_Article.jsp.

Aviva Must, PhD, MS Jennifer Spadano, MA Eugenie H. Coakley, MPH, ScD Alison E. Field, MD Graham Colditz, DrPH, and MD William H. Dietz, PhD. 1999. The Diease Burden Associated With Overweight and Obesity. *Journal of American Medical Association* 262 (No. 16):1523-1529.

Baker, Jennifer L, Lina W Olsen, and Thorkild IA Sørensen. 2007. Childhood body-mass index and the risk of coronary heart disease in adulthood. *New England journal of medicine* 357 (23):2329-2337.

Bodor, J Nicholas, Donald Rose, Thomas A Farley, Christopher Swalm, and Susanne K Scott. 2008. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public health nutrition* 11 (04):413-420.

Boone-Heinonen, Janne, Ana V Diez-Roux, David C Goff, Catherine M Loria, Catarina I Kiefe, Barry M Popkin, and Penny Gordon-Larsen. 2013. The neighborhood energy balance equation: does neighborhood food retail environment+ physical activity environment= obesity? The CARDIA Study. *PloS one* 8 (12):e85141.

Brandi Franklin, Ashley Jones, Dejuan Love, Stephane Puckett, Justin Macklin, Shelley White-Means. 2012. Exploring Mediators of Food Insecurity and Obesity: A Review of Recent Literature. *Journal of Community Health* 37:253-264.

Bundred, Peter, Denise Kitchiner, and Iain Buchan. 2001. Prevalence of overweight and obese children between 1989 and 1998: population based series of cross sectional studies. *Bmj* 322 (7282):326.

Caballero, Benjamin. 2007. The global epidemic of obesity: an overview. *Epidemiologic reviews* 29 (1):1-5.

Cannuscio, Carolyn C, Eve E Weiss, and David A Asch. 2010. The contribution of urban foodways to health disparities. *Journal of Urban Health* 87 (3):381-393.

Catherine Califano, Kennen Gross, Lance Loethen, Scott Haag, Ira Goldstein. 2012. Searching for Markets: The Geography of Inequitable Access to Healthy & Affordable Food in the United States. The Reinvestment Fund.

CDC. 2007. Behavioral Risk Factor Surveillance System, 2007 Codebook Report.

———. 2015. *Overview: BRFSS 2007* 20072015]. Available from http://www.cdc.gov/brfss/annual_data/2007/pdf/overview_07.pdf.
———. 2011. Children's Food Environment State Indicator Report, 2011. www.cdc.gov: Centers for Disease Control and Prevention.
———. 2015. *Defining Overweight and Obesity*, April 27, 2012 20122015]. Available from http://www.cdc.gov/obesity/adult/defining.html.
———. 2012. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. edited by N. C. f. H. Statistics. Hyattsville, MD.
———. 2015. *A Growing Problem*, April 17, 2013 20132015]. Available from http://www.cdc.gov/obesity/childhood/problem.html.
———. 2013. One in five adults meet overall physical activity guidelines CDC Newsroom.
———. 2015. *Facts about Physical Activity*, May 23, 2014 20142015]. Available from http://www.cdc.gov/physicalactivity/data/facts.html.
———. 2015. *How much physical activity to adults need?* , March 3, 2014 20142015]. Available from http://www.cdc.gov/physicalactivity/everyone/guidelines/adults.html.
———. 2015. *Survey Data & Documentation*, June 24, 2014 20142015]. Available from http://www.cdc.gov/brfss/data_documentation/index.htm.
———. 2015. *About Adult BMI*, May 15, 2015 20152015]. Available from http://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html.
———. 2015. *Adolescent and School Health: Childhood Obesity Facts*, April 24, 2015 20152015]. Available from http://www.cdc.gov/healthyyouth/obesity/facts.htm.
———. *Childhood Obesity Facts*, June 19, 2015 2015. Available from http://www.cdc.gov/obesity/data/childhood.html.
———. *Healthy Eating for a Healthy Weight*, May 15, 2015 2015. Available from http://www.cdc.gov/healthyweight/healthy_eating/.
———. *Obesity and Overweight*, September 30, 2015 2015. Available from http://www.cdc.gov/nchs/fastats/obesity-overweight.htm.
CDC: Division of Nutrition, Physical Activity, and Obesity, National Center for Chronic Disease Prevention and Health Promotion. 2015. *Overweight and Obesity: Facts*. Centers of Disease Control and Prevention, September 3, 2014 20142015]. Available from http://www.cdc.gov/obesity/data/facts.html.
———. 2015. *Adult Obesity Facts*, June 16, 2015 20152015]. Available from http://www.cdc.gov/obesity/data/adult.html.
Census. *State & County QuickFacts*, September 30, 2015. Available from http://quickfacts.census.gov/qfd/states/00000.html.
Cerin, Ester, Lawrence D Frank, James F Sallis, Brian E Saelens, Terry L Conway, James E Chapman, and Karen Glanz. 2011. From neighborhood design and food options to residents' weight status. *Appetite* 56 (3):693-703.
Chang, Virginia W, and Diane S Lauderdale. 2005. Income disparities in body mass index and obesity in the United States, 1971-2002. *Archives of internal medicine* 165 (18):2122-2128.

Charlotte A. Schoenborn, M.P.H; Patricia F. Adams; Patricia M. Barnes, M.A. 2002. Body Weight Status of Adults: United States, 1997-98 Advance Data No. 330. CDC: National Center for Health Statistics.

Chatterjee, Leonardo Trasande and Samprit. 2009. The Impact of Obesity on Health Service Utilization and Costs in Childhood. *OBESITY* 17 (9).

Cheadle, Allen, Bruce M Psaty, Susan Curry, Edward Wagner, Paula Diehr, Thomas Koepsell, and Alan Kristal. 1991. Community-level comparisons between the grocery store environment and individual dietary practices. *Preventive medicine* 20 (2):250-261.

Cheryl D. Fryar, Margaret D. Carroll, Cynthia L. Ogden. *Prevalence of Overweight and Obesity Among Children and Adolescents: United States, 1963-1965 through 2011-2012* 2014. Available from http://www.cdc.gov/nchs/data/hestat/obesity_child_11_12/obesity_child_11_12.pdf.

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 (2):276-296.

Cynthia L. Ogden, Ph.D., and Margaret D. Carroll, M.S.P.H., Division of Health and Nutrition. 2015. *Prevalence of Overweight, Obesity, and Extreme Obesity Among Adults: United States, Trends 1960-1962 through 2007-2008* 20102015]. Available from http://198.246.102.49/nchs/data/hestat/obesity_adult_07_08/obesity_adult_07_08.pdf.

Dare, Shadrach, Daniel F Mackay, and Jill P Pell. 2015. Relationship between Smoking and Obesity: A Cross-Sectional Study of 499,504 Middle-Aged Adults in the UK General Population.

Diez-Roux, Ana V, F Javier Nieto, Laura Caulfield, Hermann A Tyroler, Robert L Watson, and Moyses Szklo. 1999. Neighbourhood differences in diet: the Atherosclerosis Risk in Communities (ARIC) Study. *Journal of Epidemiology and Community health* 53 (1):55-63.

Dor, Avi, Christine Ferguson, Casey Langwith, and Ellen Tan. 2010. A heavy burden: The individual costs of being overweight and obese in the United States.

Edmonds, Jodi, Tom Baranowski, Janice Baranowski, Karen W Cullen, and Dawnell Myres. 2001. Ecological and socioeconomic correlates of fruit, juice, and vegetable consumption among African-American boys. *Preventive medicine* 32 (6):476-481.

ERS. 2009. Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences. In *Report to Congress*: United States Department of Agriculture, Economic Research Services

Finkelstein, Eric A, Justin G Trogdon, Joel W Cohen, and William Dietz. 2009. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health affairs* 28 (5):w822-w831.

Ford, Paula B, and David A Dzewaltowski. 2008. Disparities in obesity prevalence due to variation in the retail food environment: three testable hypotheses. *Nutrition reviews* 66 (4):216-228.

French, Simone A, Mary Story, and Robert W Jeffery. 2001. Environmental influences on eating and physical activity. *Annual review of public health* 22 (1):309-335.

Giles-Corti, Billie, Sally Macintyre, Johanna P Clarkson, Terro Pikora, and Robert J Donovan. 2003. Environmental and lifestyle factors associated with overweight and obesity in Perth, Australia. *American Journal of Health Promotion* 18 (1):93-102.

Glanz, Karen, James F Sallis, Brian E Saelens, and Lawrence D Frank. 2005. Healthy nutrition environments: concepts and measures. *American Journal of Health Promotion* 19 (5):330-333.

Green, William H. 2011. *Econometric Analysis* 7th ed. New York University.

Guo, Shumei S, and William Cameron Chumlea. 1999. Tracking of body mass index in children in relation to overweight in adulthood. *The American journal of clinical nutrition* 70 (1):145s-148s.

Hammond, Ross A, and Ruth Levine. 2010. The economic impact of obesity in the United States. *Diabetes, metabolic syndrome and obesity: targets and therapy* 3:285.

Horowitz, Carol R, Kathryn A Colson, Paul L Hebert, and Kristie Lancaster. 2004. Barriers to buying healthy foods for people with diabetes: evidence of environmental disparities. *American Journal of Public Health* 94 (9):1549-1554.

Janne Boone-Heinonen, PhD, PhD Penny Gordon-Larsen, MD Catarina I. Kiefe, PhD, DrPH James M. Shikany, MD Cora E. Lewis, and PhD Barry M. Popkin. 2011. Fast Food Restaurant and Food Store--Longitudinal Associations With Diet in Young to Middle-aged Adults: The CARDIA Study. *Archives of Internal Medicine* 171 (13):1162-1170.

Jeffrey Levi, Ph.D., MA Laura M. Segal, Rebecca St. Laurent, and MPH Jack Rayburn. 2015. *State of Obesity*. Robert Wood Johnson Foundation 20152015]. Available from http://stateofobesity.org/facts-economic-costs-of-obesity/.

Jin, Maoyong Fan and Yanhong. 2014. Obesity and Self-Control: Food Consumption, Physical Activity and Weight-Loss Intention

John Cawley, Chad Meyerhoefer. 2012. The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics*
219-230.

Koplan, Jeffrey P, Catharyn T Liverman, and Vivica I Kraak. 2005. Preventing childhood obesity: health in the balance: executive summary. *Journal of the American Dietetic Association* 105 (1):131-138.

Laraia, Barbara A, Anna Maria Siega-Riz, Jay S Kaufman, and Sonya J Jones. 2004. Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventive medicine* 39 (5):869-875.

Lauren M. Dinour, MPH, RD, Dara Bergen, MPH, RD, Ming-Chin Yeh, PhD, MEd, MS. 2007. The Food Insecurity-Obeisty Paradox: A Review of the Literature and the Role Food Stamps May Play. *Journal of the American Dietetic Association* 107 (11):1952-1961.

Lee, Helen. 2012. The role of local food availability in explaining obesity risk among young school-aged children. *Social Science & Medicine* 74 (8):1193-1203.

Liu, Gilbert C, Jeffrey S Wilson, Rong Qi, and Jun Ying. 2007. Green neighborhoods, food retail and childhood overweight: differences by population density. *American Journal of Health Promotion* 21 (4s):317-325.

Lopez, Russ P. 2007. Neighborhood risk factors for obesity. *Obesity* 15 (8):2111-2119.

Macintyre, Steven Cummins and Sally. 2006. Food environments and obesity--neighbourhood or nation? *International Journal of Epidemiology* (35):100-104.

Marder, William, and Stella Chang. 2006. Childhood obesity: costs, treatment patterns, disparities in care, and prevalent medical conditions. Thomson Medstat Research Brief New York.

Mary Gatineau, Shireen Mathrani. 2012. Obesity and alcohol: an overview. edited by D. J. F. Professor Betsy Thom: National Obesity Observatory.

Mokdad, Ali H, Earl S Ford, Barbara A Bowman, William H Dietz, Frank Vinicor, Virginia S Bales, and James S Marks. 2003. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Jama* 289 (1):76-79.

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 (2):325-331.

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 (2):206-216.

Morland, Kimberly B, and Kelly R Evenson. 2009. Obesity prevalence and the local food environment. *Health & place* 15 (2):491-495.

Morland, Kimberly, and Susan Filomena. 2007. Disparities in the availability of fruits and vegetables between racially segregated urban neighbourhoods. *Public health nutrition* 10 (12):1481-1489.

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 (4):333-339.

Morland, Kimberly, Steve Wing, and Ana Diez Roux. 2002. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *American journal of public health* 92 (11):1761-1768.

Morland, Kimberly, Steve Wing, Ana Diez Roux, and Charles Poole. 2002. Neighborhood characteristics associated with the location of food stores and food service places. *American journal of preventive medicine* 22 (1):23-29.

MS Townsend, J Peerson, B Love, C Achterberg, SP Murphy. 2001. Food insecurity is positively related to overweight in women. *Journal of Nutrition* (131):1738-1745.

Nicole I. Larson, PhD, PHS, RD, Mary T. Story, PhD, RD, Melissa C. Nelson, PhD, RD. 2009. Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S. *American Journal of Preventive Medicine* 36 (1):74-81.

Ogden CL, Lamb MM, Carroll MD, Flegal KM. 2010. Obesity and socio economic status in adults: United States 1988-1994 and 2005-2008. In *NCHS data brief no. 50*.

Ogden, Cynthia L, Margaret D Carroll, Brian K Kit, and Katherine M Flegal. 2014. Prevalence of childhood and adult obesity in the United States, 2011-2012. *Jama* 311 (8):806-814.

Ogden, Cynthia L, Katherine M Flegal, Margaret D Carroll, and Clifford L Johnson. 2002. Prevalence and trends in overweight among US children and adolescents, 1999-2000. *Jama* 288 (14):1728-1732.

Ogden CL, Lamb MM, Carroll MD, Flegal KM. Obesity and socio economic status in adults: United States 1988-1994 and 2005-2008. NCHS data brief no 50. Hyattsville, MD: National Center for Health Statistics. 2010.

Olshansky, S Jay, Douglas J Passaro, Ronald C Hershow, Jennifer Layden, Bruce A Carnes, Jacob Brody, Leonard Hayflick, Robert N Butler, David B Allison, and David S Ludwig. 2005. A potential decline in life expectancy in the United States in the 21st century. *New England Journal of Medicine* 352 (11):1138-1145.

Paeratakul, S, JC Lovejoy, DH Ryan, and GA Bray. 2002. The relation of gender, race and socioeconomic status to obesity and obesity comorbidities in a sample of US adults. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity* 26 (9):1205-1210.

Papas, Mia A, Anthony J Alberg, Reid Ewing, Kathy J Helzlsouer, Tiffany L Gary, and Ann C Klassen. 2007. The built environment and obesity. *Epidemiologic reviews* 29 (1):129-143.

Patel, Kushal, Margaret K Hargreaves, Jianguo Liu, David Schlundt, Maureen Sanderson, Charles E Matthews, Charlene M Dewey, Donna Kenerson, Maciej S Buchowski, and William J Blot. 2011. Relationship between smoking and obesity among women. *American journal of health behavior* 35 (5):627.

PE Wilde, CK Ranney. 2000. The monthly food stamp cycle: shopping frequency and food intake decisions in an endogenous switching regression framework. *American Journal of Agricultural Economics* (82):200-213.

Pearson, Tim, Jean Russell, Michael J Campbell, and Margo E Barker. 2005. Do 'food deserts' influence fruit and vegetable consumption?—A cross-sectional study. *Appetite* 45 (2):195-197.

Pi-Sunyer, F Xavier. 1992. Health Implications of Obesity. *Am J Clin Nutr*:1595S-1603S.

Policy Solutions, The Reinvestment Fund. 2015. *2011 Limited Supermarket Access Analysis:*

Summary of TRF's Methodology 20112015]. Available from http://www.trfund.com/wp-content/uploads/2013/07/LSAMethodology2011.pdf.

Poston, Walker S Carlos, and John P Foreyt. 1999. Obesity is an environmental issue. *Atherosclerosis* 146 (2):201-209.

Powell, Lisa M, Sandy Slater, Donka Mirtcheva, Yanjun Bao, and Frank J Chaloupka. 2007. Food store availability and neighborhood characteristics in the United States. *Preventive medicine* 44 (3):189-195.

Prevention, Centers for Disease Control and. 2015. *About Healthy Places*, January 23, 2014 20092015]. Available from http://www.cdc.gov/healthyplaces/about.htm.

Rahman, Tamanna, Rachel A Cushing, and Richard J Jackson. 2011. Contributions of built environment to childhood obesity. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine* 78 (1):49-57.

Richard V. Burkhauser, John Cawley. 2008. Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics* 519-529.

Rundle, Andrew, Kathryn M Neckerman, Lance Freeman, Gina S Lovasi, Marnie Purciel, James Quinn, Catherine Richards, Neelanjan Sircar, and Christopher Weiss. 2009. Neighborhood food environment and walkability predict obesity in New York City. *Environ Health Perspect* 117 (3):442-447.

S Kumanyika, RW Jeffery, A Morabia, C Ritenbaugh, VJ Antipatis. 2002. Obesity Prevention: the case for action. *International Journal of Obesity* 26:425-436.

Satcher, D. 2010. Surgeon General's call to action to prevent and decrease overweight and obesity. US Department of Health and Human Services; 2001.

Serdula, Mary K, Donna Ivery, Ralph J Coates, David S Freedman, David F Williamson, and Tim Byers. 1993. Do obese children become obese adults? A review of the literature. *Preventive medicine* 22 (2):167-177.

Services, U.S. Department of Health and Human. 2010. Executive Summary, Dietary Guidelines for Americans, 2010.

———. 2015. *Nutrition and Weight Status, An Overview*, 4/25/2014 20142015]. Available from http://www.healthypeople.gov/2020/topics-objectives/topic/nutrition-and-weight-status.

Sobal, Jeffery, and Barbara S Rauschenbach. 2003. Gender, marital status, and body weight in older US adults. *Gender Issues* 21 (3):75-94.

Sturm, Roland, and Ashlesha Datar. 2005. Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public health* 119 (12):1059-1068.

Suter, Paolo M, and Angelo Tremblay. 2005. Is alcohol consumption a risk factor for weight gain and obesity? *Critical reviews in clinical laboratory sciences* 42 (3):197-227.

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 (6):563-570.

Swinburn, Garry Egger and Boyd. 1997. An "Ecological" Approach to the Obesity Pandemic. *British Medical Journal* 312 (7106):477-480.

Tate, Karl. 2015. Americans Aren't Eating Their Fruits and Vegetables (Infographic).

Theodore B. Van Itallie, M.D. 1985. Health Implications of Overweight and Obesity in the United States. *Annals of Internal Medicine* 103 (Number 6 (Part 2)):983-988.

Thompson, Latetia V. Moore; Frances E. 2015. Adults Meeting Fuit and Vegetable Intake Recommendations--United States, 2013. *Morbidity and Mortality Weekly Report (MMWR)* 64 (26):709-713.

USDA. 2015. *How Many Vegetables Are Needed Daily or Weekly?* 2015]. Available from http://choosemyplate.gov/printpages/MyPlateFoodGroups/Vegetables/food-groups.vegetables-amount.pdf.

U.S. Department of Health and Human Services. The Surgeon General's call to action to prevent and decrease overweight and obesity. [Rockville, MD]: U.S. Department of Health and Human Services, Public Health Service, Office of the Surgeon General; [2001]. Available from: U.S. GPO, Washington.

———. 2015. *How Much Fruit is Needed Daily?* 2015]. Available from http://choosemyplate.gov/printpages/MyPlateFoodGroups/Fruits/food-groups.fruits-amount.pdf.

WHO. 2013. World Health Statistics, 2013.

———. 2015. *Obesity and overweight*, January 2015 20152015]. Available from http://www.who.int/mediacentre/factsheets/fs311/en/.

———. *Obesity and overweight fact sheet N# 311. Updated 2015* 2015. Available from http://www.who.int/mediacentre/factsheets/fs311/en/.

Zenk, Shannon N, Amy J Schulz, Teretha Hollis-Neely, Richard T Campbell, Nellie Holmes, Gloria Watkins, Robin Nwankwo, and Angela Odoms-Young. 2005. Fruit and vegetable intake in African Americans: income and store characteristics. *American journal of preventive medicine* 29 (1):1-9.

Zip Code FAQs. 2014. Available from http://www.zipboundary.com/zipcode_faqs.html.