CONTEXTUAL INFLUENCES ON OBESITY PREVALENCE:
A SPATIALLY EXPLICIT ANALYSIS

A Thesis in
Sociology and Demography
by
Nyesha Cheyenne Black

© 2011 Nyesha Cheyenne Black

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Arts

December 2011
The thesis of Nyesha Cheyenne Black was reviewed and approved* by the following:

Stephen Matthews
Associate Professor of Sociology, Anthropology, and Demography
Thesis Advisor

John Iceland
Professor of Sociology and Demography
Head of Department of Sociology & Crime, Law and Justice

Leif Jensen
Professor of Rural Sociology and Demography

*Signatures are on file in the Graduate School
ABSTRACT

This study examines contextual influences on obesity prevalence in coterminous counties in the United States. Using a variety of secondary data sources, I constructed a dataset with a rich set of county-level contextual variables. This study employs an ecological, spatially explicit perspective by exploring the influence of macro-level processes and the environment on health disparities. Traditional regression methods (OLS), along with exploratory spatial data analysis (Moran’s I) and geographic weighted regression (GWR), were utilized to thoroughly examine the relationship between obesity and a rich set of predictors. Results from traditional regression methods show that rural vs. urban residence does not significantly contribute to differences in obesity prevalence by county; whereas minority composition, features of the built and natural environment, and physical inactivity among adult residents are all significantly associated with county-level obesity prevalence. Also, county-level income inequality was found to provide a protective barrier against higher obesity prevalence. This finding suggests that the relationship between relative deprivation and health should be further explored in the health disparities literature. Furthermore, GWR confirms that place matters and the relationship between contextual influences and obesity prevalence varies substantially across place. GWR also provides an empirical basis for the public health community to design interventions that effectively target predictors of obesity at the local level.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ v

LIST OF TABLES ........................................................................................................ vi

ACKNOWLEDGEMENTS ........................................................................................... vii

Chapter I. Introduction ............................................................................................. 1
  Thesis Overview .................................................................................................... 3

Chapter II. Background .......................................................................................... 4
  Trends, Causes, and Consequences of Obesity ..................................................... 4
  Macro-level and Social Processes ......................................................................... 6
  The Social Determinants of Disease ..................................................................... 8
  Place Matters ........................................................................................................ 9
  Environmental Influences ..................................................................................... 10
    Built Environment ............................................................................................ 11
    Natural Environment ......................................................................................... 14
    Summary ............................................................................................................ 15

Chapter III. Hypotheses ......................................................................................... 16

Chapter IV. Data and Measures ............................................................................. 18
  Data ..................................................................................................................... 18
  Measures ............................................................................................................ 20
    Dependent Variable ........................................................................................ 20
    Independent Variables ..................................................................................... 20

Chapter V. Methods ............................................................................................... 26

Chapter VI. Results ............................................................................................... 32
  Descriptive Statistics ........................................................................................ 32
  Global (OLS) Model Results ............................................................................. 41
  Local (GWR) Model Results ............................................................................. 45

Chapter VII. Discussion ......................................................................................... 56

References .............................................................................................................. 61

Appendix .................................................................................................................. 72
LIST OF FIGURES

Figure 1: Adult obesity prevalence in the United States (2008) ...................... 34
Figure 2: Distribution of minority populations in the United States (2000) ............ 35
Figure 3: Distribution of income inequality (Gini) in the United States (2000) ....... 37
Figure 4: Access and proximity to grocery stores in the United States (2006) ......... 38
Figure 5: Distribution of amenity scale in the United States (2000) .................... 39
Figure 6: Estimates of physically inactive adults in the United States (2008) ......... 40
Figure 7: Local Moran’s I for standardized OLS residuals ............................ 44
Figure 8: Local r-square of GWR model .................................................. 47
Figure 9: GWR estimates of minority populations ........................................ 49
Figure 10: GWR estimates of income inequality ............................................ 52
Figure 11: GWR estimates of access and proximity to grocery stores ................. 53
Figure 12: GWR estimates of amenity scale ................................................. 54
Figure 13: GWR estimates of physical inactivity ............................................ 55
LIST OF TABLES

Table 1: Recodes of gradients of physically inactive adults in the United States........ 25

Table 2: Descriptive statistics for percentage of obese population and contextual
variables. ...................................................................................................................... 33

Table 3: Global (OLS) regression model predicting obesity ........................................ 41

Table 4: Geographically weighted regression 5-number parameter summary results
and Monte Carlo significance test for spatial variability
of parameters (N=3,108)............................................................................................ 46
ACKNOWLEDGEMENTS

First, I wish to thank God. I have learned so much about life and myself in the last two years. The most important lesson I have learned is that God has a plan for my life and if I seek his guidance and submit to his plan for me, I have no reason to stress or worry about any obstacle I may face. He has also blessed me with amazing parents who instilled the value of education in me, and have always supported me and my endeavors. My boyfriend, Kendrick, also deserves special thanks, as he has stood by me and my decision to pursue my graduate education at Penn State. He makes sure to make me laugh daily (which definitely helps me get through the long winters in Central Pennsylvania), and he gives me something to look forward to in life.

I also must thank my committee. I admire the scholarship of Leif Jensen and John Iceland, and I am honored that they agreed to serve on my committee. Stephen Matthews has been a phenomenal mentor, and I am forever grateful for his guidance and patience as I have struggled through the last two years. I also appreciate Stephen for giving me the opportunity to learn new skills, methods, and to collect data for him in Philadelphia – the best experience of my graduate career thus far.

I also must thank members of the GIA core – Carla Shoff and Yosef Bodovski—for helping me solve data and software issues.
Chapter I

Introduction

Understanding how the social conditions of individuals are connected to broader social structures is a salient tenet of sociological inquiry. Individuals are stratified into hierarchical groups which create competition for scarce resources that usually results in power struggles and social inequalities. The gross marginalization of groups in the United States hampers the ability of disadvantaged individuals to access resources that are essential to maintaining health and well-being. Hence, not having access to health-enhancing goods and services creates and perpetuates health disparities. Health disparities remain unyieldingly prominent across race, socioeconomic status, education attainment, and gender (Williams and Collins 1995; Montez et al. 2009; Elo and Preston 1996). The stratification of health in the United States places individuals at risk of exposure to detrimental health-causing agents. It is well-established that health disparities not only persist across individuals, but also across place (Geronimous, Bound, and Waidmann 1999; Murray et al. 2006; Cossman et al. 2007). Place serves as a conduit for opportunity structures that is socially constructed and shapes health disparities. “Spatially patterned health inequalities are rooted in the unequal distribution of resources” (Bernard et al. 2007:1841). Since resources are not allocated evenly across place, social scientists have an interest in how contextual-level influences affect an individual’s health status separately from compositional features of the environment (Robert 1999).

The objective of this thesis is to understand how area based differentials in health enhancing resources at the county level undergirds the spatial patterning of health across the continental United States. More specifically, I will examine contextual-level influences on
obesity by using a uniquely constructed dataset. My data set includes variables from multiple sources, many of which have yet to be examined, to my knowledge, in the health disparities literature. Furthermore, I will employ a regression-based method—geographically weighted regression (GWR) — that accounts for the unique attributes of place. This spatially explicit perspective is innovative and contributes to the health disparities literature on multiple fronts. First, whereas most studies to date have examined environmental influences on obesity at the local-level within tightly bounded study areas (Morland, Diez-Roux, and Wing 2006; Morland and Filomena 2007; Zenk et al. 2009), I examine environmental influences on obesity outcomes across all coterminous counties. The United Department of Agriculture (USDA) Food Environment Atlas offers relatively recent data on county level attributes that are related to obesity and other diet-related diseases. Few published studies have analyzed the USDA Food Environment Atlas dataset (Ahern, Brown, and Dukas 2011). Second, the dataset I have compiled offers a more comprehensive examination of the relationship between the contextual environment and obesity outcomes. In addition to including measures of the built environment, I look at temporal changes in the food retail landscape, specifically supermarkets and convenience stores. The built environment literature offers justification for such an examination, but none have been conducted to date. Also, I examine the relationship between the natural or physical environment and obesity outcomes, a relationship largely neglected in the health disparities research. Third, I extend upon ordinary least squares (OLS) regression methods by using geographically weighted regression to examine county-level obesity outcomes. I am not aware of a study that has employed GWR to predict obesity prevalence at any aggregate level, particularly the county level. Finally, I include a measure of relative deprivation to add to the debate in the health literature about the relationship between income inequality and obesity. My analyses
investigate the complicated interrelationships among residence status, socioeconomic status, and
health. This contextually-based approach can potentially influence public health interventions
that effectively target the predictors of obesity at the local level.

Thesis Overview

The thesis consists of seven chapters including several maps and tables, and followed by
references and an appendix. In chapter I, I introduce the importance of examining the
stratification of health through a contextual lens. In chapter II, I continue this synthesis, by
discussing the consequences and causes of obesity. I then connect those causes to broader
macro-social processes that are also rooted in demographic processes. I also discuss the import
of place and how various features of the environment, in which local communities are embedded,
determines access to health promoting resources thereby contributing to health disparities.
Chapter III details the hypotheses I test with empirical analyses, while chapter IV describes the
data sources used to construct my dataset, as well as the measures used in the final analyses.
Chapter V explains the methodological approaches used in this study, with a thorough discussion
of GWR. Chapter VI presents the results of the empirical analyses as well as the maps I created.
The last chapter (VII) discusses the inferences that can be gleaned from this empirical research,
the limitations of the study, and the implications for policy and future research.
Chapter II
Background

Trends, Causes, and Consequences of Obesity

Substantial public discourse highlights the increased incidence and prevalence of obesity and other diet-related diseases among Americans. Diet-related diseases, particularly obesity, are considered one of the most pressing public health concerns in the United States (Department of Health and Human Services 2001). Obesity rates have increased rapidly in the last three decades, among all ethnic groups, ages, regions, gender, income levels, and education levels. Adult obesity has increased from 13% to about 30% between the 1960s and 2009 (Wang and Beydoun 2007). Childhood obesity prevalence has also increased, tripling since 1980, to a level of approximately 17% (Ogden et al. 2010). The South and rural communities have higher rates of obesity (Jackson et al. 2001). African-Americans, especially African-American women, and Hispanics have higher rates of obesity compared to their white counterparts (Flegel et al. 2002, Hedley et al. 2004). Income and education are inversely related to obesity, and poverty is positively associated with obesity (Drewnowski 2004). These patterns continue to exist today.

Obesity has substantial economic consequences. Finkelstein and colleagues (2009) analysis of the monetary cost of obesity reveals that the medical expenditures associated with obesity increased from $78.5 billion in 1998 to an estimated $147 billion in 2008. A significant portion of the cost was absorbed by government programs such as Medicare and Medicaid. Obesity is also associated with lower productivity. It is estimated that obese workers cost private employers approximately $45 billion a year (Rosen and Barrington 2008).

The implications of obesity also have serious consequences for well-being and life expectancy. Obesity is positively associated with other chronic illnesses, thus placing individuals
with obesity at higher risk of a variety of degenerative diseases such as cardiovascular disease, strokes, diabetes, hypertension, and certain cancers (Malnick and Knobler 2006). The number of deaths attributed to obesity is difficult to estimate due to the classification of deaths under other illnesses that are contributed to obesity. However, the annual number of deaths attributed to obesity has been estimated between 100,000 to 400,000 (Allison et al. 1999; Mokdad et al. 2004; Flegal et al. 2005). The excess deaths associated with obesity may lead to an actual decline in life expectancy and longer years of life with morbidity for the current generation of adolescents (Olshansky et al. 2005; Daniels 2006).

Several personal attributes and behaviors are associated with obesity. Genetics has been proven to be associated to the risk of obesity, though the biological capability of gene transformation does not explain the rapid growth rates of obesity witnessed in the past four decades (Hill and Townbridge 1998). Sedentary lifestyles such as television viewing, computer use, and video game playing are associated with obesity among adolescents (Utter et. al 2003). Physical inactivity is also positively associated with obesity risk during the life course (Hu et. al 2003). Americans are also consuming more calorie-dense foods (Ledikwe et al. 2005), such as fast foods (Bowman et al. 2004), and sugared-sweetened beverages (Ludwig et al. 2001; Schulze et. al 2004; Popkin 2008).

While personal dietary behaviors are linked to increased trends in obesity, emerging evidence implicates contextual environments as influences on health behaviors and outcomes (Lovasi et al. 2009). Food access is an important factor in maintaining a healthy diet; one that includes consuming fruits, vegetables, and whole grains, while limiting consumption of high fat, high sodium, and processed foods. Neighborhoods that do not have adequate access to quality, healthy food options are classified as food deserts (Cummins 2002). Moreover, neighborhoods
with low food accessibility, in addition to environments that are not conducive to physical activity, are considered obesogenic environments (Swinburn et al. 1999).

*Macro-level Social Processes*

It is necessary to elaborate on the causes of obesity presented above. Directing attention to demographic, epidemiological and, nutritional processes is necessary to understand the obesity epidemic. The concepts in demography studies and social processes help elucidate the causes of obesity in the United States. Historical transitions that are seemingly inherent to all societies affect individual lifestyles and health outcomes. The demographic transition, which is linked to urbanization and industrialization, is thought to have three stages: 1) low population growth due to extremely high birth rates and death rates, 2) a period of rapid population growth with falling death rates, but still high birth rates, and 3) a period of little to no population growth due to below-replacement fertility (Kirk 1996). Consistent with the patterns of the demographic transition is the epidemiological transition. The rapid decline in deaths during the second phase of the demographic transition is a consequence of shifting patterns from deaths attributed to infectious diseases to excess mortality caused by degenerative diseases (Orman 1971). When infectious diseases were the main cause of death, contextually based interventions were the dominant approach to eliminate diseases among public health interventionists. As excess deaths became more attributed to chronic illnesses, the public health community placed more emphasis on individual lifestyles as the primary risk factor for disease onset (Macintyre and Ellaway 2000). The focus on personal behaviors has often proven to produce ineffective treatments on health outcomes. Modern public health pedagogy has shifted its emphasis to the importance of social processes on health outcomes (Osypuk and Galea 2007).
The demographic and the epidemiological transition are closely associated with the nutritional transition (Popkin and Gordon-Larsen 2004). Societies undergoing the nutritional transition eventually adopt a diet that is high in carbohydrates, processed foods, and high sugar and fat content. These dietary behaviors coupled with less leisure physical activity under the nutritional transition, greatly contribute to increased trends of nutrition-related disabilities and related deaths (Popkin and Gordon-Larsen 2004).

In summary, modernization shifts individuals from provincial lifestyles to the urban sphere which promotes less manual-intensive labor under the demographic transition. Moreover, the demographic transition aligns with the epidemiological transition while simultaneous advances in public health results in the sharp decline of contagious diseases. These advances in public health are later eclipsed by the construction of diet-related diseases that are prominent during the nutritional transition.

Popkin and Gordon-Larsen (2004) describe the final stage of the nutritional transition as a period in which personal behavioral changes results in the reduction of obesity outcomes due to a decrease in the consumption of calorie-dense foods, an increase in the consumption of fiber-rich foods, fruits and vegetables, and an increase in leisure-time activity. However, the last stage of the proposed nutritional transition is based on the questionable underlying assumption that changes in personal behaviors will directly lead to a decline in obesity and poor dietary behaviors. The changes needed to meet the final stage of the nutritional transition calls for a shift in attention to understanding how sociological processes, both at the individual and contextual levels, contributes to the social and spatial patterning of health disparities.
The Social Determinants of Disease

Social and economic inequalities are salient determinants of population health (Lynch 1996; Lynch and Kaplan 1997). The social patterning of health inequalities varies across health outcomes such as disability, nonfatal impairments, and chronic illness (Hayward et al. 2000). The direct and indirect cost of health inequalities in the United States between 2003 and 2006 was a staggering $1.24 trillion; while the annual lost of economic productivity due to racial health disparities is estimated at $309 billion (LaVeist et al. 2009).

Income inequality is associated with health and mortality (Kawachi et al. 1997; Wilkinson 1997). Hayward and colleagues (2000) also find that improving socioeconomic resources among blacks are more effective in reducing racial inequalities in health than altering health behaviors. However, other studies have shown an inconsistent relationship between relative deprivation and health outcomes (Lorant et al. 2001; Shi et al. 1999). There have also been inconsistent empirical findings to corroborate a positive or negative relationship between obesity and relative deprivation. A state-level analysis of abdominal weight gain shows that men and women are more likely to gain weight if they are residents of states with high household income inequality (Kahn et al. 1998). One study has shown that income inequality at the state-level is also associated with a higher body mass index among women (Diez-Roux et al. 2000). In contrast, Chang et al. (2005) show that among race-sex groups, there is a significant inverse relationship between income inequality and BMI for white women living in metropolitan areas. The authors conclude that relative deprivation may serve as a protective factor against obesity and overweight status among white women. These conflicting studies suggest that the relationship between relative deprivation and obesity needs to be further elucidated.
Place Matters

A better understanding of how features of the contextual environment manifest internally and affects health status is important to unravel the conundrum of eliminating health disparities (Glass and McAtee 2006). The importance of place in shaping social processes such as inequalities, poverty, and employment opportunities has garnered attention among sociologists (Gieryn, 2000; Tickameyer, 2000; Lobao, Hooks, and Tickameyer 2007). Gieryn (2000) and Tickameyer (2000) emphasize the importance of “place-sensitive sociology” that recognizes that place has downstream effects on individual behaviors, relationships, and opportunities; and perpetuates power struggles via the social construction of the environment. Simply including dummy variables that distinguish regional and metropolitan residential status may be insufficient (Tickameyer, 2000). Models must go beyond treating the effects of place as unexplained variance in the error term, and attempt to build conceptual and empirical models that fully specify place effects on health outcomes (Diez-Roux 2002).

Contextual characteristics of poor places are not conducive to health and healthy lifestyles. Place may exacerbate the effects of poor personal behaviors on health outcomes, or place may enhance the benefits of good health behaviors. Thus the mediating role that place may have on personal behaviors has been described as deprivation amplification (Macintyre et al. 1993; Macintyre 2007). Characteristics of the environment that may contribute to deprivation amplification are residential segregation, a lack of infrastructure, and income inequality. In the same vein, area and individual level attributes that may enhance deprivation amplification are also considered the fundamental causes of disease (Link and Phelan 1995). Access to resources is important to eschew deprivation amplification. Resources such as knowledge, money, power,
prestige, and social networks are safety nets against adverse health conditions (Phelan, Link, and Tehranifar 2010). It is well documented that these protective factors and disease prevalence are not evenly distributed across people or places (Geronimus et al. 1996; Murray et al. 2006; Cossman et al. 2007; Ezzati et al. 2008). The fundamental cause theory predicts that health disparities are intractable as long as resources are unevenly allocated across place (Phelan, Link, Tehranifar 2010).

Kemp’s (2010) discussion of place histories reiterates how place matters in relation to reproducing inequalities in society. Inequalities in health are pervasive among racial and ethnic groups, socioeconomic status, gender status, and metropolitan status. The mediating effects among these variables on health outcomes (LaVeist 2005; Read and Gorman 2006; Hartely 2004), vary across counties in the United States (Murray et al. 2006).

The health outcome that is the focus of this study, obesity, is not evenly distributed across people or places. The literature supports the notion that the non-random distribution of obesity is attributed to the unequal distribution of resources that are imperative to eliminating disparities in obesity rates across place.

Environmental Influences

Quality of life and health are associated with not only the social environment, but also the physical environment (Pacione 1984), including the natural and built environment (Kearns and Gesler 1998; Macintyre, Elleway, and Cummins 2002). The prevalence and incidence of disease represents a confluence of personal behaviors and contextual factors (Sooman and MacIntyre 1995). It is imperative to understand both the independent and mediating role of the physical environment on health outcomes.
Built Environment

The built environment has been recognized as a crucial marker for access to resources, such as food stores and green spaces, which may mitigate the risk of poor health outcomes such as obesity. The spatial distribution of activities dictates land utilization. The built environment circumscribes patterns of human activity within the physical environment (Handy et al. 2002). The built environment is the result of local zoning ordinances, transportation systems, segregation patterns, and investments or disinvestments in the community. It has been well documented that food access is positively associated with a more nutritious diet and lower rates of obesity and other diet-related diseases (Larson et al. 2009); and differences in food retail access may contribute to and perpetuate health disparities (Diez-Roux and Mair 2010). Empirical evidence shows a negative relationship between the consumption of nutritious foods and limited access to supermarkets (Morland et al. 2002). Convenience stores are positively associated with obesity and other comorbidities such as diabetes, hypertension, and cardiovascular disease (Morland et al. 2006; Bodor et al. 2010). Communities that are the most vulnerable to adverse health risks are typically inundated with convenience stores which offer a limited range of food options, such as fresh produce, meat, and low-fat milk (Ver Ploeg et al. 2009; Leise et al. 2007). In all, poor access to nutritious food options can be a barrier for changing poor personal dietary behaviors which increases the risk of obesity (Baker et al. 2006). Access is an important element of quality of life and social equality (Stein 2002). Accessibility is measured by the ease to which people can acquire essential goods and services (Pacione 1984). The concept of accessibility also involves dimensions of satisfaction, availability, affordability, accommodation, and acceptability of goods and services (Penchansky
and Thomas 1981). Mobility via vehicle ownership also affects accessibility to goods and services. Social barriers, both ascribed and achieved, affect vehicle ownership and access to food stores and activity spaces. Research shows that private transportation is more imperative to food access than proximity to food stores, especially when public transportation is not available (Altchueler, Somkin, and Adler 2004; Coveney and O’Dwyer 2009). A higher percentage of low-income households do not own a vehicle, and 17% of the total U.S. population has low or medium access to supermarkets (Ver Ploeg et al. 2009). In the urban sphere, the exodus of large grocery stores and supermarkets to the suburbs coupled with low vehicle ownership among low-income neighborhoods contributes to poor dietary decisions among the poor and minorities (Morland et al. 2002).

Rural residents are often geographically constrained (Ver Ploeg et al. 2009) and more heavily burdened by the additional travel costs and time allocation required to seek opportunities and services (Pacione 1984). Rural residents often must commute to metropolitan areas for many services, including accessing services at large food retailers (Ver Ploeg et al. 2009; Michimi and Wimberley 2010). On average rural residents make fewer trips to the grocer, and spend almost 16.9 minutes commuting to grocery stores compared to only 14.2 minutes traveled by their urban counterparts (Ver Ploeg et al. 2009). Accessibility to supermarkets and large grocers allow residents to take advantage of lower prices and more varied food options regardless of neighborhood poverty levels (Ver Ploeg et al. 2009; Leise et al. 2007). Supermarkets prices are approximately 4-10% cheaper than convenience stores for the same basket of goods, and rural supermarket prices are 4% higher than supermarkets located in urban areas (Kaufman 1999; Andreyeva et al. 2008).
Some researchers conclude that the poor pay more for food (Chung and Myers 1999). In contrast, other studies have shown that higher-income households pay more for groceries than low and medium income households (Broda et al. 2009; Kaufman et al. 1997). However, these studies fail to control for the shopping preferences of higher income groups for “designer supermarkets”, such as Whole Foods, which may inflate the cost of foods purchased by high income households. Also, lower food prices are not equivalent to better access in terms of quality and acceptability (Macintyre, Maciver, and Sooman 1993).

Features of the built environment also play a role in physical activity levels of local residents (Sallis and Glanz 2009). One of the differences accounting for the disparities between counties is leisure-time walking, which is a common daily activity among urban residents. The differences may be partly due to urban planning projects that promote mixed-land use designs which encourage walking among residents who are able to take advantage of the convenience of work, recreation, and services that are within a close proximity of their home (Frank, Andresen, and Schmid 2004). The presence of sidewalks, adequate lighting, and population density as opposed to heavy traffic, pedestrian unfriendly land use, and vacant land is not supportive of physical activity (Berrigan and Troiano 2002). Studies show that residents of sprawling counties are more likely to report hypertension and obesity than residents of compact counties, even after controlling for food consumption and demographic characteristics (Ewing et al. 2003). Thus, modifications to the built environment, which promote physical activity and increased access to services, are imperative (Handy et al. 2002; Lovasi et al. 2009).
Natural Environment

Natural amenities are important for psychological well-being (Ulrich 1981) and stress-reduction (Wells and Evans 2003). Researchers, including sociologists have observed a relationship between natural amenities and economic development, particularly in nonmetropolitan communities (Hunter, Boardman, and Onge 2005; Albrecht 2004). Ethnographic research highlights the effect of nature on health promotion in urban settings (Altchueler, Somkin, and Adler 2004). Dutch researchers have examined the relationship between the natural environment and health, and their findings show that the natural environment has significant positive effects on self-reported health indicators, and the effect of the natural environment on individuals with lower education attainment was significant and stronger in magnitude than those with higher education credentials (Vries et al. 2003). However, U.S.-based researchers have mostly neglected but should more thoroughly explore the relationship between various landscapes and landscape elements that encourage physical activity and promote health (Velarde, Fry, and Tveit 2007). For example, researchers in forestry and wildlife sciences have explored the relationship between natural resource amenities and life expectancy at the county level in the coterminous United States (Poudyal et al. 2009). Poudyal and colleagues find that natural amenities have a significant and positive effect on longevity. More specifically, counties with longer life expectancies also have milder weather, longer mean sunlight hours in January, and boast more water as a percentage of land area. The authors conclude that the public health community is ignoring the benefits of promoting the natural environment as an inexpensive and existing resource for physical activity which may enhance public well-being and quality of life (Poudyal et al. 2009).
Summary

The literature raises important questions about the relationship among the environment, social conditions, and obesity prevalence. My research makes two important contributions to the literature. First, while several studies have examined the relationship between the built environment and obesity these studies tend to be restricted to a few geographical areas (Morland 2007; Rundle et al. 2009). In this study, I adopt a national perspective by exploring the relationship between aggregate level attributes and obesity among all coterminous counties in the United States. Using county as my lens, I am able to examine a more complete set of variables on the social, built, and physical environment. Second, my work extends the analytical perspective, moving beyond traditional aspatial ordinary least square regression methods. Specifically, I use geographically weighted regression (GWR), an exploratory technique that accounts for heterogeneous relationships between each predictor and obesity. This spatially explicit approach can provide insight into where and how the relationship between the independent variables and obesity vary by context. This type of approach has several implications for public policy and future research. Few studies have employed GWR to examine contextual influences on obesity prevalence, at any aggregate level. This empirical analysis provides a useful demonstration of how GWR may be used to examine how context shapes a variety of health outcomes, including obesity.
Chapter III
Hypotheses

The goal of this study is to examine contextual level influences on obesity prevalence in the continental United States. Based on my literature review of research on health disparities, and the social, built, and physical environment discussed in Chapter II, I have the following related hypotheses:

Typology of place

Hypothesis 1: Residents of rural counties have higher obesity rates than urban residents; therefore, I expect that rural counties will have higher obesity rates than urban counties after controlling for other variables.

Macro-Social Processes

Hypothesis 2: Minority status is often associated with poor health outcomes and higher obesity prevalence, especially among African-American, Hispanic, and Native-American groups. Therefore, I predict that counties with a higher percentage of minority groups, with the exception of Asian-Americans, also report higher obesity prevalence after controlling for other variables.

Hypothesis 3: While there is conflicting evidence between the association of obesity and relative deprivation, most studies conclude that income inequality has deleterious effects on health outcomes. Therefore, I predict that counties with lower income inequality have lower obesity prevalence.
**Built Environment**

**Hypothesis 4**: Access to healthy food options is imperative to maintain a healthy weight status. Thus, counties that have a higher percentage of residents that do not own a vehicle and live more than one and ten miles from a supermarket will also have higher obesity prevalence after controlling for other variables.

**Hypothesis 5**: Access to supermarkets is associated with more nutritious dietary behaviors, while the inverse has been shown for access to convenience stores. Therefore, I predict that counties that have experienced a net loss of supermarkets per capita from 2000-2008, will also have higher obesity rates. The inverse will be the case for counties that witnessed a loss of convenience stores per capita after controlling for other variables.

**Physical/Natural Environment**

**Hypothesis 6**: Limited research has shown that the natural environment has positive effects on self-reported health, life expectancy, and well-being. While controlling for other variables, I predict that the natural amenity scale, which measures the attractiveness of the climate and topography, is negatively associated with obesity.

**Health Behavior/Lifestyle measure**

**Hypothesis 7**: A sedentary lifestyle is a risk factor for obesity. Hence, I hypothesize that counties whose residents report high physical inactivity also report higher obesity prevalence.
Chapter IV
Data and Measures

Data
To examine the spatial distribution of obesity in the United States, I constructed a rich county-level dataset derived for multiple sources. I data extracted from the United States Department of Agriculture (USDA) Food Environment Atlas Dataset (henceforth, FEAD), which was released in 2010. The dataset includes 168 measures of the food environment which provides contextual variables on factors that influence food choices, diet quality, and food access. All counties or county-equivalents with complete data (n=3108) were included in the analysis.

I extracted demographic information on the racial composition and socioeconomic characteristics from FEAD based on the U.S. Census Bureau’s 2008 county population estimates. Environmental variables of grocery store access and proximity and the natural amenity index were also extracted from the Food Atlas. FEAD has a dummy variable for metropolitan and nonmetropolitan counties. However, the variable does not allow one to manipulate the variable into other dichotomous features such as rural counties vs. urban counties or metropolitan adjacent vs. non-adjacent counties. Therefore, I utilized the USDA rural-urban continuum codes, and merged the data with selected variables from FEAD. The rural-urban continuum code ranges from 1-9 and identifies rural and urban counties in the United States. The 2003 rural-urban continuum, which is classified by the USDA, distinguishes counties by degree of urbanization and proximity to metropolitan counties (see Appendix A, Table 1).

1 Counties or county equivalents in Alaska and Hawaii were not included in the analysis due to missing data for several measures. There was also missing data for the obesity measure for the county-equivalent Clifton Forge, Virginia. Therefore, I merged the independent city of Clifton Forge, Virginia to Allegheny County, Virginia. Clifton Forge is an enclave of Allegheny County. The elimination or merging of 33 selected counties leaves a sample size of 3108 counties for analysis.
The U.S. Census County Business Patterns (CBP) data was used to estimate change in the food retail landscape at the county level from 2000 to 2008. The CBP provides annual data on industry for state, counties, and metropolitan areas. The CBP includes data on employment size, annual payroll, and the total number of establishments in each sector. Sectors are segmented by The North American Industry Classification System (NAICs) codes, which allows for easy identification of industries (Census, 2011). The CBP data was extracted from the Business Register (BR), which details information about current business surveys, federal income and payroll tax records, and administration records for business establishments in the United States (Census, 2011). The CBP is released 18 months following each enumeration.

The measure of physical inactivity was obtained from the Centers from Disease Control and Prevention National Diabetes Surveillance System. The measure of physical inactivity was estimated using the CDC’s Behavioral Risk Factor Surveillance System (BRFSS). According the CDC, the BRFSS is a continuous, monthly survey of the adult population. The BRFSS includes data from all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam. The BRFSS if the largest telephone health surveys in the world, with more than 350,000 adults surveyed each year. The data provides state-specific data on risk behaviors related to personal behaviors and preventive health practices. The data is an accumulation of three years of data to increase the precision of year-specific county-level estimates of risk factors. The BRFSS estimates 3-year (2007, 2008, and 2009) county-level statistics of diabetes and risk factors such as physical inactivity and obesity (CDC 2011).

The Gini coefficient is a measure of inequality calculated from county-level household income distributions. The Gini coefficient can range from 0 to 1. A value of “0” indicates complete equality, meaning that income is equally distributed among the population. In contrast,
a Gini coefficient of “1” indicates extreme inequality and income is concentrated in one household. The Gini coefficients for this study are provided by an online-published dataset (Burkey, 2006).

Measures

Dependent Variable

The dependent variable is the percentage of adults who are obese at the county level. The variable is estimated from the BRFSS using three year counts from 2007, 2008, and 2009 to determine the age-adjusted percentage of adults (age 20 or older) with obesity. Obesity is defined by a body mass index (BMI) of ≤30. The body mass index is calculated by dividing an individual’s weight by their height. The percentage of adults who are obese at the county-level in the United States ranges from 12.5 to 43.5, with a mean of 28.28 and a standard deviation of 3.61.

Independent Variables

Rural vs. Urban Counties

To examine differences in obesity by rural and urban counties, a rural-urban dummy variable was derived from the USDA rural-urban continuum. County classifications range from 1-9, with counties 8 and 9 classified as rural counties. The counties coded a “1” in this study are defined as completely rural or less than 2,500 urban population. Counties coded as “0” are designated as urban counties. There are 656 rural counties, which represent 21% of the sample.
Racial Composition and Total Population

Racial and ethnic composition of each county is examined by accounting for the percentage of Blacks, Hispanics, Asians, and Native Americans. The racial composition of each county is provided in the USDA Food Atlas dataset, which were derived from the U.S. Census 2008 County Population Estimates. The percentage of Pacific Islanders and Alaskan Natives are collapsed into the percentage of Native Americans in this study. The percentage of the white population was not included in the analysis to avoid issues of multicollinearity. The total population for each county, which can also account for variation with the rural-urban binary independent variables, was included in the analysis as a control variable.

Inequality

The Gini coefficient is used to examine the effects of inequality on county-level obesity (Burkey 2006). The values of the Gini coefficient range from 0-1, in which a value of 1 indicates complete inequality, and a value of 0 indicates complete equality across individuals in society. The coefficient appropriately captures income inequality as well as other measures of relative deprivation (i.e., Robin Hood Index, Atkinson Index, and Thiel’s entropy measure) (Kawahchi & Kennedy, 1997). There is a wide variation in the Gini coefficient by county (ranges from about .31 to .60). I divided the coefficient into tertiles to create a categorical variable of low inequality, medium inequality, and high inequality. The measure of inequality was then coded as a dummy variable to compare differences in obesity prevalence among counties with low inequality versus counties with medium and high inequality. Counties with low inequality are coded as “1” in this analysis, while counties with medium and high levels of inequality are coded as “0”. Counties with low inequality range from approximately 0.31 to about 0.47.
Proximity to Grocery stores

To examine the effects of access and spatial proximity to grocery store or supermarket, two measures of access and proximity were analyzed in this study. The first variable is the percentage of households in a county without a car and more than one mile to a supermarket or large grocery store; the second variable reveals the number of households in each county without a car and more than 10 miles to a grocery store. The variables were provided in the USDA Food Atlas Dataset, while the source of the data is from a report to Congress titled Access to Affordable and Nutritious Food – Measuring and Understanding Food Deserts and Their Consequences. The methods for the proximity measures were derived by creating a directory of supermarkets and large grocery stores from a 2006 Trade Dimensions, a directory of commercial retailers, and the STARS store directory, a directory of retailers that accepts food stamp benefits. A store must sell fresh meat and poultry, dairy, dry goods, and frozen foods; and report annual sales of $2 million or more to be classified as a grocery store or a supermarket (USDA, 2011).

Access to grocery stores in the report is determined by examining vehicle ownership at the block group level and then allocating the data aurally to 1-square kilometer grids across the continental United States. This technique allowed for easier spatial computation. Finally, the distance from the centroid of each grid to the nearest grocery store or supermarket was calculated.

Both access variables were transformed by taking the square root of the original variable to comply with the assumption of normal distribution in ordinary least squares regression.
Environmental change

Counts of grocery stores and convenience stores for each county were calculated from the CBP for years 2000 and 2008. The full 6-digit NAICS code was used to extract the data for grocery stores (445110) and convenience stores (445120) in this analysis. The variable for each store type was calculated the same for this analysis. First the count of stores for each county was divided by the total population in the county for the respective year and the square root of both variables were calculated to ensure normal distribution. Second, the number of grocery (convenience) stores per capita for 2008 was subtracted from the number of stores per capital in 2000 to create a change variable. Then, I dichotomized the variable into a dummy variable to compare the counties that experienced a net loss of grocery stores (convenience) (=1) to the counties that either gained or had the same number of grocery stores (convenience) (=0) from 2000 to 2008. About 75% of counties witnessed a decrease in the number of grocery stores per capita, while approximately 61% of counties witnessed a decrease in the number of convenience stores.

Amenity

The natural amenity resources scale is provided in the USDA Food Atlas Dataset which is constructed from the USDA Economic Research Service. The scale is based on access to natural phenomenon such as waterfronts, mountains, and temperate climates. The scale ranges from 1-7, with a value of 1 being the lowest score on the scale and 7 being the highest score on the scale. The index is based on the standard deviation of the mean value of the amenity scale for all counties. Therefore, counties at the lower end of the continuum are below the mean and counties with amenities above the mean have a high score on the continuum.
Physical Inactivity

The categorical variable of physical inactivity among adults by county was derived from the CDC’s Behavioral Risk Factor Surveillance System (BRFSS). The original variable was segmented into 9 gradations of physical inactivity by the number of physically inactive adults in a county. For this study, the variable was recoded and segmented into 3 categories of physical inactivity (see Table 1).
Table 1: Recode of gradients of physically inactive adults in the United States.

<table>
<thead>
<tr>
<th>Original Codes</th>
<th>Recodes</th>
<th>% of Physically Inactive Adults</th>
<th>Total number of Physically Inactive adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0 to 22.4%</td>
<td>0 to 70310 adults</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>22.5 to 29.4%</td>
<td>0 to 70310 adults</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>≥ 29.5%</td>
<td>0 to 70310 adults</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0 to 22.4%</td>
<td>70311 to 329000 adults</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>22.5 to 29.4%</td>
<td>70311 to 329000 adults</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>≥ 29.5%</td>
<td>70311 to 329000 adults</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0 to 22.4%</td>
<td>≥329001 adults</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>22.5 to 29.4%</td>
<td>≥ 329001 adults</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>≥ 29.5%</td>
<td>≥ 329001 adults</td>
</tr>
</tbody>
</table>

Several statistical techniques were used to examine the relationship between county level obesity prevalence and independent variables. Ordinary least squared (OLS) regression, Moran's $I$, and geographically weighted regressions (GWR) are employed in this ecological study.

OLS is a conventional approach that minimizes the sum of the squared residuals between the predicted dependent variable and the observed dependent variable (Allison, 1999). The traditional regression equation can be expressed as

$$\hat{Y}_i = \beta_0 + \sum \beta_k x_{ik} + \epsilon_i$$

where $\hat{y}$ is the predicted value of the dependent variable for observation $i$, $\beta_0$ is the intercept, $\beta_k$ is the parameter estimate for the independent variable, $x_{ik}$ represents the $i$th observation of the $k$th independent variable, and $\epsilon_i$ represents the error term.

Several assumptions of OLS regression must be met to calculate unbiased regression coefficients. The first set of OLS assumptions addresses the independent variables (Long 1997). First, the relationship between the dependent and independent variables must be linear. Variables that are nonlinear must be transformed by using a scientific notation such as a squared, square root, inverse, or logarithmic function. Highly collinear independent variables also violate assumptions of OLS regression. Multicollinearity diagnostics were ran in STATA 11.1 to test for violation of collinearity between independent variables. The variance inflation factor (VIF) (reported in Table 2) and the condition number were observed to detect potential issues with multicollinearity in this study. The threshold for the VIF is debatable. Allison (1999) suggests a
VIF cut-off of 2.5, while others argue that a VIF below 10 is appropriate (Kutner, 2004). For this study, I decided to use a more conservative VIF threshold of 2.5, because the multicollinearity condition number became highly volatile when the VIF exceeded a value of 2.5. I will elaborate on variables that presented issues with multicollinearity later in results section.

The next set of assumptions of OLS regression addresses the distribution of the error term, or the disturbance term (Long 1997; Allison 1999). The error term must be normally distributed and mean independent; therefore, the mean value of the error term should equal zero. The error term is also assumed to be homoskedastic or uncorrelated with other errors in the population sample. These set of assumptions are the most relevant to the spatial attributes of the data used in this thesis. Neighboring counties likely share similar compositional and contextual characteristics that are spatially correlated. More specifically, a set of conditions, in this case obesity and several predictors of obesity, of one county may affect obesity outcomes in neighboring counties. The OLS model assumes that the conditions in each county observed are independent of proximal counties. This assumption is likely to be violated. Thus, the parameters calculated by the OLS will likely be biased, and generate low estimates of the confidence intervals and real variance (Anselin 1988; Ward and Gleditsch 2008).

Spatial dependence occurs when social processes are not randomly distributed across place (Mennis and Jordan 2005), while spatial heterogeneity (non-stationarity) occurs when the relationship between independent and dependent variables is not consistent across place. There are several statistical techniques that account for the spatial dependence or spatial heterogeneity among places (Ward and Gleditsch 2008). Both spatial dependence and spatial heterogeneity are easier to observe with choropleth mapping and Moran’s I scatterplots (Anselin 1995). I use the Moran’s I to test for spatial autocorrelation of each variable included in the OLS model residuals.
The Moran’s $I$ is a global measure of spatial autocorrelation between the values of an observation with other local values conditional on a spatial weight matrix (Ward and Gleditsch 2008, p.23). An obesity variable with a positive Moran’s $I$ value is one with a high degree of spatial clustering (i.e., the correlation of attribute values in neighboring areas) is interpreted as counties with high obesity prevalence surrounded by counties with high obesity prevalence, or counties with low obesity prevalence surrounded by other counties with low obesity prevalence. A negative Moran’s $I$ is interpreted as counties with high obesity prevalence that are adjacent to counties with low obesity prevalence. A significant Moran’s $I$ for the dependent and independent variable suggests that spatial dependence may be important and that there is potential for significant spatial effects. Moreover, a significant Moran’s $I$ in the OLS model residuals supports evidence that some assumptions of OLS have been violated and spatial methods may provide a more robust model.

The import of place is one of the central premises of this thesis. I assume the mechanisms or social stimuli that operate on obesity outcomes within each county are not likely stationary across place. A priori assumptions of spatial dependency and non-stationarity lends to the inclusion of a GWR model. GWR corrects for both spatial heterogeneity and spatial dependence of the data by estimating local, more robust parameters that capture the distinctiveness of place, and the spatial variations between dependent and independent variables (Brunsdon et al. 1996).

One of the main differences between traditional regression methods and GWR is that traditional methods produce a global model that predict the outcome variable across the entire study area, while GWR generates a local model for each observation in the sample (i.e., drawing
on a sub-sample of the total number of observations and weighting the relationships between variables based on a distance-based criteria).

The GWR model can be expressed with the following function:

\[ y_i = \beta_{0i}(u_i, v_i) + \sum_k \beta_{ki}(u_i, v_i) x_{ik} + \epsilon_i \]

\( y_i \) is the percentage of the adult population reporting obesity for county \( i \), \((u_i, v_i)\) denotes the coordinates of the centroid of county \( i \), \( \beta_{0i} \) and \( \beta_{ki} \) represents the local estimates intercept and the effect of variable \( k \) on county \( i \).

Each county centroid, \( i \), is a regression point, and observations closer in proximity to \( i \) are weighted more than counties that are farther away (Fortheringham, Brundson, and Charlton 2002). In other words, neighboring counties have more influence than distant counties. An adaptive weighting scheme, as used in the analysis reported here, accounts for the variation in the size of counties, and the sparseness of counties in certain regions of the United States, most notably counties in the West (Nakaya, 2005). In contrast to a fixed spatial kernel, which treats the bandwidth for estimating each local model as constant, the adaptive weighting function permits a larger bandwidth when the data is sparse and a small bandwidth when data are dense (Fortheringham, Brundson, and Charlton 2002, p.46.). The bandwidth heavily influences the weighting scheme. The bandwidth is expressed in the same number of units included in the analysis. The large bandwidth approaches the parameter achieved in a global model, while a small bandwidth reveals local relationships (Charlton and Fotheringham 2009). The adaptive spatial kernel also ensures that an equal number of observations receive a non-zero weight value at each regression point (Bitter et al. 2007, p.15).

Following Fotheringham et al. (2002), the Akaike Information Criterion (AIC\(_c\)) was used to determine the optimal kernel bandwidth size. The AIC\(_c\) uses a bi-square decaying function
based on nearest neighbors used (Fortheringham, Brundson, and Charlton 2002). To test for spatial non-stationary, Monte Carlo significance testing was ran with the GWR model. The Monte Carlo approach is computationally intensive. The Monte Carlo test provides a randomized null hypothesis model, in which the locations of the observation are arbitrarily assigned to the predictor and response variables. The null hypothesis assumes there is no significant difference in patterns of parameters across place (Brunsdon, Fotheringham, and Charlton 1998). Hence, a significant Monte Carlo test reveals spatial variation in local response variables (Fortheringham, Brundson, and Charlton 2002).

This spatially explicit approach helps identify relationships that are likely masked by using traditional regression methods, which assumes that the relationship between variables are homogenous throughout the entire study area. However, GWR is not void of limitations (Wheeler and Paez 2010). One of the disadvantages of the GWR model is the trade-off between the spatial scale and statistical stability of the bandwidth. A small bandwidth size can lead to unreliable parameters, but large bandwidths present bias in estimated parameters (Fortheringham, Brundson, and Charlton 2002; Nakaya et al. 2005). GWR cannot correct a poorly specified model. Multicollinearity and model specification must be addressed beforehand.

GWR fails to provide a base regression model to estimate the local relationships observed between variables (Wheeler and Páez 2010). Also, as the number of local models increases, the probability of concluding that spatial non-stationarity exist also increases, and the relationship between parameters may not be significant (Wheeler and Páez, 2010).

The final GWR output was converted from GWR 3.0 to ArcGIS to map the local relationships of parameters at each regression point. The t-values of each variable were used to provide a visual interpretation of where (i.e., geographically speaking) variables tested
significant for spatial non-stationarity (Mennis, 2006). The local r-square was also mapped from output to ascertain where the model fits, or where the predictors best estimate the response variables.

The GWR model output also generates a 5-number summary of the local parameters. The summary table includes the median, upper, and lower quartiles, and the minimum and maximum values of parameter estimates. The varying magnitude and signs of the parameter estimates across the distribution represent spatial heterogeneity (Fortheringham, Brundson, and Charlton 2002). The 5-number summary distribution can be interpreted as follows: the minimum parameter reported is the smallest value estimated in the study area, 50% of the local parameter estimates will fall between the upper and lower quartiles, approximately 68% of parameter estimates will lie between ±1 standard deviation of the mean, and the maximum parameter reported is the largest value estimated in a local model.

Finally, the better regression method is determined by evaluating the Akaike Information Criterion (AIC) between the OLS and the GWR models. The model with the lowest AIC value signifies a better model fit. GWR is an exploratory method which may help identify a well specified global (OLS) model.
Chapter VI
Results

Descriptive Statistics (Numerical Values and Mapping)

Table 2 details the descriptive statistics of the dependent variable and all covariates. The mean percentage of adults, who are obese at the county level, as shown in Table 1, is 28.28 with a standard deviation of 3.61, with percentages ranging from 12.5% of the adult population obese to more than 43%. Figure 1 shows the distribution of obesity prevalence in the United States. This is the outcome variable and this is the distribution I would like to explain. The map reveals a concentration of obesity in the Southeast, along with the lower Appalachia region of West Virginia and Kentucky, with a few pockets of high obesity in the Dakotas and Nevada.

The dummy urban-rural variable used in the analysis mean value is .21, with a standard deviation of .40. The total population of U.S. counties varies widely. The average population count for a county in the contiguous United States is 90,000, with a standard deviation of approximately 300,000. The county with the smallest population, Loving County, Texas, has a population of 67. While the most populated county in the U.S., Los Angeles County, has a population of more than 9,500,000.

The racial composition of counties also varies widely. The mean percentage of the African-American population at the county level is 9%, while the average Hispanic population is about 8%. Native-Americans and Asian-Americans make up approximately 1% and 2% of the total population at the county-level, respectively. Figure 2 reveals the distribution of minority populations throughout the United States. African Americans are disproportionately located in the Black Belt South. The Hispanic population is highly concentrated in the West, especially along the Mexican border, as well as Southern Florida.
Table 2: Descriptive Statistics for percentage of obese population and contextual variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
<th>Moran’s I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obesity (%)</td>
<td>28.28</td>
<td>3.61</td>
<td>12.50</td>
<td>43.50</td>
<td>.65***</td>
</tr>
<tr>
<td>Urban/Rural Continuum</td>
<td>0.21</td>
<td>0.41</td>
<td>0.00</td>
<td>1.00</td>
<td>.22***</td>
</tr>
<tr>
<td>Total Population</td>
<td>89,956.06</td>
<td>293,557.60</td>
<td>67.00</td>
<td>9,519,338.00</td>
<td>.36***</td>
</tr>
<tr>
<td>Black (%)</td>
<td>9.00</td>
<td>14.32</td>
<td>0.00</td>
<td>85.50</td>
<td>.80***</td>
</tr>
<tr>
<td>Hispanic (%)</td>
<td>7.71</td>
<td>12.83</td>
<td>0.10</td>
<td>97.30</td>
<td>.81***</td>
</tr>
<tr>
<td>Asian (%)</td>
<td>0.98</td>
<td>1.89</td>
<td>0.00</td>
<td>30.90</td>
<td>.60***</td>
</tr>
<tr>
<td>Native Americans (%)</td>
<td>1.62</td>
<td>6.20</td>
<td>0.00</td>
<td>83.50</td>
<td>.37***</td>
</tr>
<tr>
<td>Gini</td>
<td>0.33</td>
<td>0.47</td>
<td>0.00</td>
<td>1.00</td>
<td>.35***</td>
</tr>
<tr>
<td>HH with no car, &gt; 1mile to grocery store (%)</td>
<td>1.90</td>
<td>0.61</td>
<td>0.00</td>
<td>5.28</td>
<td>.52***</td>
</tr>
<tr>
<td>HH with no car, &gt; 10 miles to grocery store (%)</td>
<td>0.61</td>
<td>0.63</td>
<td>0.00</td>
<td>4.04</td>
<td>.43***</td>
</tr>
<tr>
<td>△ in # of Grocery stores/capita</td>
<td>0.75</td>
<td>0.43</td>
<td>0.00</td>
<td>1.00</td>
<td>.11***</td>
</tr>
<tr>
<td>△ in # of Convenience stores/capita</td>
<td>0.60</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
<td>.13***</td>
</tr>
<tr>
<td>Amenity Scale</td>
<td>3.49</td>
<td>1.04</td>
<td>1.00</td>
<td>7.00</td>
<td>.79***</td>
</tr>
<tr>
<td>Physical Inactivity</td>
<td>0.17</td>
<td>0.38</td>
<td>0.00</td>
<td>1.00</td>
<td>.54***</td>
</tr>
</tbody>
</table>
Figure 1: Adult obesity prevalence in the United States (2008)

Adult Obesity Prevalence in the United States (2008)

Quantiles (%)

- 12.50 - 26.20
- 26.21 - 27.79
- 27.80 - 29.20
- 29.21 - 30.79
- 30.80 - 43.50

Source: Food Environment Atlas, USDA
Map Prepared by: Nyesha Cheyenne Black
Figure 2: Distribution of minority populations in the United States (2000)

Distribution of Minority Populations in United States (2008)

African-Americans

Hispanics

Asian Americans

Native Americans

Source: Food Environment Atlas, USDA
Map Prepared by: Nyesha Cheyenne Black
The Native American population is largely represented in Oklahoma and the Great Plains. The Asian-American population is concentrated in the West and scattered throughout metropolitan areas of the U.S.

Figure 3 features the distribution of income inequality in the U.S. Patterns of spatial clustering of inequality are evident in the band on Black Belt counties, Appalachia, South Texas, and South Florida. Interestingly, as shown in Figure 4, the spatial patterning of income inequality is similar to the spatial clustering of counties with percentage of households with no car, and more than one mile to a grocery store. Figure 4 also shows that counties with the least access to grocery stores are in the West, with small pockets in Louisiana, and the central counties of Alabama stretching into East Mississippi.

The mean amenity index in the continental United States is about 3.5. Counties with the highest amenity score are along the Pacific Northwest, South Florida, and a band of counties stretching from Southern Wyoming to Northern New Mexico. In contrast, there is a cluster of counties with a low amenity scale in the Midwest, as shown in Figure 5. Physical inactivity, which is the variable with the highest correlation with obesity, has a mean value of 2.13, which suggests that on average 22.5-29.4% of adults reported physical inactivity. There are clear patterns of reports of little to no exercise in the Southeast, small pockets in Northern Maine, and the Dakotas, as shown in Figure 6.
Figure 3: Distribution of income inequality (Gini) in the United States (2000)
Figure 4: Access and proximity to grocery stores in the United States (2006)

Access and Proximity to Grocery Stores in the United States (2006)

Percentage of Households with no car & >1 mile to a Grocery Store

Percentage of Households with no car & >10 miles to a Grocery Store

Quantiles (%)
- 0.00 - 1.39
- 1.39 - 1.70
- 1.71 - 1.98
- 1.99 - 2.37
- 2.38 - 5.28

Quantiles (%)
- 0.00
- 0.01 - 0.34
- 0.35 - 0.62
- 0.62 - 1.05
- 1.06 - 4.04

Source: Food Environment Atlas, USDA
Map Prepared by: Nyasha Cheyenne Black
Figure 5: Distribution of amenity scale in the United States (1999)
Figure 6: Estimates of physically inactive adults in the United States (2008)

Source: Centers for Disease Control and Prevention: National Diabetes Surveillance System.
Map Prepared by: Nyesha Cheyenne Black
Global Model Results

Results from the global regression model are provided in the Table 2. I explored several other response variables to examine the relationship between the social determinants of disease and obesity. Including measures of poverty, unemployment, education attainment, and relative deprivation raised concerns due to high multicollinearity between the independent terms. Since I did not want to compromise the integrity of the model, I decided to only retain the Gini coefficient in the final model.

Table 2: OLS Regression model predicting obesity (Global Regression Model)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>28.00***</td>
<td>0.26</td>
<td>---</td>
</tr>
<tr>
<td>Urban/Rural Continuum</td>
<td>-0.20</td>
<td>2.68</td>
<td>1.46</td>
</tr>
<tr>
<td>Total Population</td>
<td>-4.52E-07**</td>
<td>1.72E-07</td>
<td>1.45</td>
</tr>
<tr>
<td>Black (%)</td>
<td>0.08***</td>
<td>0.00</td>
<td>1.46</td>
</tr>
<tr>
<td>Hispanic (%)</td>
<td>-0.02***</td>
<td>0.00</td>
<td>1.27</td>
</tr>
<tr>
<td>Asian (%)</td>
<td>-0.13***</td>
<td>0.03</td>
<td>1.65</td>
</tr>
<tr>
<td>Native Americans (%)</td>
<td>0.09***</td>
<td>0.01</td>
<td>1.13</td>
</tr>
<tr>
<td>Gini †</td>
<td>0.40***</td>
<td>0.10</td>
<td>1.24</td>
</tr>
<tr>
<td>HH with no car, &gt; 1mile to grocery store (%)</td>
<td>1.46***</td>
<td>0.10</td>
<td>1.97</td>
</tr>
<tr>
<td>HH with no car, &gt; 10 miles to grocery store (%)</td>
<td>-0.36***</td>
<td>0.90</td>
<td>1.79</td>
</tr>
<tr>
<td>Δ in # of Grocery stores/capita†</td>
<td>0.26**</td>
<td>0.10</td>
<td>1.09</td>
</tr>
<tr>
<td>Δ in # of Convenience stores/capita †</td>
<td>0.20*</td>
<td>0.09</td>
<td>1.10</td>
</tr>
<tr>
<td>Amenity Scale</td>
<td>-0.78***</td>
<td>0.05</td>
<td>1.33</td>
</tr>
<tr>
<td>Physical Inactivity †</td>
<td>-3.05***</td>
<td>0.12</td>
<td>1.26</td>
</tr>
<tr>
<td>Adjusted R-Square</td>
<td>.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akaike Information Criterion</td>
<td>14115.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p≤0.05; ** p≤ 0.01; ***p≤ 0.001

†Dummy Variables: Gini (1=low inequality); Δ in # of Grocery stores/capita (1=loss of grocery stores); Δ in # of Convenience stores/capita (1=loss of convenience stores); and Physical Inactivity (1=low physical inactivity)
The OLS results show that obesity prevalence is not significantly different between rural and urban counties, and the effect of the total population on adult obesity is significant, but negligible. All of the race variables in the model are significant. A presence of Native-Americans and African-Americans at the county-level is positively associated with obesity. In fact, a one percentage point increase in the African-American population and the Native-American population is associated with a .08 and a .09 percentage increase in obesity, respectively.

The dummy variable for the Gini coefficient reveals that counties with low income inequality have an adult obesity prevalence that is .40 percent higher than counties with medium or high income inequality. This result contradicts the premise that low inequality, relative to medium and high inequality, is a safety net for poor health outcomes.

The results of the access and proximity variables show that a percentage increase in the number of county residents without a car, and more than one mile from a grocery store is associated with a 1.46 percent increase in adult obesity at the county level. However, as the percentage of the county population who lack a car, and live more than 10 miles from a grocery store increases by one percentage point, the percentage of obese adults decreases by .36 percent. These somewhat contradictory results suggest a curvilinear relationship between proximity to grocery stores and obesity.

Counties that witnessed a per capita loss in the number of grocery stores reported obesity prevalence .26 percent higher than counties that did not experience a net change or a net gain in the number of grocery stores per capita. It is expected that counties that witnessed a decrease in
the number of convenience stores will expect to have a lower prevalence of obesity. However, results from the global model reveal the opposite. Counties that witnessed a per capita loss in the number of convenience stores have an estimated obesity prevalence that is .19 percent higher than counties that did not experience a net change or a net gain in the number of grocery stores per capita. These findings suggest that convenience stores are imperative to food retail landscape.

The amenity scale reveals a negative and significant effect on adult obesity. A one point increase on the amenity scale is associated with a .78 percent decrease in adult obesity. While having access to natural resources has a negative effect on obesity prevalence, the largest effect on obesity is observed by counties with lower rates of physical inactivity. Adult residents of counties who reported the highest rates of physical activity have obesity prevalence that is almost 3 percent lower than counties with lower percentages of physically active adults.

The variance inflation factor for each variable is listed in Table 2, and all measures have a VIF of 1.40 or less, which is below the threshold of 2.5 (Allison 1999). The global model explains 58% of the total variance in the percentage of adults who are obese with an AIC of 14,115.13. The AIC has little meaning without a model for comparison. The Moran’s I presented in the Table 1 shows that there is significant spatial autocorrelation. The percentage among all minority groups is highly spatially correlated between counties. Obesity, physical inactivity, and proximity measures of supermarket access is also notably spatially correlated among counties. Mapping the local Moran’s I of the standardized OLS residuals (see Figure 7) facilitates the identification of neighboring counties with similar residuals (i.e., a spatial
patterning of the residuals), and isolate outliers in the sample which may help explain differences in the relationship between obesity and predictors (Waller and Gotway 2004).
Figure 7: Local Moran’s I for standardized OLS residuals
Local Model Results

Further analysis is needed to determine spatial heterogeneity in the outcome variables, and whether a more local, nuanced model is more robust. GWR 3.0 software was used to compute the results for the local regression models. The GWR 5-number summary and Monte Carlo significance tests for spatial heterogeneity of parameter test results are featured in Table 4. The Monte Carlo significance test in the model indicates that most of the parameters in the model are non-stationary across space. The AIC in the GWR model is 12,635.10, and the adjusted r-square indicates that the model explains 78% of the variance, both of which confirms that the GWR model is a better fit than the global model. The AIC determined bandwidth in the GWR model is 257 counties. That is, each local GWR model is estimated on data from 257 counties, which is far less than the 3,108 counties in the global analysis. The results of the Monte Carlo test reveal that the differences in adult obesity between rural and urban counties are stationary across space. Both measures of change in the food retail landscape are also stationary across the continental United States. In other words, there is no spatial patterning to the relationship between obesity and metropolitan or non-metropolitan status or change in the food retail landscape across the coterminous United States.

Figure 8 profiles the local r-square of the GWR model. The explained variance in the model varies greatly from about 30% to approximately 88%; recall the global “stationary” model predicts a single adjusted r-square of 58%. The map shows that the model predicts adult obesity relatively poorly in much of the Appalachia region, the Midwest, and the Pacific Northwest. In contrast, the model is very predictive of adult obesity in the Southeast, in several of the Plain States and the Southwest. The parameters significant according to the Monte Carlo test are discussed. All of the racial/ethnic composition variables were significant in the GWR model.
### Table 4: Geographically Weighted Regression 5 number parameter summary results and Monte Carlo significance test for spatial variability of parameter (N=3,108)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Lower Quartile</th>
<th>Median</th>
<th>Upper Quartile</th>
<th>Maximum</th>
<th>Monte Carlo Testing Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>21.00</td>
<td>26.86</td>
<td>28.26</td>
<td>29.54</td>
<td>34.03</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Urban/Rural Continuum</td>
<td>-2.29</td>
<td>-0.40</td>
<td>-0.12</td>
<td>0.15</td>
<td>1.85</td>
<td>Stationary</td>
</tr>
<tr>
<td>Total Population</td>
<td>-2.00e-05</td>
<td>-3.00e-06</td>
<td>-1.00e-06</td>
<td>0.00e+00</td>
<td>1.10e-05</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Black (%)</td>
<td>-0.57</td>
<td>0.06</td>
<td>0.10</td>
<td>0.13</td>
<td>1.57</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Hispanic (%)</td>
<td>-0.27</td>
<td>-0.04</td>
<td>-0.00</td>
<td>-0.03</td>
<td>0.16</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Asian (%)</td>
<td>-3.61</td>
<td>-0.34</td>
<td>-0.18</td>
<td>-0.06</td>
<td>0.89</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Native Americans (%)</td>
<td>-1.18</td>
<td>0.07</td>
<td>0.13</td>
<td>0.28</td>
<td>4.67</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Gini †</td>
<td>-1.19</td>
<td>-0.39</td>
<td>0.34</td>
<td>0.84</td>
<td>2.87</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>HH with no car, &gt; 1 mile to grocery store (%)</td>
<td>-0.66</td>
<td>0.01</td>
<td>0.38</td>
<td>1.08</td>
<td>3.22</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>HH with no car, &gt; 10 miles to grocery store (%)</td>
<td>-2.17</td>
<td>-0.40</td>
<td>-0.15</td>
<td>0.14</td>
<td>1.31</td>
<td>Non-stationary*</td>
</tr>
<tr>
<td>Δ in # of Grocery stores/capita †</td>
<td>-1.73</td>
<td>-0.16</td>
<td>0.07</td>
<td>0.38</td>
<td>1.74</td>
<td>Stationary</td>
</tr>
<tr>
<td>Δ in # of Convenience stores/capita †</td>
<td>-1.28</td>
<td>-0.16</td>
<td>0.04</td>
<td>0.07</td>
<td>0.34</td>
<td>Stationary</td>
</tr>
<tr>
<td>Amenity Scale</td>
<td>-2.29</td>
<td>-0.69</td>
<td>-0.34</td>
<td>-0.02</td>
<td>0.07</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Physical Inactivity †</td>
<td>-7.38</td>
<td>-3.17</td>
<td>-2.25</td>
<td>-1.55</td>
<td>3.51</td>
<td>Non-stationary***</td>
</tr>
<tr>
<td>Adjusted R-Square</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akaike Information Criterion</td>
<td>12635.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p≤0.05; ** p≤ 0.01; ***p≤ 0.001

†† Dummy Variables: Gini (1=low inequality); Δ in # of Grocery stores/capita (1=loss of grocery stores); Δ in # of Convenience stores/capita (1=loss of convenience stores); and Physical Inactivity (1=low physical inactivity)
Figure 8: Local r-square of GWR model
Figure 9 reveals areas in the continental United States where the association between obesity and the percentage of each minority group is significant. African-Americans are positively associated with obesity in most of the United States. The positive association is the highest in some of the Western states that have small African-American populations. There is a strong positive effect between the percentage of Hispanics and obesity in the Southwest, Northwest, and Central and South Florida. Interestingly, these regions host the largest Hispanic population in the U.S. However, there is a significant negative relationship between obesity and the Hispanic population along a band of counties stretching from Northern South Carolina, the lower Appalachia region stretching from East Tennessee to Central West Virginia. There is also a noticeable significant negative population in Southeast New York and the Northern Great Lake area of Wisconsin. In contrast to where there is a positive effect of the Hispanic population on obesity, the areas of the U.S. with a negative association between Hispanics and obesity have small Hispanic populations. The strongest significant positive effect between obesity and the Asian-American population is in Central Nebraska and Kansas, as well as some of the panhandle region of Oklahoma. This spatial pattern is surprising given the small racial composition of this group in the area. There is significant negative relationship between obesity and the Asian-American population in much of Wyoming and East Montana. The negative effect is all observed in the Alabama Black Belt and the Florida Panhandle. The significant effect of the Asian variable is also observed in Pacific Northwest, the New England States, and the Chicago and D.C. metropolitan areas. These areas are also home to relatively large Asian populations. The percentage of Native-Americans at the county level is significantly associated with obesity for a large portion of the United States west of the Mississippi River. The largest positive effect
is observed in Florida and Southern Georgia, as well as the Illinois-Missouri border, and much of West Virginia.
Figure 9: GWR estimates of minority populations

GWR Estimates for Minority Populations (%)

African American Population

Effect
High: 1.57
Low: -0.57

Hispanic Population

Effect
High: 0.12
Low: -0.27

Asian Americans

Effect
High: 0.89
Low: -3.61

Native Americans

Effect
High: 4.67
Low: -1.18

Map and Calculations Prepared by: Nyesha Cheyenne Black
The effect of income inequality on obesity is significant in the West and along the Atlantic Coast stretching from South Carolina to the New England. The results are presented in Figure 10. While the counties with lower income inequality in the West have higher adult obesity prevalence, there is an inverse relationship in Middle West Virginia. The GWR results show that low inequality is operating differently in West Virginia than in other parts of the country, in which low inequality has baleful consequences for adult obesity.

Both of the proximity measures in the study are significant according to the Monte Carlo test; however, the effect of the parameter is not significant in most of the United States. Figure 11 highlights the effect of vehicle ownership and proximity to a grocery store on obesity prevalence. The association between the one mile distance measure and obesity is significant and positive in California, with the exception of Northern and Southern California, as well as the lower Appalachia Mountain region. The ten mile distance parameter is negative in the global model, but there are some notable places where the variable has a positive association with obesity. The positive effects are seen in East Arizona, stretching into parts of New Mexico; most of Utah, parts of Southeast Idaho and Southwest Wyoming; a small region of Alabama on the Georgia state line, and the metropolitan area spanning from D.C. to Philadelphia. In contrast, the map shows that residing more than 10 miles from a grocery store is negatively associated with obesity in a band of counties along the Illinois-Indiana state line, Southeast Missouri, Southwest Kentucky, Central West Virginia, and much of the Northwest.

Figure 12 features the amenity scale, and the physical inactivity variable is displayed in Figure 13. The ability to have access to natural resources is significant in some regions of the contiguous United States. The effect of the amenity scale on obesity is significant, and negative in parts of Colorado, Wyoming, Idaho, and Utah. It is not surprising given that this region has
low obesity prevalence and also scores high on the amenity index. In contrary, the amenity scale has a positive effect on obesity and along the Iowa-Missouri-Illinois border. The physical inactivity dummy variable is significant for the majority of the continental United States. The OLS model reveals that counties where adults report high leisure-time physical activity have an obesity prevalence that is more than 3% lower than counties where adults report high levels of physical inactivity. The negative effect of physical activity on obesity is more prominent in the Texas Panhandle; the Western Nebraska and Western Kansas; much of New Mexico, Arizona, and Utah; and parts of Eastern Kentucky.
Figure 10: GWR estimates of income inequality

GWR Estimates of Income Inequality

Effect
High : 2.87
Low : -1.18

Map and Calculations Prepared by:
Nyesha Cheyenne Black
Figure 11: GWR estimates of access and proximity to grocery stores

GWR Estimates for Access and Proximity to Grocery Stores

Percentage of Households with no car & >1 mile to a Grocery Store

Effect
High : 3.22
Low : -0.66

Percentage of Households with no car & >10 miles to a Grocery Store

Effect
High : 1.31
Low : -2.17

Map and Calculations Prepared by:
Nyesha Cheyenne Black
Figure 12: GWR estimates of amenity scale
Figure 13: GWR estimates of physical inactivity
Chapter VII
Discussion and Conclusions

This paper has examined how aggregate social characteristics and other contextual variables shape obesity prevalence at the county level. The influence of context on health outcomes is complex, and health researchers have been attempting to unravel this complexity for more than a decade. I also extend upon traditional aspatial regression analyses by employing GWR. This approach accounts for the unique attributes of counties and the spatially varying relationships between the outcome and predictor variables across the continental United States. The results from this exploratory GWR study yield several interesting findings, many of which capture the salience of place.

First, I find that obesity prevalence is not significantly different in urban counties and rural counties. This finding was confirmed in the GWR model. In addition, other typologies of place—metropolitan versus nonmetropolitan and adjacent metropolitan counties and non-adjacent counties—were tested, and did not result in significant differences in obesity prevalence, controlling for other covariates. These findings suggest that other characteristics of counties explain aggregate level obesity prevalence.

The racial composition of the county is associated with obesity prevalence. However, the relationship between racial and ethnic composition and obesity varies by race. The percentage of African-Americans and Hispanics is positively associated with obesity prevalence, while the opposite relationship is observed among Asians and Hispanics. Hispanics and Asian Americans are more likely than other minority groups to have ties to immigration; therefore, the relationship between obesity and the percentage of Asian Americans and Hispanics may be moderated by immigrant status. On average, immigrants tend to be healthier due to selection effects (Wingate
and Alexander 2005). Studies have shown that acculturation and personal health behaviors may undergird differences among minorities with immigrant ties (Gordon-Larsen et al. 2003).

My third hypothesis was not supported by the OLS model. Counties with low income inequality have higher obesity prevalence in the global model. However, the relationship between income inequalities operates differently in parts of Appalachia than it does in much of the West. This finding may shed some light on the conflicting findings that have examined the association between obesity and income inequality (Chang and Christakis 2005; Diez-Roux 2000).

Access and the food retail landscape are important features of the built environment, and are associated with obesity prevalence at the county level. The percentage of households in a county that own a vehicle and are more than one mile to a grocery store had the largest effect on obesity prevalence, second only to physical inactivity. It was not corroborated that living further from a grocery store would have worse effects on obesity prevalence, however people who live more than 10 miles from a supermarket may shoulder against poor access through other means. This conflicting relationship may be further explored by comparing counties where the relationship between obesity and low access differs (e.g., selected counties in New Mexico and Utah vs. selected counties in Illinois, Indiana, and Kentucky). Furthermore, counties that witnessed a decline in the number of grocery stores and convenience stores from 2000-2008, also have higher obesity prevalence. This may suggest that diversity in the food retail landscape is imperative to diet, regardless of food store type. However, the availability and quality of healthy foods of the same food store type vary by the racial and socioeconomic composition of the neighborhood (Franco et al. 2008).
The results of this GWR based study have several implications for future research. First, the factors associated with obesity prevalence appear to operate differently across place. Prior studies using OLS and other traditional models have not explored the possibility of spatially varying relationships between obesity and other contextual variables. GWR offers a new exploratory method for uncovering heterogeneous relationships across space and could be applied to study health outcomes at a variety of spatial scales.

The second implication of this study is the need to further explore how income inequality affects obesity outcomes at the aggregate level. Inconsistent findings from other empirical studies confirm this need (Chang and Christakis 2005; Diez-Roux 2000). A conceptual model is needed to draw hypotheses about how income inequality may operate differently on obesity across place. In other words, why does income inequality serve as protection against obesity in much of the West, but exacerbates obesity outcomes in parts of Appalachia? These results suggest that relationship between income inequality and obesity may vary across counties depending on where the counties lie on the obesity distribution. Future studies should continue to examine the complicated interrelationships between income inequality, place, and health outcomes. Empirical studies, alone, may not be able to unravel these relationships. Cynthia Duncan’s place histories of three impoverished rural regions of the U.S. may serve as a useful model for understanding the relationship between community cohesion, reciprocity, and health (Duncan 1999).

This study also reveals several potential policy prescriptions to combat the epidemic. First, culturally sensitive programs that target minorities, particularly African-Americans and Native Americans, will likely be effective in counties where the GWR model shows a strong positive relationship between racial composition and obesity. Second, mobility and living at least
one mile from a food store are associated with obesity prevalence at the county level. Access to grocery stores and convenience stores are important features of the built environment, which protect against obesity. Although grocery stores and convenience stores are pertinent to the food retail landscape, policymakers and researchers alike must acknowledge that the quality of food stores is not consistent across neighborhoods.

Finally, changes in lifestyles are perhaps the most salient factor to achieve decreased obesity rates. Physical inactivity was the most predictive factor in obesity prevalence. Technological innovations, occupational shifts, and commuting patterns promote sedentary behaviors, which negatively affects obesity outcomes. Results from this study suggest that more efforts are needed to increase physical activity among Americans. Targeting this predictor alone, may pay substantial dividends in terms of health care costs and workplace productivity.

This study is not without limitations. In general, the analysis is cross-sectional, and does not contribute to the understanding of how macro-level processes are responsible for the increased prevalence of obesity overtime. There are also potential concerns about using aggregate level data to make generalizations about individuals or places. This study does not account for the possible selection of people into counties. For example, wealthy individuals, who are more likely to be physically active and able to afford to live in counties with more natural amenities (e.g., the beach), may select to live in counties that meet their preferences. Thus, the spatial patterns observed may be an artifact of the preferences of individuals. The model fit, or explained variance, varies substantially across the study area. This is evidence that additional variables may be included to improve model specification. For example, other socioeconomic variables (e.g., poverty and education) may improve the local r-square in some areas. Also, the modifiable area unit problem (MAUP) is well-known in the spatial analysis literature
(Fotheringham and Wong 1991; Jelenski and Wu 1996). It is possible that different parameters and inferences can be drawn depending on the spatial scale (e.g., block groups, census tracts, counties, and states) used in the unit of analysis, or when aggregate units are divided into different zonal areas (Openshaw and Taylor 1981). However, counties, the unit of analysis used in this study, are less arbitrary than other spatial units (e.g., census tracts and block groups). Counties are meaningful in the sense that they serve as political units, and many public health services are administered at the county level. However, the size and shape of counties vary across the United States; therefore, counties may be less relevant in some areas than in others (Goodchild 2011).

In conclusion, future studies should adopt an approach that is sensitive to the uniqueness of place. As discussed earlier, GWR is not void of limitations, but it provides a useful alternative to traditional regression methods. Contextual influences on obesity prevalence, and variations in the environmental resources that serve as health enhancing resources, is associated with the spatial stratification of health. A local, more nuanced model helps identify factors that contribute to obesity prevalence, and unravel the relationships between place and health. Thus, this study elucidates the relationship between place and obesity. Hence, the associations between most predictors observed in this study vary across the continental United States at the county level. The information gleaned from this study may be useful for grassroots policy interventions, and provides an empirical basis for the public health community to target specific predictors of obesity that will be the most effective.
References


Baker E., M. Schootman, E. Barnidge and C. Kell. 2006. The role of race and poverty in access to foods that enable individuals to adhere to dietary guidelines, *Prev Chronic Dis* 3


LaVeist, Thomas. 2005. "Disentangling race and socioeconomic status: A key to understanding health inequalities." *Journal of Urban Health* 82:26-34.


Macintyre, Sally. 2007. "Deprivation amplification revisited; or, is it always true that poorer places have poorer access to resources for healthy diets and physical activity?" *International Journal of Behavioral Nutrition and Physical Activity* 4:32.


U.S. Bureau of the Census. *County Business Patterns.*


Table 1: Description of Rural-Urban Continuum Codes, 2003

<table>
<thead>
<tr>
<th>Codes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Metro Counties:</td>
</tr>
<tr>
<td>1</td>
<td>Counties in metro areas of 1 million population or more</td>
</tr>
<tr>
<td>2</td>
<td>Counties in metro areas of 250,000 to 1 million population</td>
</tr>
<tr>
<td>3</td>
<td>Counties in metro areas of fewer than 250,000 population</td>
</tr>
<tr>
<td></td>
<td>Nonmetro counties:</td>
</tr>
<tr>
<td>4</td>
<td>Urban population of 20,000 or more, adjacent to a metro area</td>
</tr>
<tr>
<td>5</td>
<td>Urban population of 20,000 or more, adjacent to a metro area</td>
</tr>
<tr>
<td>6</td>
<td>Urban population of 2,500 to 19,999, adjacent to a metro area</td>
</tr>
<tr>
<td>7</td>
<td>Urban population of 2,500 to 19,999, not adjacent to a metro area</td>
</tr>
<tr>
<td>8</td>
<td>Completely rural or less than 2,500 urban population, adjacent to a metro area</td>
</tr>
<tr>
<td>9</td>
<td>Completely rural or less than 2,500 urban population, not adjacent to a metro area</td>
</tr>
</tbody>
</table>

Source: United State Department of Agriculture
<table>
<thead>
<tr>
<th></th>
<th>Obesity</th>
<th>Rural-Urban</th>
<th>Total Pop.</th>
<th>Black</th>
<th>Hispanic</th>
<th>Asian</th>
<th>Native</th>
<th>Inequality</th>
<th>no car, 1 mi</th>
<th>no car, 10 mi</th>
<th>Supermkt</th>
<th>Corner</th>
<th>Amenity</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obesity</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural-Urban</td>
<td>0.106</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Pop.</td>
<td>-0.201</td>
<td>-0.332</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.469</td>
<td>-0.138</td>
<td>0.076</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.284</td>
<td>-0.067</td>
<td>0.178</td>
<td>-0.107</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-0.304</td>
<td>-0.399</td>
<td>0.525</td>
<td>0.030</td>
<td>0.167</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native</td>
<td>0.124</td>
<td>0.128</td>
<td>-0.040</td>
<td>-0.099</td>
<td>-0.017</td>
<td>-0.050</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality</td>
<td>0.190</td>
<td>0.185</td>
<td>0.064</td>
<td>0.428</td>
<td>0.152</td>
<td>-0.027</td>
<td>0.046</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no car, 1 mi</td>
<td>0.487</td>
<td>0.296</td>
<td>-0.202</td>
<td>0.456</td>
<td>-0.129</td>
<td>-0.285</td>
<td>0.165</td>
<td>0.468</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no car, 10 mi</td>
<td>0.077</td>
<td>0.417</td>
<td>-0.132</td>
<td>0.011</td>
<td>0.038</td>
<td>-0.171</td>
<td>0.313</td>
<td>0.164</td>
<td>0.433</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supermkt</td>
<td>-0.110</td>
<td>-0.086</td>
<td>0.071</td>
<td>-0.086</td>
<td>0.050</td>
<td>-0.007</td>
<td>-0.096</td>
<td>-0.152</td>
<td>-0.033</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corner</td>
<td>-0.024</td>
<td>-0.098</td>
<td>0.052</td>
<td>0.033</td>
<td>0.039</td>
<td>-0.057</td>
<td>-0.010</td>
<td>-0.043</td>
<td>-0.028</td>
<td>-0.048</td>
<td>0.053</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amenity</td>
<td>-0.372</td>
<td>-0.072</td>
<td>0.148</td>
<td>-0.033</td>
<td>0.375</td>
<td>0.171</td>
<td>0.060</td>
<td>0.209</td>
<td>-0.031</td>
<td>0.000</td>
<td>-0.026</td>
<td>0.008</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>0.630</td>
<td>0.198</td>
<td>-0.179</td>
<td>0.290</td>
<td>-0.199</td>
<td>-0.318</td>
<td>0.048</td>
<td>0.316</td>
<td>0.426</td>
<td>0.110</td>
<td>-0.124</td>
<td>-0.046</td>
<td>-0.201</td>
<td>1.000</td>
</tr>
</tbody>
</table>