A SPATIAL ANALYSIS OF THE RELATIONSHIP BETWEEN OBESITY AND THE BUILT ENVIRONMENT IN SOUTHERN ILLINOIS

by

Shiloh L. Deitz

M.A. Southern Illinois University Carbondale, 2014
B.A. Whitworth University, 2009

A Thesis
Submitted in Partial Fulfillment of the Requirements for the
Master of Science

Department of Geography and Environmental Resources
in the Graduate School
Southern Illinois University Carbondale
May 2016

THESIS APPROVAL

A SPATIAL ANALYSIS OF THE RELATIONSHIP BETWEEN OBESITY AND THE BUILT ENVIRONMENT IN SOUTHERN ILLINOIS

By

Shiloh L. Deitz

A Thesis Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Masters of Science

in the field of Geography and Environmental Resources

Approved by:
Dr. Leslie Duram, Chair-Advisor
Dr. Guangxing Wang
Dr. Charles Leonard
Dr. John Jackson

Graduate School
Southern Illinois University Carbondale
March 9, 2014

AN ABSTRACT OF THE THESIS OF

Shiloh L. Deitz, for the Master of Science degree in Geography and Environmental Resources, presented on March 9, 2016, at Southern Illinois University Carbondale.

TITLE: A SPATIAL ANALYSIS OF THE RELATIONSHIP BETWEEN OBESITY AND THE BUILT ENVIRONMENT IN SOUTHERN ILLINOIS

MAJOR PROFESSOR: Dr. Leslie Duram

Scholars have established that our geographic environments – including infrastructure for walking and food availability - contribute to the current obesity epidemic in the United States. However, the relationship between food, walkability, and obesity has largely only been investigated in large urban areas. Further, many studies have not taken an in-depth look at the spatial fabric of walkability, food, and obesity. The purpose of this study was two-fold: 1) to explore reliable methods, using sociodemographic census data, for estimating obesity at the neighborhood level in one region of the U.S. made up of rural areas and small towns – southern Illinois; and 2) to investigate the ways that the food environment and walkability correlate with obesity across neighborhoods with different geographies, population densities, and socio-demographic characteristics. This study uses spatial analysis techniques and GIS, chiefly geographically weighted multivariate linear regression and cluster analysis, to estimate obesity at the census block group level. Walkability and the food environment are investigated in depth before the relationship between obesity and the built environment is analyzed using GIS and spatial analysis. The study finds that the influence of various food and walkability measures on obesity is spatially varied and significantly mediated by socio-demographic factors. The study concludes that the relationship between obesity and the built environment can be studied quantitatively in study areas of any size or population density but an open-minded approach toward measures must be taken and geographic variation cannot be ignored. This

work is timely and important because of the dearth of small area obesity data, as well an absence of research on obesogenic physical environments outside of large urban areas.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Duram, for her guidance throughout this project. I would also like to thank my committee – Drs. Charlie Leonard, Guangxing Wang, and John Jackson - for their careful readings and suggestions. I also could not have completed this project without the support of the Paul Simon Public Policy Institute and the inclusion of built environment questions on their Southern Illinois and Simon polls. Lastly, I would like to thank my husband Thomas for providing technical assistance, advice, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Over one-third of the adult population in the U.S. is obese (34.9%) which has been estimated to cost the country $147 billion a year in medical expenses (CDC 2014). Obesity is fast becoming one of the largest preventable health issues in the United States – creating myriad other health complications including increased risk for heart disease, stroke, type 2 diabetes, and certain cancers. The most common cause of obesity is energy imbalance, that is, calories ingested exceed calories expended in exercise (Hill, Wyatt, and Peters 2013).

The inactive lifestyles of many Americans can be in large part attributed to the environments that we live in (NHLBI 2012). Infrastructure and design, particularly those characteristics of the built environment[1] that encourage car use and discourage walking or biking may contribute to the obesity epidemic. In rural areas and small towns rates of exercise are particularly low[2], and rates of obesity are particularly high – 39.6% of rural residents in the U.S. are obese compared to 33.4% of urban residents (Befort et al. 2012). Regionally, obesity is a particularly serious problem in the South and Midwest (see figure 1).

[1] The 'built environment' refers to the settings for human activities that are human-made, such as parks, buildings, neighborhoods, and downtown areas.

[2] In both the Midwest and Southern United States *inactivity* was highest in non-metropolitan counties (37.7% in the Midwest and 44.9% in the South) (Meit et al. 2014).

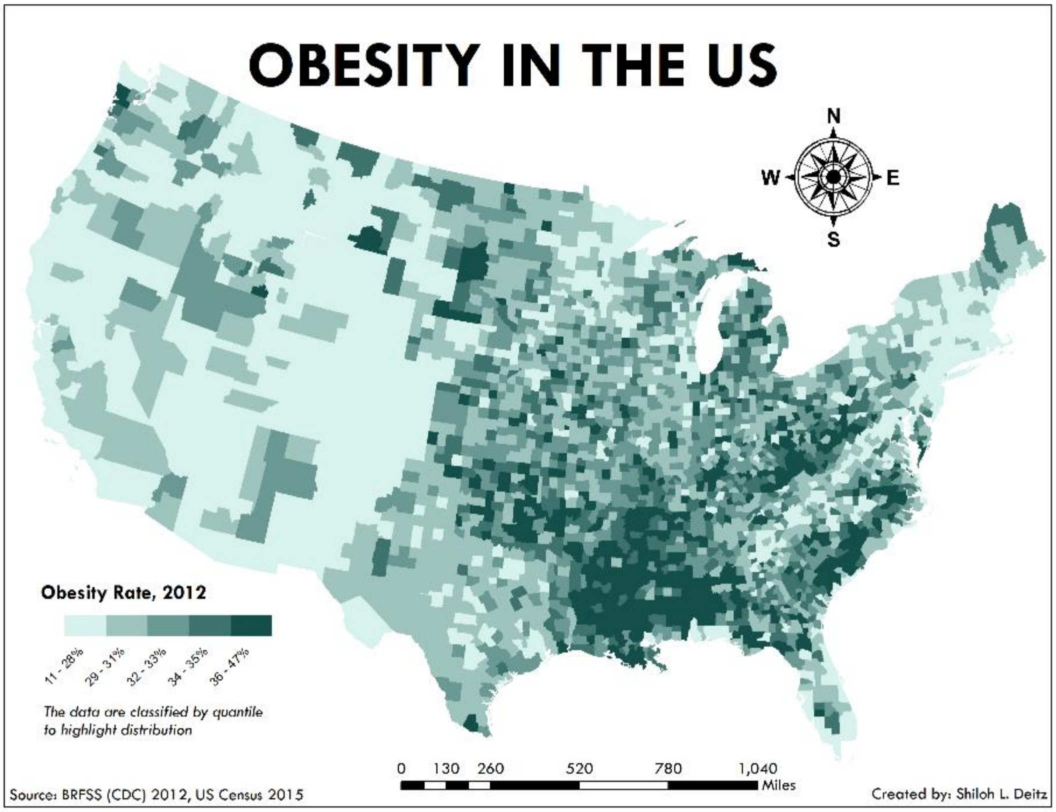

FIGURE 1. OBESITY IN THE UNITED STATES

The built environment in small towns often encourages car use, provides restricted food shopping options, and is plagued by a low ratio of healthy food to unhealthy food outlets (Hartley 2004; Meit et al. 2014). If environmental factors contribute to obesity rates, it is quite possible that the design of small towns and cities in otherwise rural areas contributes to higher rates of obesity.

Non-metropolitan and rural areas have largely been overlooked in walkability, food environment, and healthy community studies. While it is well-established that local geographic or environmental factors – such as region, terrain, walkability, and culture - contribute to the obesity epidemic, there is a dearth of small-area obesity data in rural and small town areas making it nearly impossible to quantitatively explore this connection in those communities (Swinburn et al. 1999).

Broadly, this study aims to contribute to understanding the relationship between obesity and the built environment in southern Illinois, a region made up of rural areas and small towns. The purpose of this study is to explore reliable methods for estimating obesity at the neighborhood level in one region of the U.S., to explore walkability and food access in southern Illinois, and to investigate the ways that the food environment and walkability correlate with obesity across different geographic settings. This study is timely and important foremost due to the seriousness of the obesity epidemic, but also because of the dearth of small area obesity data (as well as the lack of clear methodologies for interpolating obesity data), and the absence of research on obesogenic environments outside of large urban areas. This research could be used as a methodological guide for obesity and healthy community studies, or to propose healthy design policy for non-metropolitan areas in the U.S.

In the following pages, I will provide an in-depth review of previous scholarship on obesity estimation (including the relationship between obesity and sociodemographic factors), obesogenic environments, the relationship between walkability and obesity, and the relationship between food environments and obesity. I will then outline the purpose of this study, the research questions that guide it, and the methods used to answer the research questions. I will conclude by commenting on the methodological and theoretical importance of this study.

II. LITERATURE REVIEW

This study is based on the presupposition that there are particular environmental factors that contribute to obesity rates. In academic research, environments that promote excessive unhealthy intakes and discourage physical activity are called obesogenic (Hill and Peters 1998). In 1999, Swinburn and his colleagues described the obesogenic environment as the “sum of influences that the surroundings, opportunities, or conditions of life have on promoting obesity in individuals or populations” (564). Political, sociocultural, economic, and physical environments can be linked to obesity (Swinburn et al. 1999, see figure 2). In this project, the estimation of obesity draws on the *sociocultural* and *economic* factors that contribute to obesity. For the rest of the analyses the focus is on walkability, density of healthy food to unhealthy food, and distance to food as factors that may contribute to an obesogenic *physical* environment.

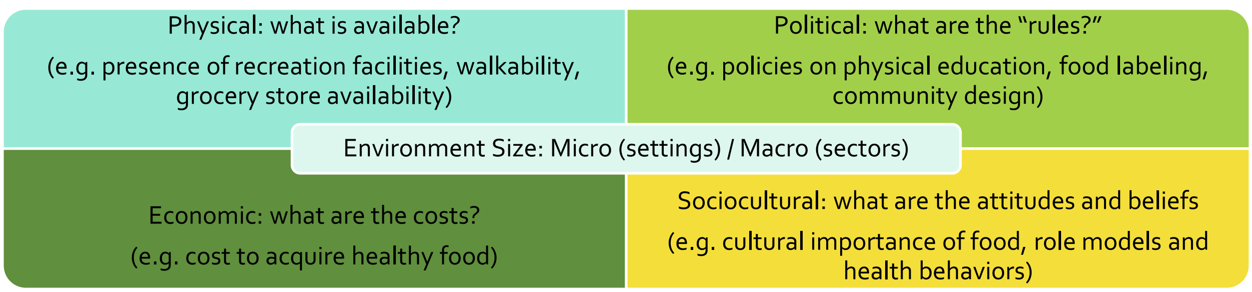

FIGURE 2. ANALYSIS GRID FOR ENVIRONMENTS LINKED TO OBESITY

A. OBESITY ESTIMATION

This study uses sociodemographic factors to estimate obesity using regression models. The methods used in the analysis build upon pre-existing literature on both sociodemographic trends in relation to obesity and small area estimation techniques for health measures.

1. Sociodemographic Factors and Obesity

Past research has established links between various social or demographic characteristics and obesity including: marital status, income, educational attainment, characteristics of work commute, employment status, occupation, race, age, sex, housing characteristics, and family type (Boone et al. 2013; Wang and Beydoun 2007).

Overall, studies have shown that married individuals tend to have better physical and mental health than non-married people (Manzoli et al. 2007; Pienta et al. 2000). Consistently, studies have documented a positive relationship between obesity and age, a negative relationship between educational attainment and obesity, a negative relationship between income and obesity, a positive relationship between commuting to work by car and obesity, and a higher prevalence of obesity among African Americans (Frank et al. 2004). In their nation-wide analysis of the National Health and Nutrition Examination Survey (NHANES) data Wang and Beydoun (2007) found that: overweight prevalence increased with age, more men than women were overweight or obese, non-Hispanic Blacks were more likely to be obese than other racial or ethnic groups, and the relationship between obesity and socio-economic status (SES[3]) varied across racial and ethnic groups. Other research has found a strong inverse relationship between socio-economic status and obesity for women, but a positive association between the two among men (Zhang and Wang 2004). This research also found that after controlling for SES minorities are not more likely to be obese than whites.

Finally, associations between the type of neighborhood a person lives in and obesity have been consistently noted. Neighborhood characteristics can be conceptualized using characteristics that include both housing and households. Obesity has been observed to be positively correlated with a high prevalence of single parent households (Huffman et al. 2010). Multiple studies have found a significant negative relationship between high walkability and the prevalence of obesity (Frank et al. 2005; Boone-Heinonen et al. 2013; Booth et al. 2005; Glazier

[3] Socio-economic status can include educational attainment, income, occupation, and employment status.

et al. 2014; Mackenbach et al. 2014, Sallis et al. 2009; Wang et al. 2013; Zhang et al. 2014). Research has also linked the year houses were built and tenure of residents to a neighborhood's walkability, and thus obesity (Zick et al. 2009). Specifically, Zick and her team found a significant negative relationship between the year houses were built and obesity.

With so many potential sociodemographic factors correlating with obesity – these studies beg the question: how could a researcher go about reducing factors to create a parsimonious regression model for obesity estimation? In addition, these studies largely ignore geography – a factor that may in large part explain the variation in sociodemographic relationships or seemingly contradictory conclusions drawn from other studies on sociodemographics and obesity – leading to the question, how do the relationships between sociodemographic factors and obesity vary geographically? And what, if any, trends can be observed regarding geography and sociodemographics?

2. Small-Area Estimation of Obesity

Numerous studies have contributed to addressing the problem of reducing factors to create a parsimonious and reliable regression model for small area estimates. Public health data is seldom collected at the local level and small-area estimation of public health indicators is common practice. Scholars have created small-area obesity estimates using sociodemographic and less often geographic relationships (Adu-Prah and Oyana 2015; Bell 2014; Boone-Heinonen et al. 2013; Cataife 2014; Lee et al. 2014; Li et al. 2009; Malec et al. 1999; Merchant et al. 2011).

Regression equations have been widely used to estimate obesity. Guido Cataife used regression models with census data to estimate obesity at the neighborhood level in Rio de Janeiro, Brazil (2014). Boone-Heinonen and his colleagues (2013) used sociodemographic and food environment data to estimate the BMI change that could be expected from neighborhood changes – informing policy interventions. Many of these studies have determined regression models for estimating obesity through a theoretical and/or automatic approach. One study conducted in Massachusetts, drew potential variables of interest for estimating obesity from the literature and then selected those most significant for a given study area using backward elimination regression (F-for-removal greater than 0.1) (Li et al. 2009). This method was useful for estimating tobacco use as well (Li et al. 2009a). However, research has also shown that models derived from backward elimination methods are not always the best because of interactions between the variables (Braun and Oswald 2011). Braun and Oswald (2011) suggest that running all possible regression subsets to find the best model might be a better technique for finding the set of variables with the greatest predictive power. This method allows for all possible relationships to be explored and is useful when independent variables are exhibit collinearity.

Other studies have added a spatial component to regression or relied solely on spatial relationships. Adu-Prah and Oyana (2015) used spatial interpolation and regression models to estimate obesity. Another team used national datasets with variables measuring both environmental and social factors to estimate obesity using spatial interpolation techniques (Merchant et al. 2011).

In an analysis of the effectiveness of multiple estimation techniques commonly employed with disease data, Goovaerts (2006) found that spatial interpolation methods alone where less valid than those that also incorporated regression equations. In later work he used geographic regression to estimate lung cancer mortality rates (Goovaerts 2010). In a general analysis of various prediction methods, Gao, Asami, and Chung (2006) evaluated the predictive power of a simple linear regression model, a spatial dependency model, a combined spatial dependency and geographically weighted regression (GWR) model and a simple GWR model. Using numerical cross-validation, they found that the simple GWR model resulted in the most reliable predictors. Zhang, Gove, and Heath (2005) compare six modeling techniques (ordinary least squares, linear mixed model, generalized additive model, multi-layer perceptron neural network, radial basis function neural network, and geographically weighted regression). They also found that the geographically weighted regression model was the best for prediction.

When estimating a phenomenon, the validity of those estimates is of paramount concern. Common techniques for accuracy analysis are analysis of the residuals and re-aggregation of the data back to larger areas where more is known about the variable of interest (e.g. obesity). If a small area estimate does not closely match estimates or measures for a larger geographic area when aggregated, red flags should be raised regarding the reliability and accuracy of the estimate. Large deviations might suggest model failure (Bell et al. 2013). For this reason, some researchers have narrowed down models using regression equations and theory but they made final model decisions based on aggregated error.

In a study of the spatial variation of housing attribute prices, Bitter and his team (2007) re-aggregated their small-area estimates back to the larger geographic areas that the estimates were based upon to check the accuracy of their models. This method has also been used by Pfefferman and Barnard (1991) in a farmland value study, Wang, Fuller, and Qu (2008), Datta, Ghosh, Steorts, and Maples (2011), Zhang, Gove, and Heath (2005), and Pfefferman and Tiller (2006). When data points are not available at smaller aggregations this might be the best method for accuracy assessment.

The field of small area estimation has made headway in many areas. However, consensus has yet to be reached on methods for developing a parsimonious model, integration of a spatial component to regression models, and methods for assessing the reliability of estimates. In this research, I explore the application of a combination of these approaches to find the best prediction model.

B. WALKABILITY

The second stage of this project involved investigating the physical environment in southern Illinois and analyzing the relationship between the physical environment and obesity. One aspect of the physical environment that may correlate with obesity rates and health is walkability. A walkable neighborhood has been broadly defined as one which combines population density, pedestrian-friendly design, and diversity of destinations (Cervero and Kockelman 1997). These factors have been found to correlate with actual walking behavior the most.

Walkability is associated with higher rates of physical activity (Berke et al. 2007; Frank et al. 2005; Freeman et al. 2012; Humpel, Owen, and Leslie 2002; Sallis et al. 2009). Correspondingly, scholars have found that walkability is associated with lower rates of obesity (Boone-Heinonen et al. 2013; Booth et al. 2005; Frank et al. 2004; Mackenbach et al. 2014, Sallis et al. 2009; Wang et al. 2013; Zhang et al. 2014). Predictably, car dependence is associated with higher rates of obesity (Glazier et al. 2014; Hinde and Dixon 2005). Research has also found that these associations are stronger for people with high socioeconomic status, men, and whites (Casagrande et al. 2011; Frank et al. 2008, Humpel et al. 2004, Suminiski et al. 2005; Wang et al. 2013).

A number of studies have found a relationship between commuting behavior and obesity. Frank and his colleagues (2004) found that each additional hour spent in the car per day was associated with a 6% increase in the likelihood of obesity. This finding was supported by later work (Frank et al. 2008; Glazier et al. 2014; Zhang et al. 2014).

Fewer studies have examined the relationship between walkability and obesity in non-metropolitan areas. Instead of focusing on infrastructure, some studies have compared the physical activity levels of children in urban and rural settings – or actual behavior (Sandercock et al. 2010). These studies have had divergent results – some found that there is no difference in the physical activity levels of urban and rural children (McMurray et al. 1999; Felton et al. 2002; Springer et al. 2009); while others have found that rural children are more active (Joens-Matre et al. 2008; Liu et al. 2008); and still others have found that suburban children are more active (Springer et al. 2006; Nelson et al. 2006).

One national quantitative study of adults looked at the relationship between rural-urban location, walkability, and obesity at the county and state level. They measured walkability by street connectivity and found that the influence of individual level variables in the models (e.g. race, class, gender) varied across urban, rural, and suburban areas (Wang et al. 2013). Overall, this study found that the relationship between street connectivity and physical activity was weaker than the relationship between street connectivity and obesity. This study was focused on finding a good measure of walkability for counties across the entire United States.
At the county and state level of aggregation the findings do not capture local diversity, but rather, provide broader generalizations or trends.

In another study, Zhang and his colleagues (2014) examined the relationship between commuting behavior and obesity in rural and urban areas. They found that automobile dependence was correlated with obesity in urban areas but not rural ones. They also found that longer commuting times were associated with obesity in areas of any size. Suggesting that using a car may be less significant than the influence on lifestyle of spending long hours in that car, particularly in car-dependent rural areas. Numerous studies have looked at the influence of long car commute on quality of life and health. More specifically, these studies have often found that time spent commuting leads to lower levels of life satisfaction, time pressure, and reduced time for physical activity and leisure (Hilbrecht et al. 2014). These factors all suggest reduced well-being and higher prevalence of obesity.

Other studies of walkability and obesity in rural areas have used qualitative methods and examined a wider variety of factors such as perceived crime, loitering behavior, trash, and the presence of gangs (Hennessy et al. 2010)[4]. Hennessy and her team found that factors unmeasurable quantitatively, influence behavior and perceptions of walkability despite infrastructure.

Overall, these studies suggest that there are differences between rural and urban areas but these differences have yet to be fully explored and methods for assessing rural or small town walkability and its relationship to obesity need to be refined. The literature raises the questions: how can walkability be quantified in non-urban areas? How walkable are small towns? What is the association between walkability and obesity in small towns? And finally, what interventions regarding walkability would be feasible in small towns?

---

[4] While these are interesting topics to analyze, the purpose of this study is to examine rural and small town areas using the same analysis techniques established for urban areas.

C. Food Environment

The food environment is another factor contributing to obesogenic physical environments. By food environment, I am referring to the environmental factors that influence food choices and diet quality (ERS 2014). For example, distance to food sources, density of healthy food to unhealthy food stores, and cost of food (CDC 2014a; 2014b). The food desert approach and the healthy food density approach are two commonly used modes of conceptualizing the food environment. A food desert, is an area where distance to an affordable and healthy food source is great (CDC 2014a). This mode of conceptualization captures issues of access to healthy food. The second approach to assess the food environment – the healthy food density approach or modified food retail environment – consists of simply calculating the ratio of healthy food retailers to all food retailers within any chosen area (CDC 2014b). This method aims to capture the negative health effects of living in a place with a high density of unhealthy (and often inexpensive) food options.

1. Food Deserts

Findings have been inconsistent on the relationship between food access and obesity. Some have found that there is no relationship between food access and obesity (Alviola, Nayga, and Thomsen 2013; Budzynski et al. 2013; Caspi et al. 2012), while others have found that living in a food desert predicts obesity (Chen et al. 2010; Edwards et al. 2009; Giskes et al. 2011; Hilmers, Hilmers, and Dave 2012; Inagami et al. 2006; Morland et al. 2002; Schafft et al. 2009). The inconsistency of results may be in part due to differences in conceptualization and methodology. Food access does not lend itself easily to quantitative research. The influence of access on health is mediated by social and geographic factors and thus we cannot expect to find one rule that applies to all areas.

While few studies have been conducted in rural areas, the ones conducted have generally found that rural residents have lower food access (Larson, Story, and Nelson 2009) and that some rural residents rely on non-traditional methods for acquiring healthy food such as growing food, sharing, hunting, and buying in bulk (McPhail et al. 2013; Morton and Blanchard 2007; Scarpello et al. 2009; Sharkey et al. 2010; Yousefian et al. 2011). Depending on where they are geographically and also depending on social factors, rural residents interact with their environment in unique ways suggesting that studies should focus on local areas or pay special attention to geographic variation. Further, qualitative work should be approached as a compliment to quantitative anaylsis – perhaps serving to explain quantitative finding or anomalies.

2. Healthy Food Density

Results have been fairly consistent on the relationship between supermarket or healthy food source density and obesity – an increase in supermarket density and/or a decrease in convenience store density predicts a decrease in obesity rates (Boone-Heinonen et al. 2013; Casey et al. 2014; Frank et al. 2012; Giskes et al. 2011; Hutchinson et al. 2012; Morland and Evenson 2009; Morland et al. 2006; Powell et al. 2007; Powell and Bao 2009). Specifically, Morland, Diez Roux, and Wing (2006) found that obesity and overweight prevalence was lowest in areas with supermarkets only, then areas with a combination of supermarkets and groceries, and was highest in areas with a combination of grocery stores and convenience stores but no supermarkets. This research suggests that food environments are complicated. Demographically, minorities and the poor are more likely to live in a food desert and are more likely to live in an area with higher convenience store and fast food density (Alviola, Nayga, Thomsen, and Wang 2013; Bellinger and Wang 2011; Choi and Suzuki 2013; Powell et al. 2007; Sohi et al. 2014). All of these studies were conducted in urban areas, we know little about the impact of healthy food density in non-urban areas.

3. Food Access and Availability

While many studies have looked at just the density of healthy food or the presence of food deserts and the impact of those factors on obesity and other diet related health outcomes, there are studies that have analyzed the overall food environment. Many of the studies have had the same findings as studies which have only analyzed one factor or the other – that is, living in a food desert, in an area with a high ratio of unhealthy food to healthy food stores is associated with obesity (Morland et al. 2006). While generally it has been found that as distance to food increases and density of healthy food decreases obesity goes up, this is dependent on multiple factors, particularly conceptualization, sociodemographics, and geography.

Data on obesity rates is difficult to acquire, and for that reason many assessments of food environments have looked at sociodemographic disparities – particularly disparities according to race and class. Pedro Alviola and his colleagues (2013) modeled sociodemographic neighborhood characteristics and the food environment. They found that in urban low income blocks in Arkansas with higher minority populations, residents faced a higher density of convenience and fast food outlets compared to higher-income urban blocks (Alviola, Nayga, Thomsen, and Wang 2013). Rural communities with declining populations were found to be at risk for lower access to healthy food in their study. Sharkey and Horel (2008) found that neighborhoods with the greatest socioeconomic and racial disparity had greater spatial access to supermarkets, grocery stores, convenience stores and discount stores. They also found that rural residents have low access to food sources. Finally, a national study that included urban and rural classifications as a control variable found that low income and black neighborhoods had significantly fewer supermarkets (Powell et al. 2007). Aside from health, these disparities point to a certain social injustice in the area of access to food for healthy living.

The public health and policy communities were quick to accept the food desert concept and since the 1990s (when the idea was introduced) little has been done to truly understand the difference that various distances make. Further, non-urban areas have been largely overlooked by these analyses. The current state of the literature on food access leaves room for investigation into the relationship between obesity and food environments as well as a more robust understanding of these relationships in non-metropolitan neighborhoods. These investigations would be incomplete without an in-depth understand of both geographic and social factors.

D. WALKABILITY AND FOOD ENVIRONMENT

A few studies have used the obesogenic environment hypothesis to analyze the impact of both walkability and food environments on obesity. Many studies have concluded that interventions are best targeted at the neighborhood level (Booth et al. 2005), suggesting that studies should be conducted at this level as well. A study conducted at the zip code level by Russ Lopez (2007) found that access to healthy food and walkability are negatively associated with obesity in a metropolitan area. The same results were found in a study in Utah (Zick et al. 2009). However, while the same negative correlation was observed between obesity and healthy food density or walkability by Rundle and his colleagues (2009) in their analysis of the impact of the food environment and walkability in New York City, they found no relationship between unhealthy food density and obesity.

Other studies have looked at trends among children rather than adults. A study that looked at childhood obesity found that walkability and access to healthy food had a negative association with obesity (Rahman et al. 2011). A study in Washington State that looked at child and parent obesity came to the same conclusions (Saelens et al. 2012). Another study used qualitative methods to understand the effect of the built, social, and natural environments on obesity (Maley et al. 2010). They found that residents of the rural community they studied perceived that their rural community promoted obesity - specifically due to car dependence and the social value of ‘comfort food’ (i.e. foods high in calories, fat, or sugar). This qualitative study emphasized residents’ perception of their environment and suggests something about how the physical and social environments of a region might interact to produce obesity. Again, this literature has largely focused on metropolitan areas and suffered a lack of consistent assessment tools.

The current state of the literature on obesity and built environment leaves room for further investigation of obesogenic environments in non-urban areas.

E. **RESEARCH STATEMENT**

The purpose of this study was two-fold: 1) to explore reliable methods, using sociodemographic census data, for estimating obesity at the neighborhood level in one region of the U.S. made up of rural areas and small towns – southern Illinois; and 2) to investigate the ways that the food environment and walkability correlate with obesity across neighborhoods with different geographies, population densities, and socio-demographic characteristics. This study uses spatial analysis techniques and GIS, namely geographically weighted multivariate linear regression and cluster analysis, to estimate obesity at the census block group level. Walkability and the food environment are investigated in depth before the relationship between obesity and the built environment is analyzed using GIS and spatial analysis. This work is timely and important because of the dearth of small area obesity data, as well an absence of research on obesogenic physical environments outside of large urban areas.

F. Research Questions

The research is guided by the following questions:

1. How can small area obesity estimates be reliably interpolated from data sets at coarse spatial resolutions?

2. How does obesity vary geographically in southern Illinois at the block group level?

3. How can walkability be quantitatively measured in non-urban areas and how walkable is southern Illinois?

4. How can the food environment be quantified in non-urban areas and what is the food environment in southern Illinois?

5. What is the correlation between the built environment and obesity in southern Illinois, after controlling for socio-demographic covariates?

III. METHODS

The research questions were answered with a cross-sectional quantitative research design. The data came from the adult population in the 18 southernmost counties of Illinois. The first stage of research involved estimating and understanding obesity in the region, the second stage involved quantifying the food and walkability environment, and the final stage focused on the quantitative relationships between obesity, the built environment, and sociodemographic factors. In the following pages I will describe the study area, the variables used in these analyses and their conceptualization, and finally the data analysis procedures.

A. STUDY AREA

The target study area for this research was the 18 southernmost counties of Illinois. The counties included in the study are: Alexander, Franklin, Gallatin, Hamilton, Hardin, Jackson, Jefferson, Johnson, Massac, Perry, Pope, Pulaski, Randolph, Saline, Union, Washington, White, and Williamson. In these counties, analysis was undertaken at the block group level because of the research suggesting that interventions and analysis of obesity and the built environment are best conducted at the neighborhood level (Booth et al. 2005). These are the counties that the Paul Simon Public Policy Institute uses for their Southern Illinois Poll. For obesity estimation the 600 counties of Illinois and its neighboring states (Indiana, Iowa, Kentucky, Missouri, and Wisconsin) were used (see figure 3).

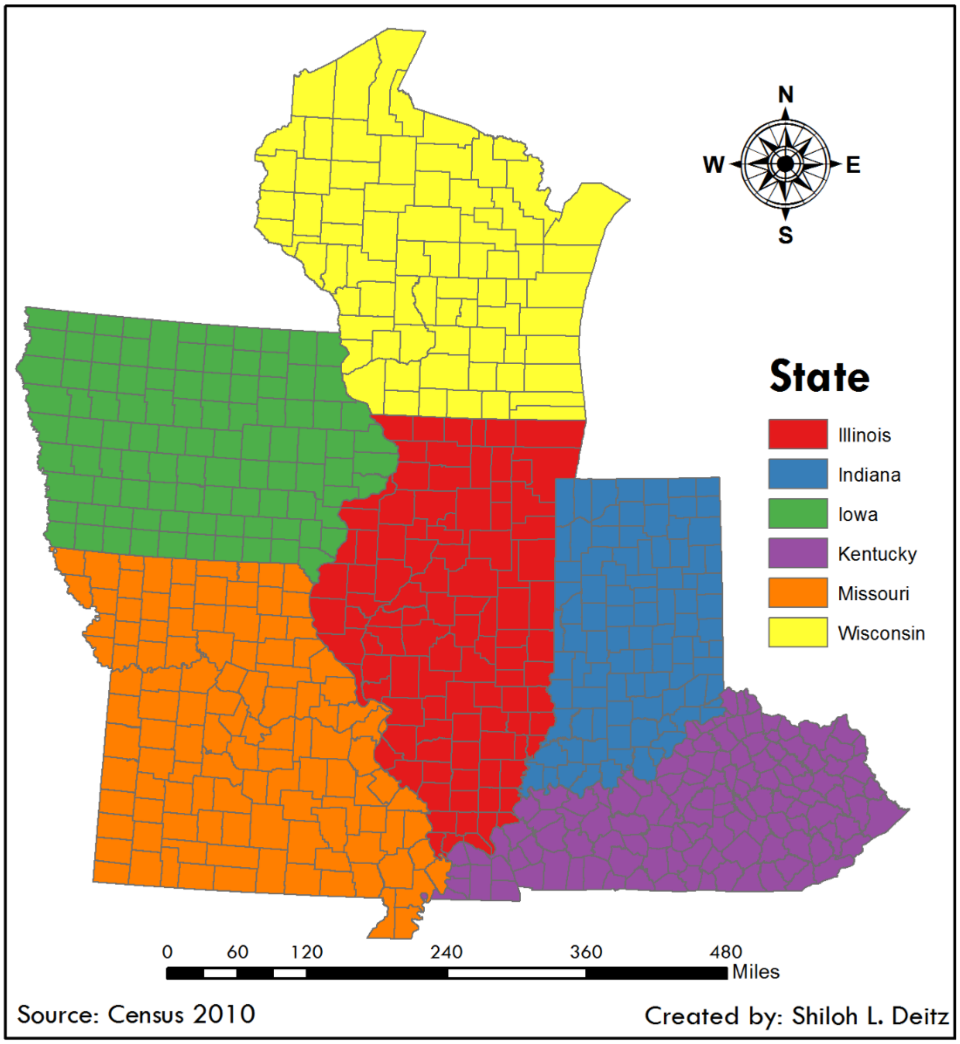

FIGURE 3. ILLINOIS, INDIANA, IOWA, KENTUCKY, MISSOURI, AND WISCONSIN

The average income in the southern Illinois study area is \$21,829 (s_x=6984.95), and on average 18.5% (s_x=14.36) of the population is in poverty. About one in five have a college degree or more on average (18.19%; s_x=12.79). The area is mostly white (90.07% on average; s_x=15.51), and about half women (50.13%; s_x=7.51). The average median age across all block groups is 41.31 (s_x=7.66).

The first stages of research comprised of looking at the geographic distribution of obesity in southern Illinois, and understanding the food environment and walkability in the region. The final stage of research involved looking at the relationship between obesity and the built environment both for the entire study region and block groups of the region that are classified as urban, urban clusters, and rural[5] (see figure 4). The only urban areas in the region are Cape Girardeau and the Carbondale-Marion metropolitan area which spans from Carbondale to Marion. While no individual city in the Carbondale-Marion urban area has a population large enough to classify it as urban, the contiguous nature of the cities leads to the urban classification.

[5] According to the U.S. Census Bureau urban clusters are any areas where there is a contiguous population settlement of at least 2,500 people and less than 50,000 (i.e. at least 2,500 people live in one area without jumping (uninhabited area) to the next settlement) (Groves 2011). These areas are identified through census tract and census block population density and other land cover characteristic. First, census tracts with a land area less than three miles and at least 1000 persons per square mile (ppsm) are identified and joined with contiguous tracts also meeting the criteria. Next, tracts that are contiguous to the tracts identified in the first step and that have at least 500 ppsm and a land cover of less than three miles are identified and joined with other tracts meeting the criteria. Next, contiguous census blocks with at least 1000 ppsm are identified and joined. The remaining census blocks are identified until no more meet the criteria if: they have a population of at least 500 ppsm, or at least one-third of the block has territory with imperviousness of at least 20% and is sufficiently compact, or at least one-third of the block has territory with imperviousness of at least 20% and at least 40% of its boundary is contiguous with an already identified urban boundary (Groves 2011). A rural area would not meet that classification due to having less population, while an urban area has more.

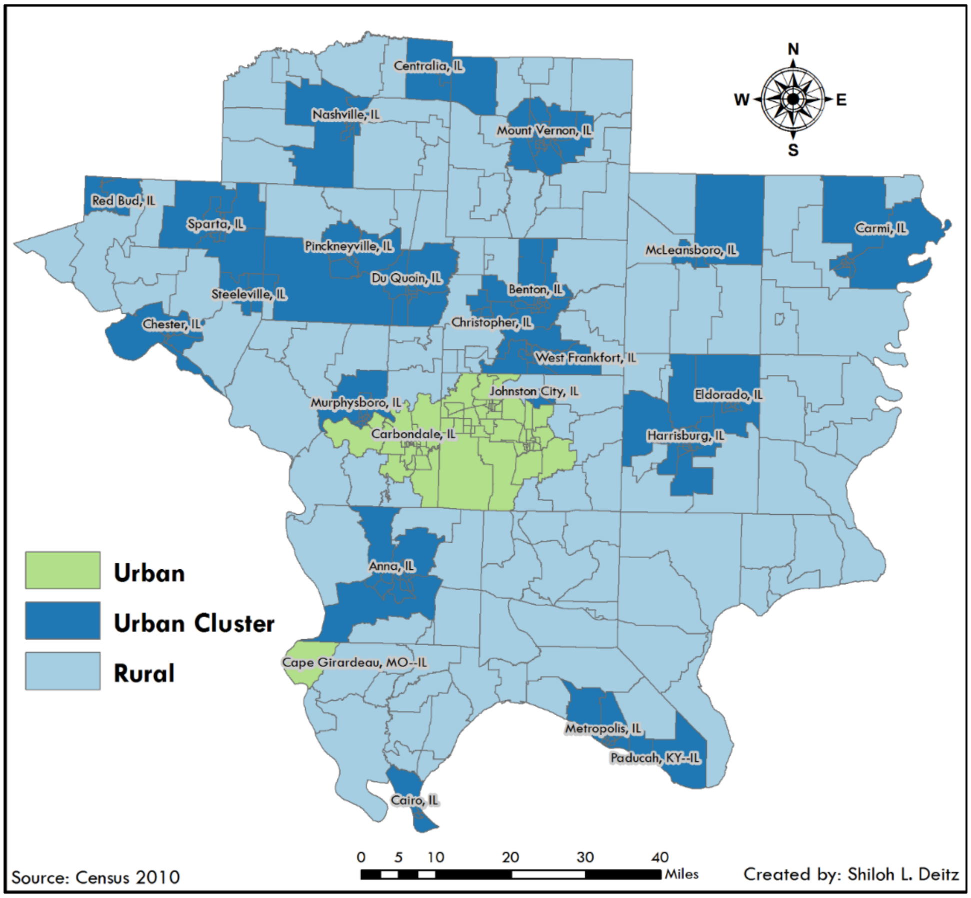

FIGURE 4. SOUTHERN ILLINOIS URBAN, URBAN CLUSTER, AND RURAL AREAS

B. DEFINITION OF TERMS, VARIABLE CONCEPTUALIZATION, AND DATA

The key variables in this study are obesity, walkability, food access, healthy food density, race, class, age, educational attainment, and gender. Geographic aggregations are at the county and census block group level. All spatial data were projected to Universal Transverse Mercator North American Datum 1983 zone 16 north. In the following section I will define key terms, describe how they were conceptualized for quantitative analysis, and describe the data sources.

1. Obesity

Body mass index was used to measure obesity in this study. Body mass index (BMI) is a number calculated using a person's height and weight. It is thought to be a fairly reliable indicator of body fatness for most people and is both easy and inexpensive to calculate. It is also the only estimate of body fat that can be taken over the phone (CDC 2014c). While there has been some debate about whether BMI is actually a good measure of obesity, studies have shown that overall it is a reasonably accurate measure – with the benefits of easy and inexpensive collection outweighing potential inaccuracies (Baile and Gonzalez-Calderon 2014; Dietz and Bellizzi 1999).

The health community typically defines obesity using BMI ranges. A person is considered overweight if their BMI is between 25 and 29.9 and obese if their body mass index is 30 or higher (CDC 2012). These thresholds were determined due to their connection with obesity-associated morbidity.

In this study, the percentage obese within a geographic unit is used – meaning the percentage of persons with a BMI over 30. County level obesity count and percentage data comes from the 2012 Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is a yearly telephone survey that collects data on health-related risk factors (BRFSS 2014). Over 400,000 telephone interviews are conducted each year in all 50 states and it is the largest continuous health survey in the nation. County obesity data was gathered from all 600 counties in Illinois and its neighboring states (Indiana, Iowa, Kentucky, Missouri, and Wisconsin) and used to estimate obesity at the block group level using socio-demographic covariates.

2. Walkability

As mentioned previously, a walkable neighborhood is one which combines population density, pedestrian-friendly design, and diversity of destinations (Cervero and Kockelman 1997). Research has suggested that commuting behavior is highly correlated with characteristics of the built environment (Wang and Chen 2015). Further, distance to food is suggestive of the overall walkability of an area – or diversity of destinations. Distance to food and number of food stores within an area are also highly correlated with *WalkScore*[6] data in the study area (see table 3). *WalkScore* data has been found to be highly correlated with street connectivity, residential density, access to public transportation, and access to walkable amenities (Carr et al. 2010; Carr et al. 2011).

For this study walkability is conceptualized in terms of commuting behavior (mode and travel time), distance to food, and number of food stores within an area. WalkScore data for the urbanized areas of southern Illinois was used to assess good walkability measures. Commuting behavior data came from the American Community Survey 2013 five year estimates (United States Census Bureau 2013). Specifically, the percentage that walk, bike or take public transportation to work and the average work commute time were used. The percentage that used other modes of travel (e.g. car, motorcycle) were also available in the data but a significantly different relationship between those that walk, bike, or take public transportation and the rest of the sample was observed so only that variable was retained. The source and process of cleaning the food data is described below. To provide guidance for conceptualizing walkability in similar study areas, this research explores multiple measures of walkability.

[6] *WalkScores* are based on walking proximity along multiple routes to 13 amenities (grocery stores, coffee shops, restaurants, bars, movie theatres, schools, parks, libraries, book stores, fitness centers, drug stores, hardware stores, and clothing/music stores). Amenities within a 5 minute walk (0.25 miles) are awarded maximum points and a decay function is used to give points at distances beyond that. Any amenity a 30 minute walk or more away is awarded zero points (*WalkScore* 2014).

3. Food

i. Access

Food access is most often conceptualized with food desert measures. A food desert is a residential area where the distance to affordable and healthy food is great, decreasing residents' ability to have a healthy diet including fresh fruits and vegetables (USDA 2009). There is lack of consensus on what a "great" distance to food is. The CDC calculates the population weighted mean center of any given geographic aggregation (e.g. tract or block) and uses that point to calculate the distance to the nearest grocery carrying a wide variety of foods including fruits and vegetables (CDC 2014a). They define a food desert as any urban area where the distance to the nearest grocery is more than one mile and any rural area where the distance is more than ten miles. This measure of ten miles is commonly used in rural areas because the average distance that an American travels for food is eight miles (McEntee and Agyeman 2010). This measurement cut off is arbitrary and has not been investigated in depth. According to this logic, a person could be living in a food desert because they are 10.1 miles from the grocery store, while their neighbor is not considered at risk for living in a food desert because they are 9.9 miles from the grocery.

Due to the arbitrary nature of the measures mentioned above, the distance to the nearest healthy food, unhealthy food, and any food (healthy or unhealthy) was calculated from the population weighted centroid along a road network. The number of healthy food stores, unhealthy food stores, and stores of any kind were also calculated at service area buffers of 800 meters, 1600 meters, 3200 meters, 8 kilometers, and 16 kilometers. This variety of measures aided in better defining the relationship between distance and health outcomes in the study region.

ii. Healthy Food Density

The modified retail food environment index is a measure created by the CDC for measuring local food environments. It represents the percentage of all food stores (including grocery stores, convenience stores, and fast food stores) that are healthy food stores (grocery stores or supermarkets). It can be calculated at any geographic aggregation with the following formula:

$$mRFEI = 100 \times \frac{\#\ Healthy\ Food\ Retailers}{\#\ Healthy\ Food\ Retailers + \#\ Less\ Healthy\ Food\ Retailers}$$

Healthy food retailers include grocery stores, supercenters and produce stores. Less healthy food retailers include convenience stores, and fast food restaurants. These classifications are based on the *typical* foods offered in such stores (CDC 2014b; Frank et al. 2012). In this study, the MRFEI was calculated at a road network service area of 800 meters, 1600 meters, 3200 meters, 8 kilometers, and 16 kilometers.

iii. Data

Data for measuring the food environment came from a database of retailers that accept SNAP, data from the market research company InfoUSA, local directories, and field work (ERS 2015, InfoUSA 2015). Road network data came from TIGER/Line® shapefiles (United States Census Bureau 2013). When food point data did not coincide with the road network a line was drawn from the point to the nearest road to allow for network analysis and better capture distance – essentially, a missing road was added to connect food locations to the network.

Food location data from InfoUSA and SNAP were merged together and checked line by line. Particularly, entries that were not in both datasets were analyzed. In total 192 stores were in the SNAP data but not InfoUSA. Conversely, 24 convenience stores and 21 grocery stores were in the InfoUSA data but not SNAP. Missing data in the SNAP database could be attributed to those stores not accepting SNAP benefits. InfoUSA and other marketing data sources have been known to be incomplete (Liese et al. 2013), for this reason field work and local directories were employed.

Food locations were then classified as unhealthy or healthy. It is generally agreed upon that a healthy store should carry fresh fruits and vegetables, bread, eggs, and dairy products – a store that does not carry fresh fruits and vegetables should be classified as a convenience store. Fast food restaurants included both national chains and local businesses. Research has shown that the draw of fast food is the low price and restaurants that do not have meal options for under $5 should not be considered fast food (McDermott and Stephens 2010). Mcdermott and Stephens found that financial limitations for low-income populations can overpower adherence to recommended dietary guidelines, so when the price of fast food is not significantly low the draw to that food will diminish. Examples of restaurants that were removed due to cost are: Godfathers Pizza, China Buffet, Subway, Quiznos, Moe's, and Lonestar. Chain stores were fairly easily classified but local businesses were called or researched online to ascertain appropriate classification. Specifically, 373 stores were chain stores and classified in bulk, while 200 had to be checked with the use of local directories and field work. This is an important thing to note as it may have affected the results – chain stores were classified in bulk - the actual inventory of these stores was not checked due to time and financial constraints on this study.

Food location address data was geocoded using four sources (Texas A&M geocoding service, google geocoder, SNAP geocode locations, and Census geocode locations). Coordinates from all three sources were compared for accuracy – priority was given to the SNAP geocodes, followed by Census, then Texas A&M, and then google. This priority rating was determined after checking a random sample of geocodes produced by each source. These locations were then plotted as points, outliers were identified and corrected, and a random sample was checked for accuracy. Population weighted centroid coordinates were also plotted as points. These points were placed in a feature dataset with the road network mentioned above.

The network feature dataset allowed for the calculation of distance from each population weighted centroid to the nearest healthy store, unhealthy store, and store of any kind by road. Service areas along the road network were also created from each centroid at distances of 800 meters, 1600 meters, 3200 meters, 8 kilometers, and 16 kilometers. These service areas were then spatially joined to food location data in order to calculate the number of food stores within each service area and the MRFEI.

4. Sociodemographic Factors

As mentioned previously studies on the relationship between the built environment and obesity have shown significant differences across various demographics (Boone et al. 2013; Choi and Suzuki). The sociodemographic variables used for estimation came from the 2013 American Community Survey (ACS) 5-year estimates (United States Census Bureau 2013a). The American Community Survey is a yearly survey run by the federal government in an effort to give communities current information. The ACS provides 1-year, 3-year, and 5-year estimates. The 5-year estimates include 60 months of collected data, provide data for all areas, have the largest sample size of any of the ACS programs, are most reliable and least current. They are best suited when precision is more important than being current, and when the researcher desires to look at areas smaller than tracts. ACS data was collected at both the county and block group levels. Block groups typically contain between 600 and 3000 people. The block group is a statistical division of census tracts larger than a census block and typically contiguous. Most block groups were delineated by local participants.

C. DATA ANALYSIS PROCEDURES

This research involved multiple distinct research processes in order to explore the data, understand the study area, and adequately answer the research questions. As is suggested in the research questions, much of the analysis involved exploring new ways to investigate and quantify relationships. The data were messy and in many cases resisted being quantified so open-minded and exploratory methods were used. Below, the methods used to answer each question are described.

1. Obesity Estimation

In this research the application of a combination of approaches to small area estimation were employed in an effort to improve upon methods and find the best predictive model. Obesity estimation from the county level to the census block group level was done in a series of steps. First, sociodemographic variables that have been identified in previous literature as having a significant relationship with obesity were gathered. These variables included: marital status, income, poverty, educational attainment, rural/urban classifications, commuting behavior, occupation, race, sex, age, housing characteristics, and family type. In total, there were 52 variables. Measures of central tendency, distributions, and bivariate correlations between variables and with obesity were analyzed to better understand the data and its potential fit in a linear regression. With these 52 variables, stepwise backward elimination ordinary least squares (OLS) regression was conducted ($p>0.2$ for removal). This method involved simply removing the least significant variable one at a time until all variable coefficients were desirably significant. Stepwise regression was employed to reduce variables and reduced the number of variables to 19. A coefficient p-value of significance of 0.2 for removal was used to aid in the elimination of some variables without eliminating too many. The variables exhibited a lot of interaction and I desired to retain the maximum number of variables for further modeling.

Next, OLS regression was run for every combination subset of the variables. While time consuming, this method provided the opportunity to eliminate multicollinearity (VIF>7.5) while also capturing unique variable interactions (Braun and Oswald 2011). A stricter VIF value (7.5) to eliminate multicollinearity was used because the effects of multicollinearity on GWR models are considerably stronger and correlation between local regression coefficients can lead to invalid interpretation (Wheeler and Tiefelsdorf 2005).

As mentioned previously, GWR was found in multiple studies to be the most reliable method for predictive models. A geographically weighted regression is simply a type of regression model with geographically varying parameters. The basic GWR equation is as follows:

$$y(u,v) = \sum_k \beta_k(u,v) x_k(u,v) + \varepsilon(u,v)$$

where $y\ (u, v)$ is the dependent variable, $x_k(u, v)$ is the Kth independent variable at locations u and v, ε_i is the Gaussian error at locations u and v, and coefficients $\beta_k(u_, v)$ are varying conditionals on that location (Fotheringham et al. 2002). The Gaussian kernel used to solve each local regression was fixed and the extent of the kernel was determined using the corrected Akaike Information Criterion.

A python script was written to run a geographically weighted regression on all 11,230 passing models. The script created an output table with the following diagnostic values: bandwidth, effective number, residual squares, sigma, AICc, r-squared, and adjusted r-squared.

Geographically weighted regression was conducted to capture spatial variability and honor Tobler's first law[7].

The AICc and adjusted r-squared values of the resultant GWR models were used to select the top 14 models (see table 2). These models were used to predict obesity at the block group level using GWR. The predicted obesity rates were then re-aggregated back to the county level and the standardized residuals were analyzed. The relative root mean square error (rRMSE), or sample mean normalized square root of the mean square of all errors in the region, was calculated and analyzed (ii). This formula takes the RMSE (i), multiplies it by 100 and divides by the sample mean. Relative RMSE is more comparable over many study regions.

$$\text{(i)} \qquad \text{RMSE}=\sqrt{\frac{\Sigma_{i=1}^{n}\left(X_{obs,i}-X_{model,i}\right)^2}{n}}$$

$$\text{(ii)} \qquad \text{rRMSE} = RMSE \times \frac{100}{\bar{x}}$$

Spatial autocorrelation of the standardized residuals was checked for the best model using the Global Moran's I test. While AICc and adjusted r-squared results were used to select the top 14 models, the final model decision was based on the rRMSE because this value suggests that the actual localized predictions were most accurate. As suggested in the literature, if a small area estimate does not closely match estimates or measures for a larger geographic area when aggregated, red flags should be raised regarding reliability and accuracy (Bell et al. 2013).

The geographic variation of obesity in southern Illinois was then explored using local cluster measures including Anselin Local Moran's I. The Anselin Local Moran's I test identifies clusters of similar values and spatial outliers. The cluster and outlier analysis was run with an

[7] "Everything is related to everything else, but near things are more related than distance things."

optimal fixed distance band of 23552.4083 meters – this distance was determined based on peak clustering of obesity rates.

2. **Walkability**

A walkable area is one that combines population density, pedestrian friendly design, and diversity of destinations. Multiple methods have been proposed for quantifying walkability, but some were not appropriate for the study region. In order to find a good method for quantitatively measuring walkability in the largely non-urban study area, the correlation between potential measures of walkability and *WalkScore* data was analyzed (see figure 10). Measures of central tendency and distributions of these variables were also observed, as well as their bivariate correlations with each other, food measures, and obesity. Measures highly correlated with *WalkScore*s were then mapped and local cluster and outlier analysis was conducted (Anselin Local Moran's I). These measures included food distance and density measures, as well as commuting behavior.

3. **Food Environment**

With no knowledge of the influence of healthy food density on obesity in non-urban areas and little knowledge of the relationship between food access and health outside of urban areas, it was necessary to take an open-minded approach to understanding their influence. Eighteen measures were created to quantify the food environment. These measures were the distance to healthy food, distance to unhealthy food, distance to any food, healthy food within a service area (800 meters, 1600 meters, 3200 meters, 8 kilometers, and 16 kilometers), food of any kind within a service area, and MRFEI for each service area. The relationship between these variables and obesity as well as the variable distributions and measures of central tendency were analyzed in depth. The interactions between variables were also analyzed. This allowed for a more robust understanding of the relationship between the food environment at various distances, rather than a strict cut off point. The food environment was then mapped and local cluster analysis was conducted.

4. Built Environment and Obesity

The final stage of research involved looking at the relationship between walkability, food, and obesity in southern Illinois – after controlling for sociodemographic variables. Due to strong multicollinearity between variables, an OLS was run for all model subsets eliminating models with high VIF values (VIF>7.5)[8]. Models with significant spatial autocorrelation (p<0.1) of standardized residuals (revealed by the Global Moran's I test) were also removed. Of the passing models the best one was selected based on model fit (adjusted r-squared) and theoretical importance.

The best model was then run in GWR with an adaptive Gaussian kernel meaning where the feature distribution was dense the spatial context was smaller and where it was sparse the spatial context was larger. The geographically weighted regression was run for the entire study area, urban block groups of the study area only, urban cluster block groups only, and rural block groups only. The best model was found with the same GWR and OLS procedures for urban block groups, urban cluster block groups, and rural block groups[9].

The results are reported below for the: best model for the entire study region, that model in the subareas only (urban, urban cluster, rural), the best urban model, the best urban cluster model, and the best rural model. Global and local regression results as well as maps of the

---

[8] Again, a VIF cut-off of 7.5 was used due to the danger of introducing multicollinearity into GWR models.
[9] OLS was run for all model subsets of urban (n=76), urban cluster (n=176), and urban areas (n=140). OLS models with high VIF values (VIF>7.5) or spatially autocorrelated residuals (p<0.1) were removed. Then all passing models were run in GWR and the best one was selected for the sub-area based on adjusted r-squared and theoretical value.

standardized residuals for each of these seven models are reported. Modeling for rural, urban cluster, and urban areas separately shows how the relationship between the built environment and obesity varies across areas of different sizes and assists in shedding light on the many studies conducted in urban areas only.

D. DELIMITATIONS

This study was geographically confined to the eighteen southernmost counties of Illinois. It focuses on the relationship between the built environment and obesity using data from one year, without attempting to account for factors that cannot be quantitatively measured. Finally, it should be kept in mind throughout that all variables are estimates of a study population and merely suggestive of population trends or patterns.

IV. RESULTS

A. Obesity

As mentioned above, 52 sociodemographic variables from the American Community Survey (2013a) were initially considered for small area obesity estimation. These variables measured: marital status, household income, educational attainment, mode of transportation to work, average travel time to work, occupation, rural/urban area, employment status, race, sex, age, family type (single parent), and housing characteristics (see table 1).

TABLE 1. STUDY AREA FOR OBESITY ESTIMATION: CENTRAL TENDENCY AND PEARSON CORRELATION WITH OBESITY

	Mean (s_x)	Pearson Correlation with Obesity (p-value)
% Married	54.33 (5.25)	0.011 (0.782)
% Previously Married*	20.72 (3.17)	0.405 (0.000)
% Never Married*	24.95 (5.30)	-0.253 (0.000)
% Household income < $10k	7.99 (3.76)	0.347 (0.000)
% Household income $10k-15k	6.60 (2.19)	0.373 (0.000)
% Household income $16-25k	12.98 (2.92)	0.374 (0.000)
% Household income $26-35k	12.07 (2.01)	0.217 (0.000)
% Household income $36-50k	15.51 (2.08)	0.040 (0.331)
% Household income $51-75k	19.36 (2.81)	-0.232 (0.000)
% Household income $76-100k	11.80 (2.71)	-0.341 (0.000)
% Household income $101-150k	9.51 (3.45)	-0.351 (0.000)
% Household income $151-200k	2.36 (1.43)	-0.352 (0.000)
% Household income > $200k	1.84 (1.41)	-0.362 (0.000)
% Less than HS education	14.40 (6.23)	0.496 (0.000)
% HS Graduate*	38.28 (5.79)	0.335 (0.000)
% Some college no BA/BS*	29.48 (4.37)	-0.227 (0.000)
% BA/BS +*	17.85 (7.38)	-0.547 (0.000)
Rural	61.54 (28.76)	0.300 (0.000)
% Car to work*	91.09 (3.60)	0.273 (0.000)
% Walk, bike, take pub. Trans to work*	3.30 (2.16)	-0.299 (0.000)
% Other transportation to work*	1.32 (0.82)	-0.031 (0.444)
% Work at home*	4.30 (2.21)	-0.138 (0.001)
Average travel time to work (minutes)	23.13 (4.41)	0.144 (0.000)
% Management*	29.09 (4.88)	-0.396 (0.000)
% Service industry	17.56 (2.88)	0.086 (0.034)

TABLE 1. CONTINUED

% Sales industry	22.49 (2.95)	-0.185 (0.000)
% Blue collar industry	30.86 (6.26)	0.356 (0.000)
% Employed*	55.59 (7.60)	-0.446 (0.000)
% Unemployed*	8.59 (2.83)	0.283 (0.000)
% Hispanic	3.23 (3.66)	-0.173 (0.000)
% White	91.08 (8.50)	0.133 (0.001)
% Black	2.99 (5.12)	-0.049 (0.231)
% American Indian*	0.52 (3.45)	0.083 (0.042)
% Asian*	0.78 (1.18)	-0.404 (0.000)
% Pacific Islander	0.03 (0.13)	-0.006 (0.881)
% Other race	0.06 (0.12)	-0.110 (0.007)
% 2+ races*	1.31 (0.73)	-0.059 (0.147)
% Women	0.62 (0.03)	0.018 (0.665)
% Women < 18 years	25.45 (4.32)	-0.015 (0.718)
% Women 20-24 years	5.82 (2.27)	-0.221 (0.000)
% Women 25-34 years	11.55 (2.38)	-0.013 (0.746)
% Women 35-44 years	11.85 (1.28)	0.089 (0.030)
% Women 45-54 years	14.30 (1.82)	-0.021 (0.611)
% Women 55-64 years	13.35 (1.46)	0.123 (0.003)
% Women 65-74 years*	9.22 (1.63)	0.186 (0.000)
% Women > 75 years	7.59 (2.67)	0.050 (0.221)
% Vacant house	14.40 (8.58)	0.104 (0.011)
% Renter occupied house	25.77 (6.10)	-0.085 (0.038)
Median house value*	110031.39 (35923.01)	-0.485 (0.000)
Median year house occupied	2001.53 (1.41)	-0.249 (0.000)
Land area (square miles)*	543.71 (267.76)	-0.219 (0.000)
% Single parent families*	21.28 (4.77)	0.324 (0.000)

The stepwise backward removal regression (p-value for removal > 0.2) reduced the number of variables to 19 (one was retained for theoretical purposes and measures the area of land; see starred variables in table 1). These variables measured marital status, educational attainment, mode of transportation to work, occupation, employment status, race, age, sex, housing characteristics, and family type. All of these variables are continuous, there is a linear relationship with obesity, there were no significant outliers, the observations were independent, the data was homoscedastic, and the errors were normally distributed.

As reviewed above, an OLS regression was then run for every possible combination of the 19 variables. The resultant 11,230 models had between 1 and 12 variables with adjusted r-squared values ranging from 0.005-0.401 and AICc values ranging from 2688.226 to 2982.094. The OLS model with the highest adjusted r-squared value[10] included the area land, and the following percentage values: previously married, never married, college degree or more, other transportation used to get to work, management occupation, employed, unemployed, American Indian, two or more races, females 65 to 74, and single parent families. This model did not have spatially autocorrelated standardized residuals (p>0.05).

A GWR was run for all 11,230 of the passing OLS models. The adjusted r-squared values for geographically weighted regression models ranged from 0.176 to 0.433 and AICc values ranging from 2674.774 to 2885.446[11]. Of the 11,230 GWR results, 14 were selected based on their corrected Akaike Information Criterion value and adjusted r-squared value (see table 2). Obesity predictions at the block group level were computed for these top 14 models using GWR, they were re-aggregated back to the county level, and the relative root mean square error (rRMSE) was calculated. The model with the lowest rRMSE for the study region (model 12 in table 2) was selected as the best prediction model (see figure 5). The rRMSE for southern Illinois of the best model was 7.836. This model was chosen despite the fact that it did not have the highest adjusted r-squared or lowest AICc because it produced the lowest level of error in block group predictions for the study area. In OLS this model did not have spatially autocorrelated residuals (p>0.05).

---

[10] The AICc for this model was 1.26 points above the lowest AICc value (2689.486 and 2688.226 respectively).
[11] The highest adjusted r-squared value for GWR was higher (0.433 compared to 0.401) and the lowest AICc was lower (2674.774 compared to 2688.226). These results suggest, among other things, that GWR is a better fit than OLS. A differences of more than 3 points in AICc values suggests that the model is a better fit (ESRI 2016).

TABLE 2. TOP GWR MODELS FOR PREDICTING OBESITY

Model	Variables	AICc	Adj. R^2
1	Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Walk, Bike, or Take Public Transportation; American Indian; Two or More Races; Median House Value; Single Parent Family	2672.825	0.433
2	Area Land; Previously Married; Never Married; College Degree or More; Commute by Car; Commute Another Way; Management Occupation; Employed; Unemployed; American Indian; Two or More Races; Single Parent Family	2670.899	0.432
3	Area Land; Never Married; High School Graduate; College Degree or More; American Indian; Two or More Races; Female 65-74; Single Parent Family	2671.472	0.430
4	Area Land; Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Employed; Unemployed; American Indian; Two or More Races; Single Parent Family	2674.774	0.433
5	Never Married; College Degree or More; Employed; American Indian; Two or More Races; Female 65-74; Median House Value; Single Parent Family	2672.415	0.429
6	Never Married; High School Graduate; College Degree or More; American Indian; Two or More Races; Female 65-74; Median House Value; Single Parent Family	2672.905	0.428
7	Previously Married; Never Married; College Degree or More; Employed; American Indian; Two or More Races; Female 65-74; Median House Value; Single Parent Family	2675.279	0.429
8	Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Employed; Unemployed; American Indian; Two or More Races; CFEM65TO74	2673.544	0.427
9	High School Graduate; College Degree or More; Commute by Car; Commute Another Way; Work at Home; American Indian; Asian; Female 65-74; Median House Value; Single Parent Family	2672.352	0.426
10	Area Land; Area Water; Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Employed; Unemployed; American Indian; Two or More Races; Single Parent Family	2676.236	0.429
11	Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Walk, Bike, or Take Public Transportation; Two or More Races; Median House Value; Single Parent Family	2674.765	0.426
12	*Area Land; Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Employed; Unemployed; Two or More Races; Single Parent Family*	*2675.991*	*0.428*
13	Previously Married; Never Married; High School Graduate; Some College no Bachelors; College Degree or More; Employed; Unemployed; Two or More Races; Single Parent Family	2674.15	0.424
14	Area Land; Previously Married; Never Married; College Degree or More; Walk, Bike, or Take Public Transportation; American Indian; Two or More Races; Median House Value; Single Parent Family	2674.725	0.424

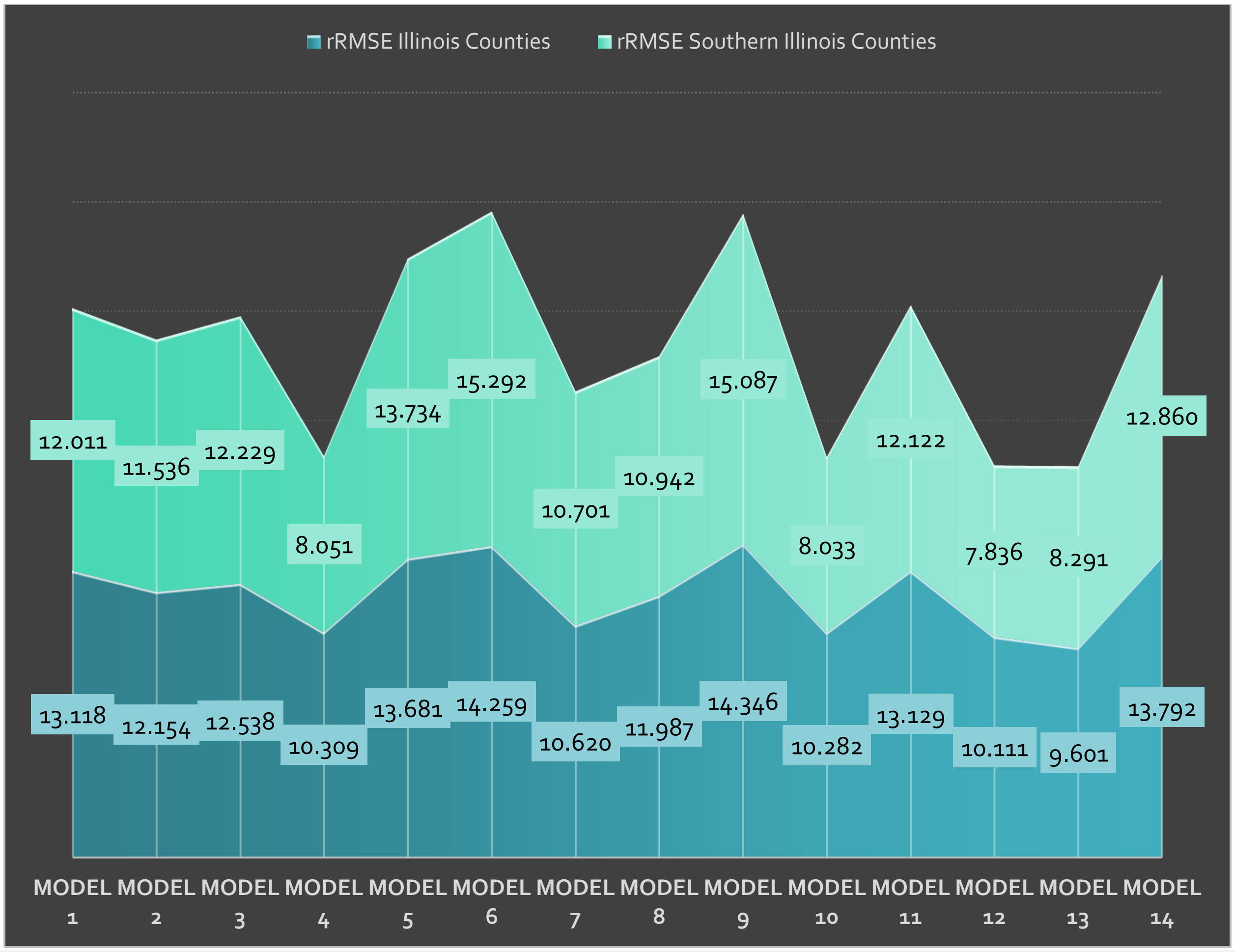

FIGURE 5. RELATIVE RMSE OF TOP PREDICTION MODELS, SOUTHERN ILLINOIS AND ILLINOIS

The best model for predicting obesity in southern Illinois included 10 variables: area of land, percent previously married, percent never married, percent with a high school education only, percent with some college education but no bachelor's degree, percent with a college degree or more, percent employed, percent that identify as two or more races, and the percentage of single parent families (see figure 6)[12]. The adjusted r-squared for this model was 0.428 and the AICc was 2675.991.

[12] The best model was model 12 in table 2 and figure 5.

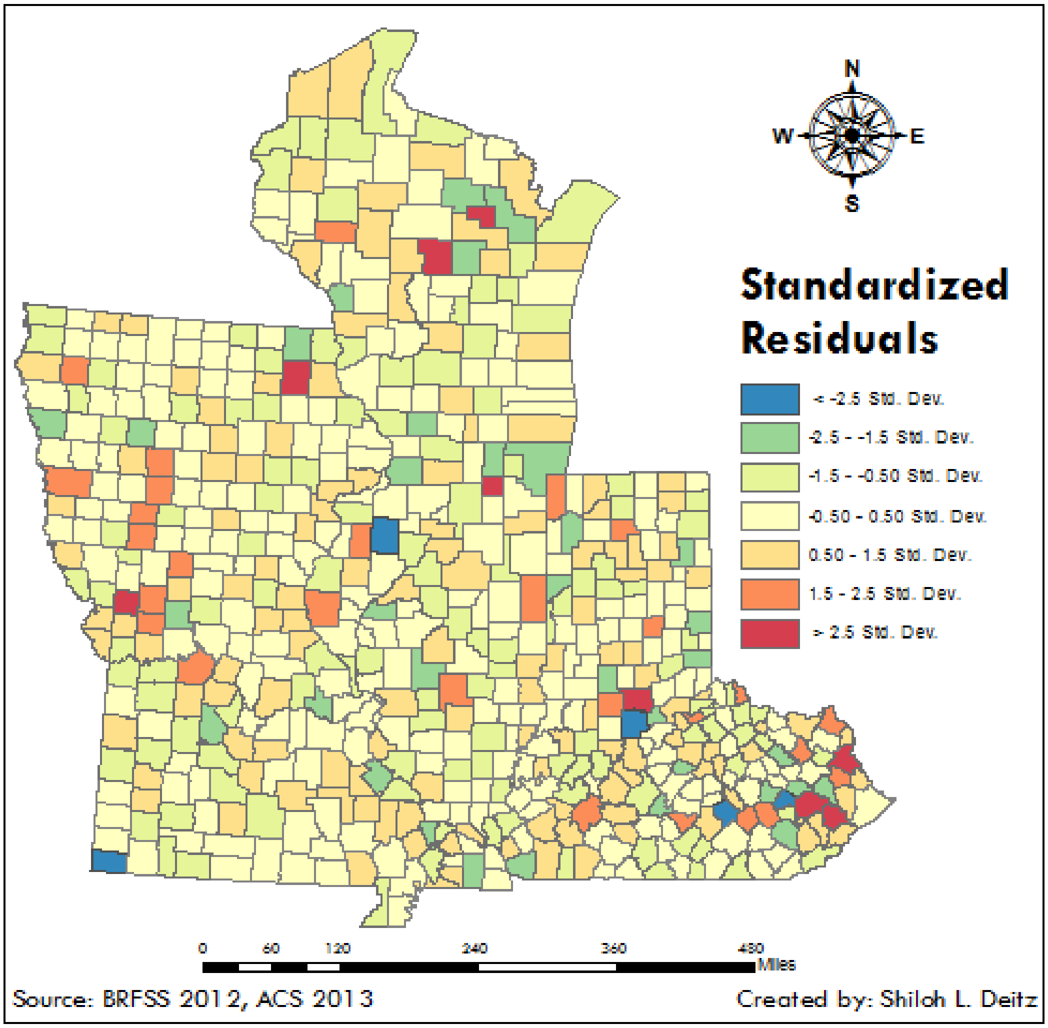

FIGURE 6. GWR FOR PREDICTION RESULTS (M12)

In southern Illinois the block group percentage obese ranges from 20.24% to 46.28%, the average rate is 32.76% (s_x=3.57) and the median is 33.08%. Sixty percent of the block groups have obesity rates between 30.30% and 35.34% (see figure 7). There is a significant correlation between the percentage that have a college degree or more within a block group and obesity (-0.724; p<0.001), and average income and obesity (-0.236; p<0.001). There is a lesser correlation between obesity and age, sex, race, or income. The correlation is also strong between urban/rural classification and obesity (-0.348; $p<0.001$)[13] (see figure 8).

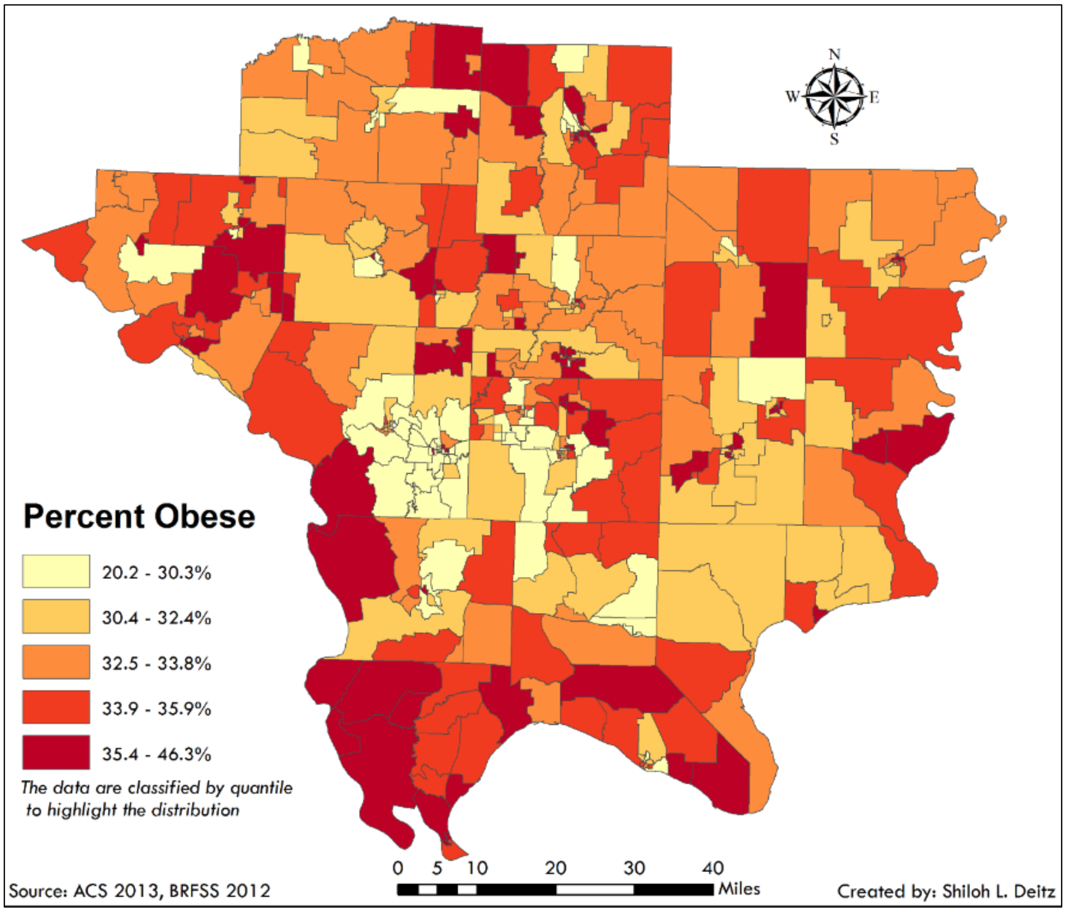

FIGURE 7. OBESITY RATES IN SOUTHERN ILLINOIS BY BLOCK GROUP

---

[13] The rural/urban variable has been classified as follows: rural (1), urban cluster (2), and urban area (3).

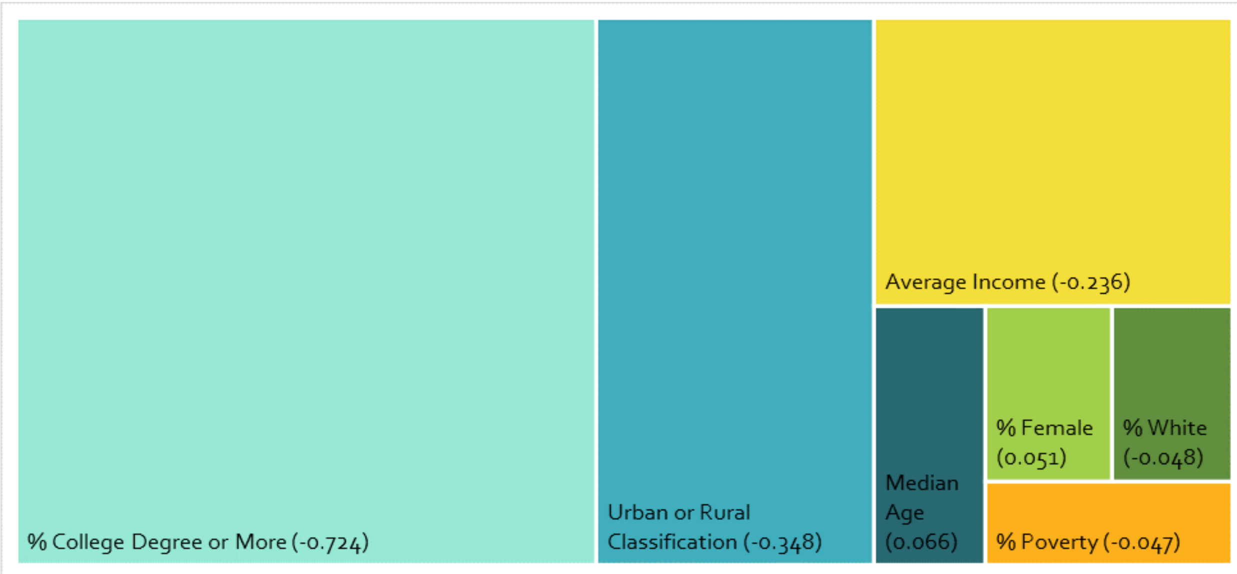

FIGURE 8. TREEMAP OF TOP SOCIODEMOGRAPHIC PEARSON CORRELATIONS WITH OBESITY

There are significant clusters of neighborhoods with low levels of obesity in Anna, Carbondale, and Murphysboro (see figure 9). The obesity rate in these neighborhoods ranges from 20.24% to 31.34%. Of the neighborhoods that cluster with other neighborhoods with low obesity rates four are rural, eight are urban clusters, and 42 are considered urban. There are significant clusters of neighborhoods with high obesity rates in Benton, Cairo, and Mount Vernon. In these neighborhoods the obesity rate ranges from 35.46% to 45.74%. There are also significant outliers with high obesity rates near the Carbondale area and outliers due to low obesity rates near the Mount Vernon area. In Mt. Vernon there are low outlier neighborhoods with obesity rates ranging from 25.45% to 28.62%. In the Carbondale area there are high outlier neighborhoods with obesity rates ranging from 34.68% up to 46.28%.

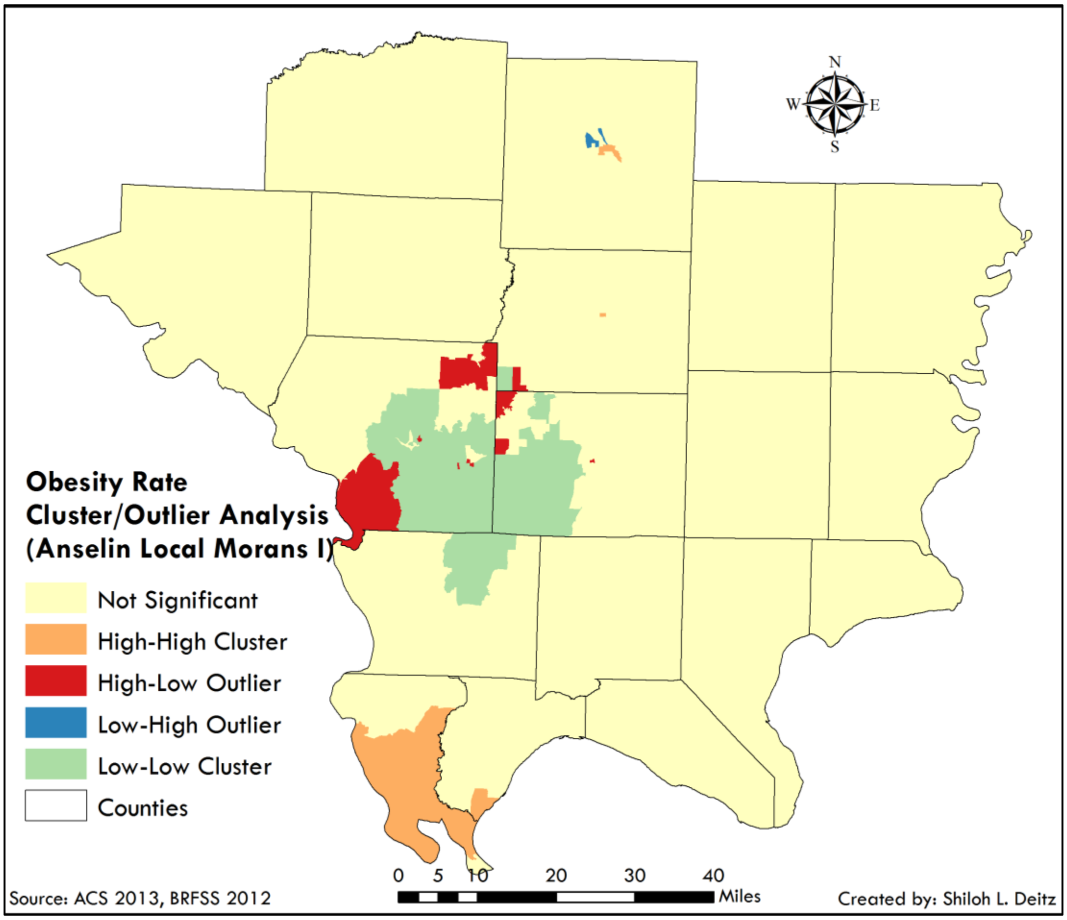

FIGURE 9. CLUSTER/OUTLIER ANALYSIS OF OBESITY RATES

B. WALKABILITY

As mentioned above, measures of walkability were assessed according to their correlation with *WalkScore* data (see figure 10). This was done because *WalkScore* has been established as a good indicator of walkability (Carr et al. 2010; Carr et al. 2011). Sequentially, the highest correlations are for number of food stores within 1600 meters (0.839; $p<0.001$), food within 3200 meters (0.702; $p<0.001$), food within 800 meters (0.699; $p<0.001$), distance to any food (-0.552; $p<0.001$), percentage that walk, bike, or take public transportation to work (0.3131; $p<0.001$), average commute time (-0.312; $p<0.001$), number of food stores within 8 kilometers (0.304; $p<0.001$), and number of food stores within 16 kilometers (0.171; $p<0.01$). All of these measures are significantly correlated with *WalkScore* (see table 3). There are also significant correlations amongst the food measures.

FIGURE 10. PEARSON CORRELATIONS OF POTENTIAL WALKABILITY MEASURES AND *WALKSCORE*

TABLE 3. PEARSON CORRELATIONS BETWEEN WALKABILITY MEASURES

	Walk Score	Food800	Food1600	Food3200	Food8km	Food16km	AllDistance	AVGTRTIME	%WalkBike PublicTrans
Walk Score	1.00								
Food800	0.70	1.00							
Food1600	0.84	0.73	1.00						
Food3200	0.70	0.50	0.80	1.00					
Food8km	0.30	0.21	0.44	0.73	1.00				
Food16km	0.17	0.13	0.27	0.42	0.60	1.00			
AllDistance	-0.55	-0.36	-0.47	-0.49	-0.33	-0.24	1.00		
AverageTravelTime	-0.31	-0.20	-0.36	-0.46	-0.54	-0.36	0.31	1.00	
%WalkBikePublicTran	0.31	0.35	0.41	0.40	0.30	0.14	-0.14	-0.33	1.00

**Significant values are highlighted (p<0.001)*

The average number of stores within 1600 meters of the block group population weighted centroid is 3.41 (s_x=4.92). In 46% of the neighborhoods there are no food sources within that distance. Sixteen neighborhoods have fifteen or more food stores within about a mile. These neighborhoods are in Benton, Carbondale, Harrisburg, and Mount Vernon (see figure 11).

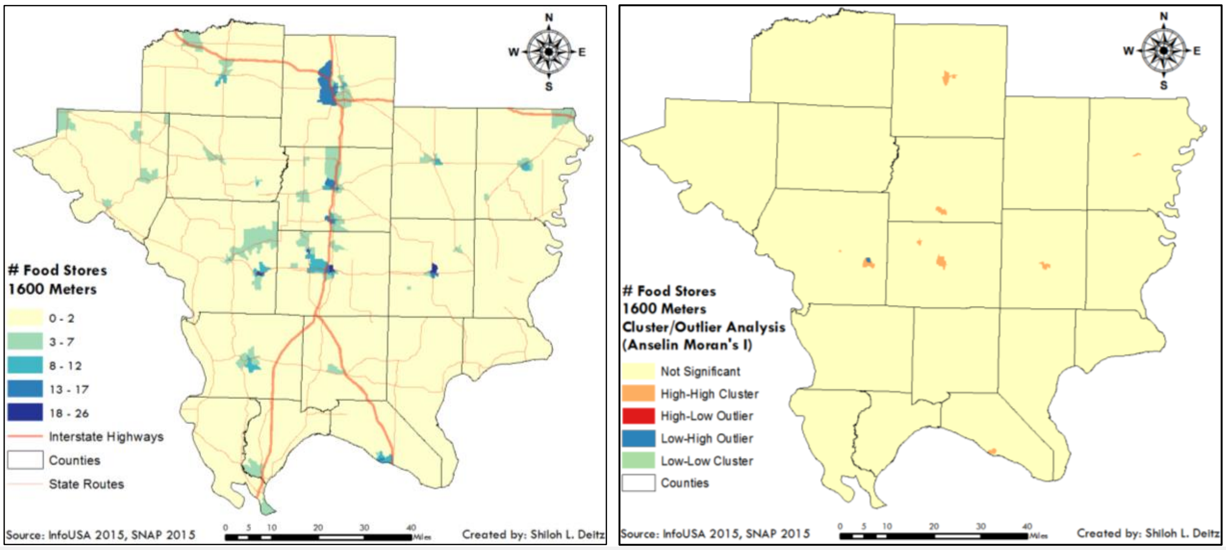

FIGURE 11. FOOD STORES WITHIN 1600M AND CLUSTER/OUTLIER ANALYSIS[14]

There is one low outlier in the city of Carbondale which otherwise has significantly high numbers of food stores within 1600 meters of individual neighborhoods. For this neighborhood, the nearest food source is a RollnUp and it is 1660 meters away. There are 43 neighborhoods that cluster with others with high numbers of food stores within 1600 meters. These neighborhoods are along interstate or state highways in the cities of Carbondale, Carmi, Harrisburg, Metropolis, Mount Vernon, Murphysboro, and West Frankfort. The number of food stores within 1600 meters ranges from six to twenty-six. There is a significant Pearson correlation between the number of food stores within 1600 meters and both the percentage of the population in poverty

[14] Data in the number of food stores within 1600 meters map are classified by natural jenks.

(0.290; p<0.001) and the percentage white (-0.288; p<0.001) suggesting that there are more food stores in locations that are less white with higher poverty rates.

Theoretically speaking, the number of food stores within 800 meters is the best measure of walkability. The average number of stores within 800 meters of the block group population weighted centroid is 0.98 (s_x=1.95). However, in at least 69% of the neighborhoods there are no food sources within that distance. Twenty-eight neighborhoods have five or more food sources within 800 meters (see figure 12). These neighborhoods are found in Benton (2), Carbondale (10), Carmi (1), Harrisburg (2), Johnston City (1), Metropolis (3), Mount Vernon (3), Murphysboro (1), Pinckneyville (2), and West Frankfort (2).

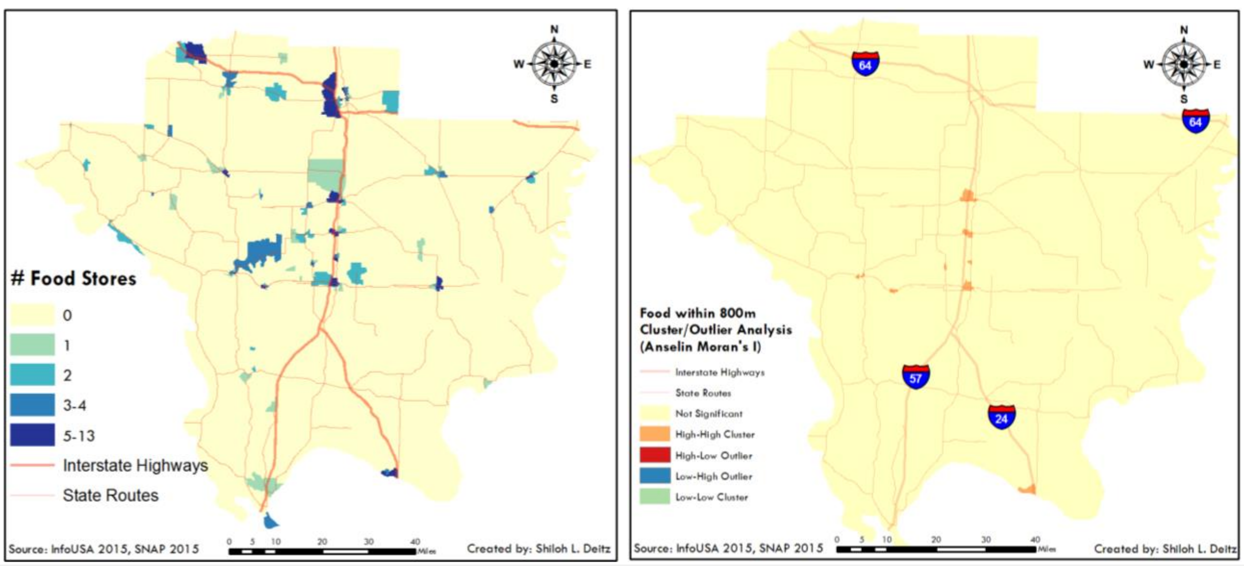

FIGURE 12. FOOD STORES WITHIN 800M AND CLUSTER/OUTLIER ANALYSIS[15]

There are a few significant clusters of block groups with high numbers of food stores within 800 meters of the population weighted block group centroids (see figure 12). The majority of these are along interstate highway 57 in a location where a state highway crosses the interstate. There are also a few along state route 13 through the Carbondale-Marion metropolitan area. The number of

[15] Data in the number of food stores within 800 meters map are classified by quantile.

food stores within 800 meters is also correlated with poverty rates (0.290; $p<0.001$) and percentage white (-0.288; $p<0.001$) suggesting that there are more food stores within this distance in areas with high poverty and minority populations.

While the average rate of workers who walk, bike, or take public transportation in southern Illinois is 3.78% (s_x=7.79), in 72% of the block groups the rate is below the average. In over 20% of the neighborhoods no one gets to work that way, in another fifth of the neighborhoods only zero to 1.81% do, in the top fifth of neighborhoods the rate varies from 6.48% to 81.41%. The top eight neighborhoods are all in Carbondale and Murphysboro (see figure 13).

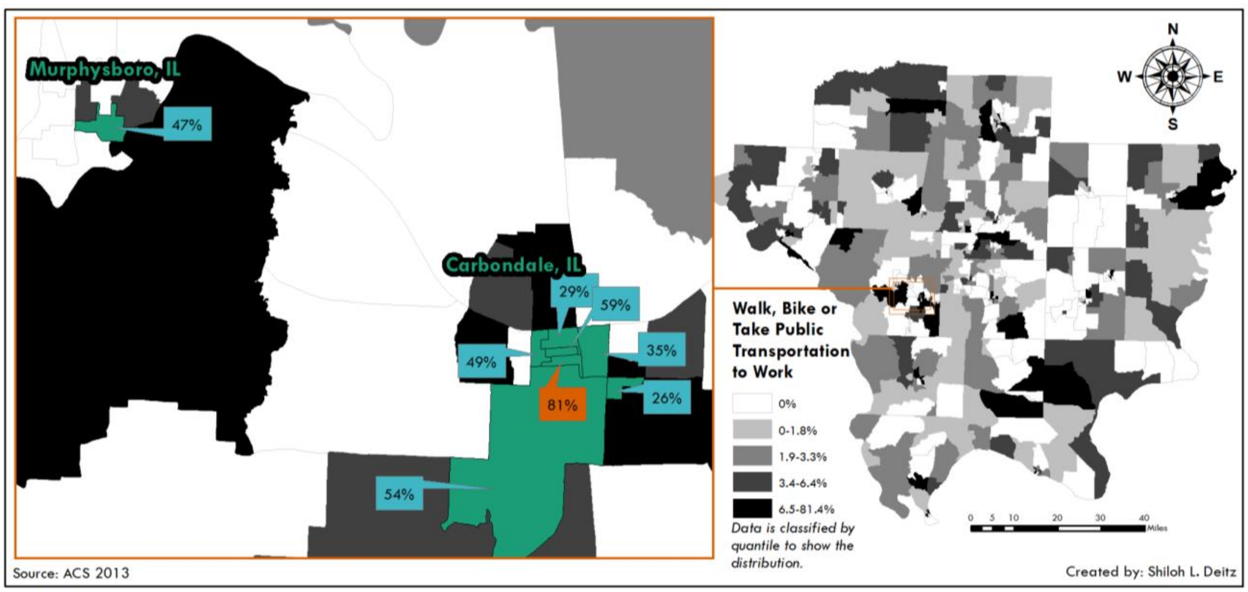

FIGURE 13. % THAT WALK, BIKE, OR TAKE PUBLIC TRANSPORTATION TO WORK

The high percentage of persons walking, biking or taking public transportation in the Carbondale area as noted above, forms a significant spatial cluster of high values. However, there are also significant outliers in the area. While some of the neighborhoods have a high rate of persons getting to work on foot, by bike, or on public transportation, other neighborhoods nearby exhibit the low rates that are typical of the region as a whole (see figure 14). The percentage that walk, bike or take public transportation to work is highly correlated with the percentage in poverty (0.475; $p<0.001$) suggesting the people with less money are more likely to get around without a car. It is also highly correlated with age (-0.327; $p<0.001$) suggesting that younger people are less likely to rely on cars.

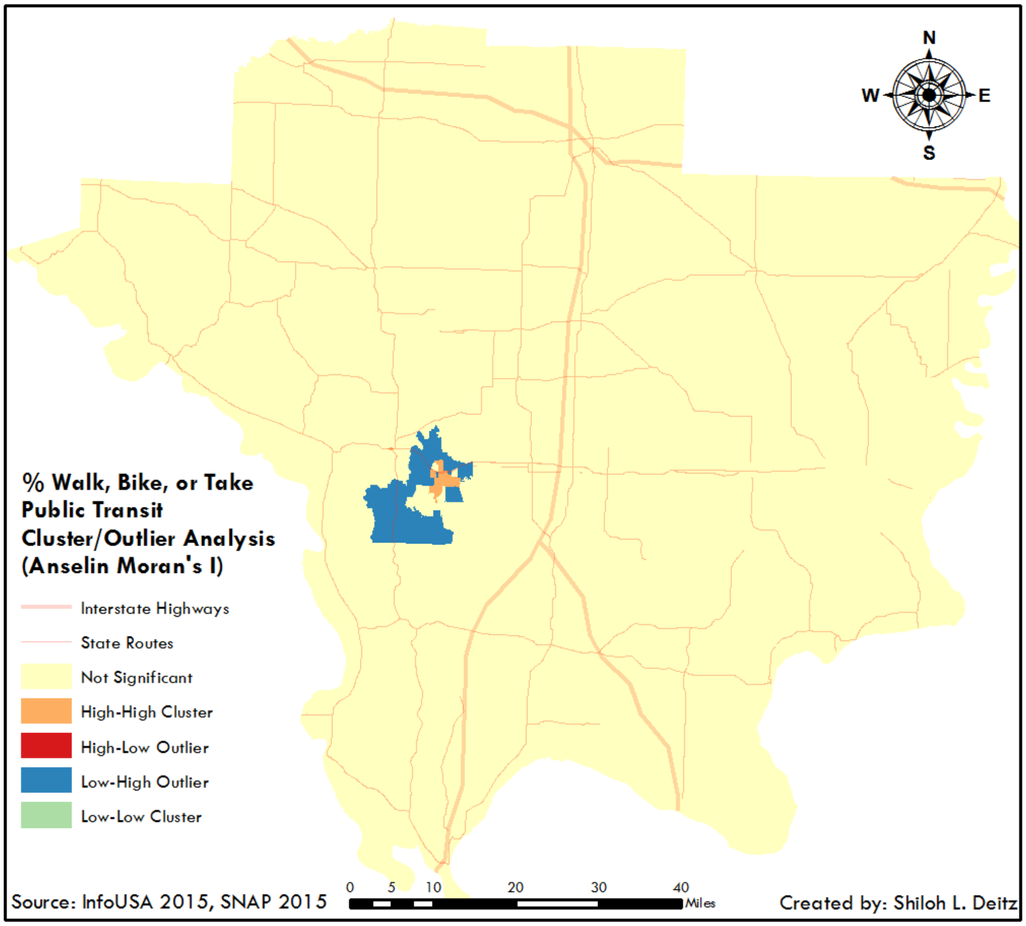

FIGURE 14. CLUSTER/OUTLIER ANALYSIS FOR WALK, BIKE, TAKE PUBLIC TRANSPORTATION TO WORK

The average commute time in southern Illinois is 24.20 minutes (s_x=6.55). The lowest average commute time for a neighborhood is 10.72 minutes, while the neighborhood with the longest commute travels 52.00 minutes to work on average. The top eight neighborhoods for shortest commute time are found in Carbondale. Specifically, of the top 20, 13 are in Carbondale, 4 are in Mt. Vernon, 2 are in Chester, and one is in West Frankfort. The average travel time in these neighborhoods with the shortest time ranges from 10.72 to 13.88 minutes. The 20 neighborhoods with the longest average travel time to work are majority rural (13). The small towns with the highest commute times are Christopher, Pinckneyville, Du Quoin, Mcleansboro, and Harrisburg. In these neighborhoods, the average commute ranges from 34.54 minutes to 52.00 minutes (see figure 15).

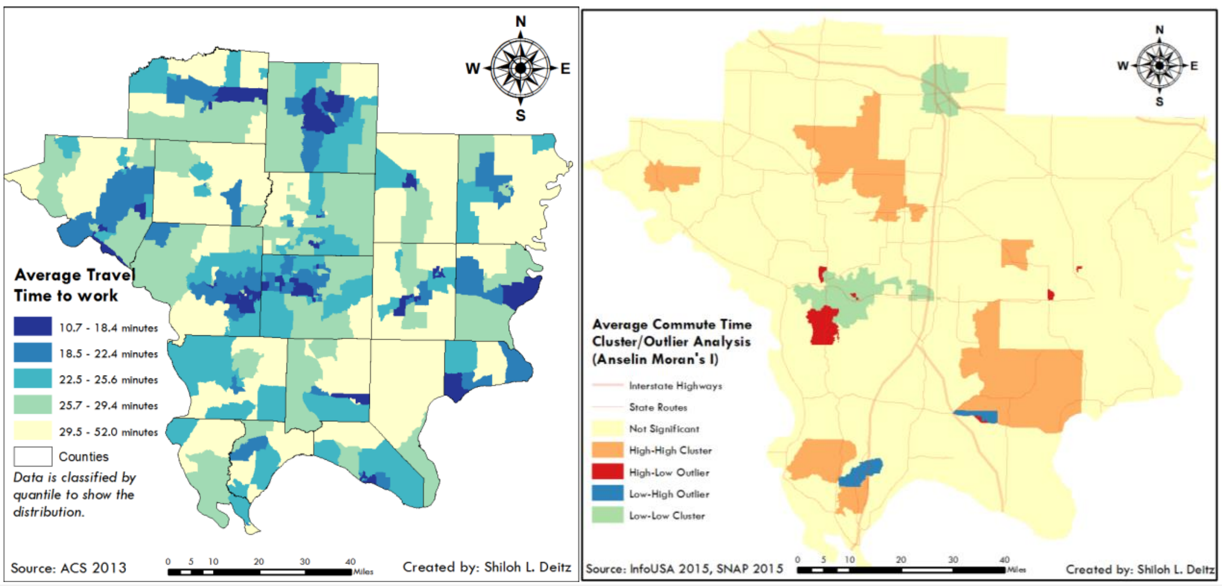

FIGURE 15. AVERAGE TRAVEL TIME TO WORK AND CLUSTER/OUTLIER ANALYSIS

There are significant clusters of high average travel times in 17 rural neighborhoods and the small towns of Christopher (2), Pinckneyville (1), and Du Quoin (1). The range of travel time for these 21 areas is 30.07 to 44.09 minutes. 57 neighborhoods cluster with other neighborhoods with relatively low average travel times to work. These neighborhoods are in Carbondale (32), Mount Vernon (19), and Murphysboro (1). The average travel time in these low clusters ranges from 10.72 to 21.25 minutes. Significant high outliers next to neighborhoods with lower average travel times are found in Carbondale (2), Eldorado (1), Harrisburg (1), Murphysboro (2), and two rural neighborhoods. The average travel time in these outlier neighborhoods ranges from 26.80

to 52.00 minutes. There are two significant low outliers next to areas with high average commute times. These neighborhoods are in a rural part of the southern area of the study region and their average travel times are 13.98 and 18.79 minutes (see figure 15). Average commute time is significantly correlated with the percentage that have a college degree or more (-0.335; $p<0.001$) suggesting that areas with higher average education tend to have lower commute times. It is also correlated with median age (0.255; $p<0.001$) suggesting that older areas tend to have longer average commute times.

C. FOOD ENVIRONMENT

As is typical across the U.S., only 21% of the 573 food sources in southern Illinois are healthy. In some areas of the study region people have to travel much more than 10 miles to reach healthy food and many travel at least 10 miles. The coverage of the region by unhealthy food is much greater. In fact, there are only 13 neighborhoods where the distance to unhealthy food is 10 miles or more. Figure 16 shows the distance to food stores (both healthy and unhealthy) for residents. Living outside of the colored service areas suggests a travel distance of over 16 kilometers to food.

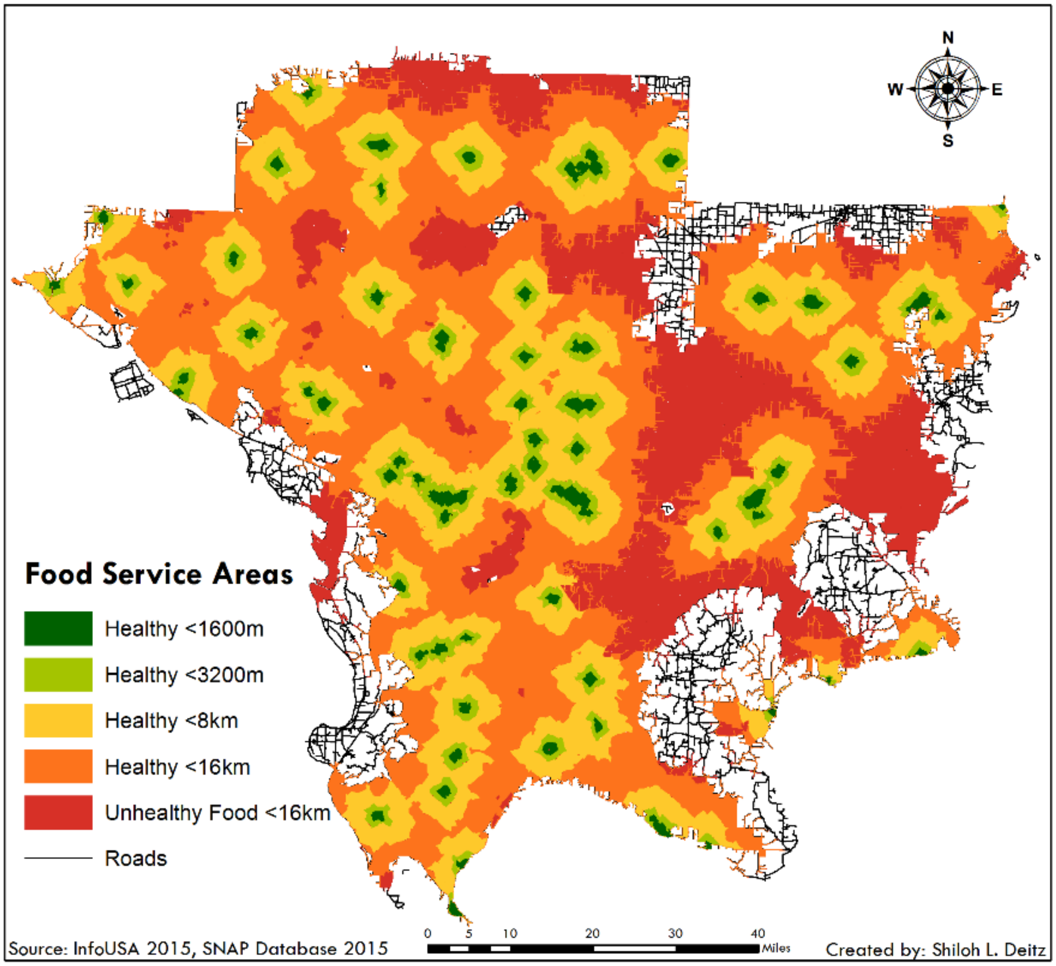

FIGURE 16. HEALTHY AND UNHEALTHY FOOD SERVICE AREAS

On average there are 0.98 (s_x =1.95) food stores within 800 meters of southern Illinois residents' homes, 3.41 (s_x=4.92) stores within 1600 meters, 8.30 (s_x=10.51) within 3200 meters, 15.33 (s_x=16.95) within 8 kilometers, and 35.82 (s_x=31.67) within 16 kilometers (see table 4). The average number of healthy food stores within 800 meters is 0.18 (s_x=0.47), within 1600 meters it is 0.61 (s_x=1.06), within 3200 meters it is 1.54 (s_x=2.08), within 8 kilometers there are 3.08 (s_x=3.69) healthy food stores, and within 16 kilometers there are 6.88 (s_x=6.38). The average distance to any food is 3422.19 (s_x=4151.77) or 2.13 miles, the average distance to unhealthy food is 3731.51 meters (s_x=4587.54), and the average distance to healthy food is 5703.88 meters (s_x=6031.47) or 3.54 miles (see figure 17). The distance to the nearest food of any kind ranges from 19 meters to 22547 meters (14 miles). The distance to healthy food ranges from 141 meters to 34063 meters (21.17 miles). The average modified retail food environment at 800 meters is 6.37 (s_x=18.70), at 1600 meters it is 9.20 (s_x=16.71), at 3200 meters it is 13.25 (s_x=18.42), at 8 kilometers the MRFEI average is 16.35 (s_x=15.39), and at 16 kilometers the average MRFEI is 20.59 (s_x=13.57) – meaning about 1 in 5 food stores are healthy (see figure 18).

TABLE 4. CENTRAL TENDENCY AND DISPERSION OF FOOD VARIABLES

	Mean (s_x)	Range
Food800m	0.98 (1.95)	0-13
Healthy800m	0.18 (0.47)	0-2
Unhealthy800m	0.79 (1.66)	0-12
MRFEI800m	6.37 (18.70)	0-100m
Food1600m	3.41 (4.92)	0-26
Healthy1600m	0.61 (1.06)	0-6
Unhealthy1600m	2.81 (4.10)	0-20
MRFEI1600m	9.20 (16.71)	0-100m
Food3200m	8.30 (10.51)	0-44
Healthy3200m	1.54 (2.08)	0-9
Unhealthy3200m	6.76 (8.64)	0-37
MRFEI3200m	13.25 (18.42)	0-100m
Food8km	15.33 (16.95)	0-55
Healthy8km	3.08 (3.69)	0-13
Unhealthy8km	12.25 (13.54)	0-44
MRFEI8km	16.35 (15.39)	0-100m
Food16km	35.82 (31.67)	0-128
Healthy16km	6.88 (6.38)	0-25
Unhealthy16km	28.94 (25.50)	0-103
MRFEI16km	20.59 (13.57)	0-100m
UnhealthyDistance	3731.51 (4587.54)	19-24482
HealthyDistance	5703.88 (6031.47)	141-34063
AllDistance	3422.19 (4151.77)	19-22547

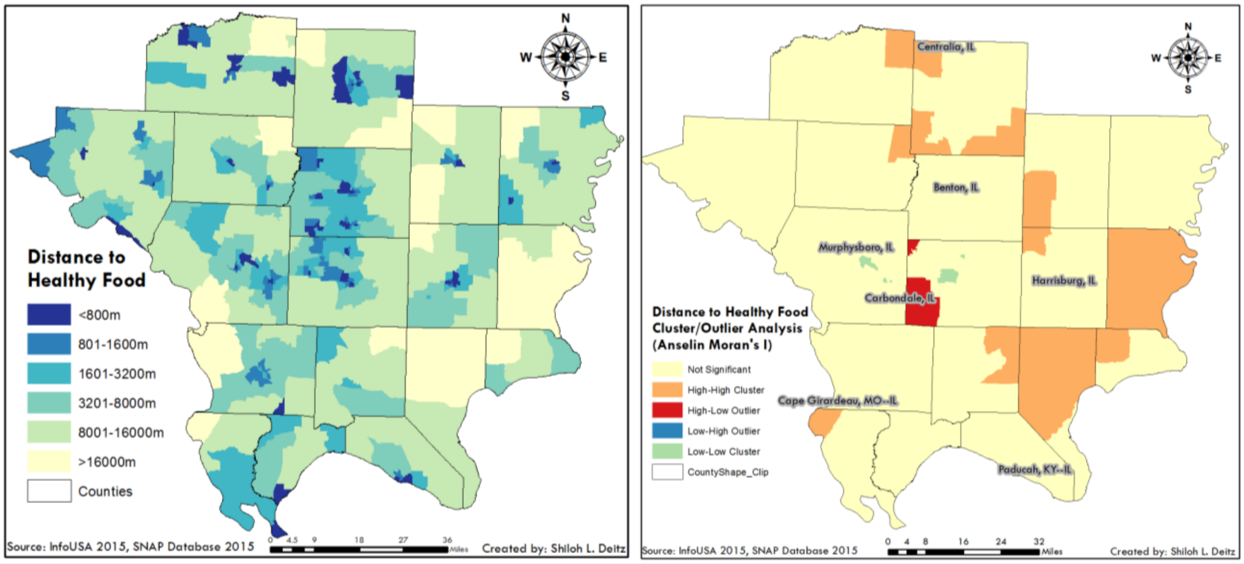

FIGURE 17. DISTANCE TO HEALTHY FOOD AND CLUSTER/OUTLIER ANALYSIS

Twenty three neighborhoods cluster with other areas with high distances to healthy food. These areas are majority rural (19), but also found in Centralia and Cape Girardeau[16]. The distance to healthy food ranges from 13708 meters to 34063 meters in these areas. Five neighborhoods form significant clusters with other neighborhoods with low distances to healthy food. These neighborhoods are in Carbondale (3), Murphysboro (1), and Johnston City (1). The distance to healthy food in these neighborhoods ranges from 239 meters to 954 meters. There are also two high outliers near the Carbondale area – in these neighborhoods the average distance to healthy food is 12494 meters and 15659 meters. These outlier neighborhoods are near to Makanda and Pulleys Mill. The average distance to healthy food is correlated with the percentage in poverty (-0.245; $p<0.001$), percentage with a college degree or more (-0.226; $p<0.001$), and percentage white (0.216; $p<0.001$) suggesting that poorer areas, areas with higher educated individuals, and higher minority populations tend to have lower average distances to the nearest healthy food.

[16] These could be false results because they are both at the edges of the study area.

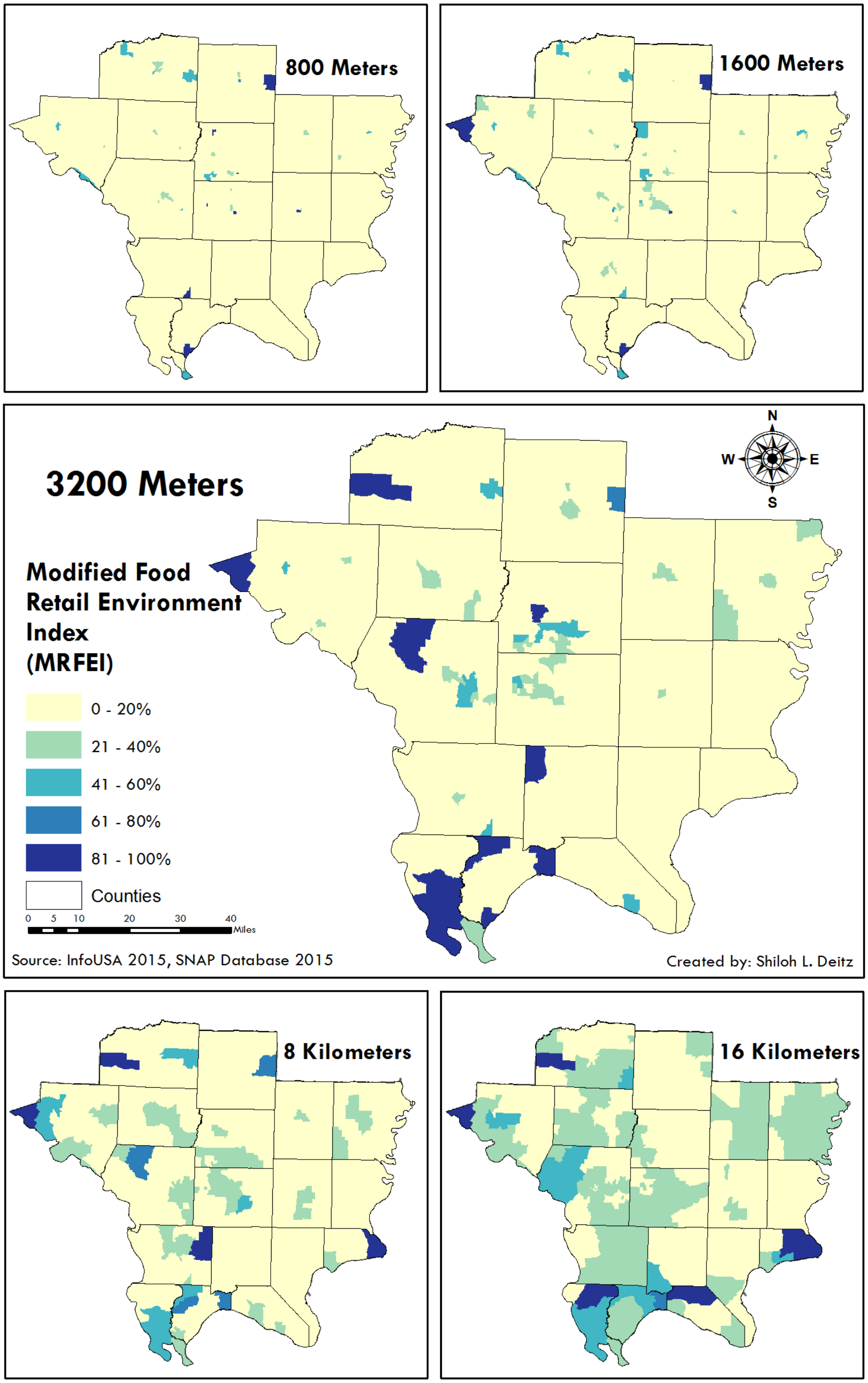

FIGURE 18. MRFEI AT MULTIPLE SERVICE AREAS

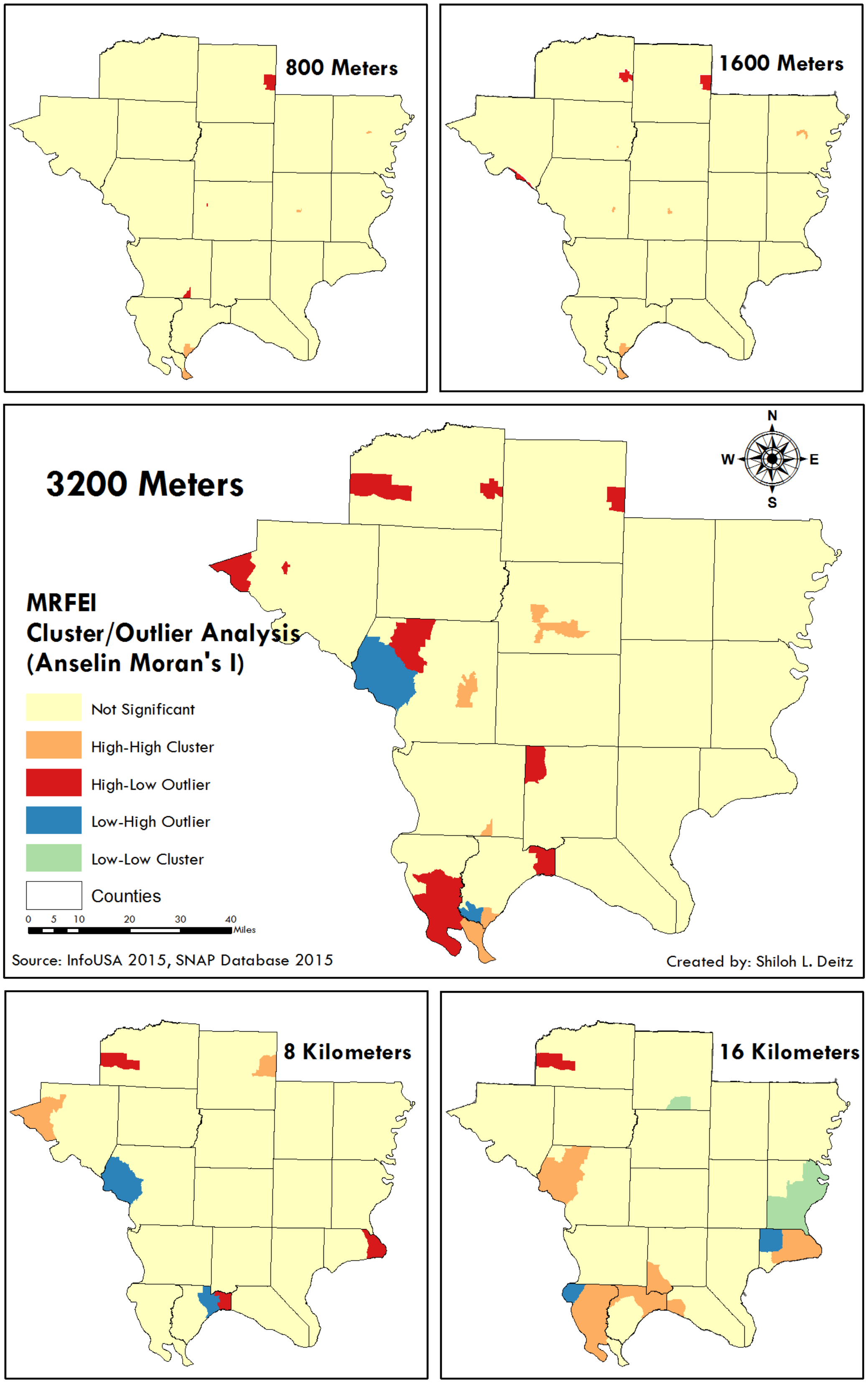

FIGURE 19. CLUSTER/OUTLIER ANALYSIS OF MRFEI

As one would expect, the MRFEI increases as the service area gets larger (see figure 18). There are significant clusters of neighborhoods with high MRFEI at 800 meters in Cairo, Carmi, and Harrisburg. There are significant high outliers in the Carbondale-Marion area, a rural neighborhood north east of Anna, and a neighborhood east of Mt. Vernon. At 1600 meters there are significant clusters of high MRFEI in the Carbondale-Marion area, Du Quoin, Carmi, and Cairo. There are significant high outliers in Mt. Vernon, between Mt. Vernon and Nashville, and in Chester. There are significant clusters of high MRFEI at 3200 meters in West Frankfort, Christopher, Carbondale, and Cairo. At 3200 meters there are significant high outliers between Cairo and Metropolis, west of Cairo, east of Anna, Murphysboro, west of Nashville, south of Centralia, and east of Mt. Vernon. There are significant low outliers north of Cairo and west of Murphysboro. There are significant clusters of high MRFEI at 8 kilometers in Mt. Vernon and between Red Bud and Chester. Clusters of low MRFEI are found west of Murphysboro and between Cairo and Metropolis. High outliers can also be observed. There are significant clusters of high MRFEI at 16 kilometers west of Murphysboro, east of Cape Girardeau, and in the northeast part of the study region. There are significant clusters of low MRFEI in rural parts of the region (see figure 19).

The food distance measures are all negatively correlated with all food service area measures (see table 5). The distance to food of any kind is highly correlated with the distance to unhealthy food (0.931; $p<0.001$). The strongest correlation among food service area measures is the correlation between food available within 16 kilometers and the healthy food within that distance (0.973; $p<0.001$). Other strong correlations are between food within 8 kilometers and healthy food within the same distance (0.939; $p<0.001$), and food within 3200 meters and healthy food within 3200 meters (0.919; p<0.001). The largest MRFEI correlation is between the MRFEI at 800 meters and 1600 meters (0.592; p<0.001).

TABLE 5. PEARSON CORRELATIONS BETWEEN FOOD MEASURES

	Food800m	Healthy800m	MRFEI800m	Food1600m	Healthy1600m	MRFEI1600m	Food3200m	Healthy3200m	MRFEI3200m	Food8km	Healthy8km	MRFEI8km	Food16km	Healthy16km	MRFEI16km	UnhealthyDistance	HealthyDistance	AllDistance
Food800m	1.00																	
Healthy800m	0.70	1.00																
MRFEI800m	0.33	0.77	1.00															
Food1600m	0.75	0.49	0.24	1.00														
Healthy1600m	0.60	0.56	0.38	0.82	1.00													
MRFEI1600m	0.29	0.49	0.59	0.36	0.66	1.00												
Food3200m	0.55	0.34	0.17	0.84	0.68	0.32	1.00											
Healthy3200m	0.51	0.37	0.21	0.78	0.74	0.41	0.92	1.00										
MRFEI3200m	0.16	0.24	0.28	0.19	0.28	0.46	0.23	0.40	1.00									
Food8km	0.32	0.19	0.09	0.57	0.48	0.22	0.81	0.76	0.21	1.00								
Healthy8km	0.31	0.21	0.11	0.56	0.55	0.27	0.78	0.82	0.28	0.94	1.00							
MRFEI8km	0.12	0.18	0.19	0.17	0.25	0.34	0.20	0.30	0.57	0.22	0.34	1.00						
Food16km	0.24	0.09	0.03	0.42	0.37	0.13	0.56	0.54	0.12	0.70	0.70	0.15	1.00					
Healthy16km	0.23	0.09	0.04	0.39	0.37	0.15	0.51	0.53	0.14	0.64	0.69	0.19	0.97	1.00				
MRFEI16km	-0.04	0.00	0.01	-0.07	0.00	0.09	-0.10	-0.02	0.28	-0.12	-0.04	0.38	-0.12	-0.01	1.00			
UnhealthyDistance	-0.36	-0.27	-0.21	-0.47	-0.37	-0.28	-0.50	-0.45	-0.25	-0.46	-0.41	-0.19	-0.42	-0.38	0.30	1.00		
HealthyDistance	-0.37	-0.34	-0.29	-0.50	-0.46	-0.44	-0.55	-0.55	-0.50	-0.54	-0.53	-0.56	-0.47	-0.45	-0.15	0.62	1.00	
AllDistance	-0.37	-0.28	-0.24	-0.47	-0.39	-0.37	-0.50	-0.46	-0.39	-0.46	-0.42	-0.41	-0.40	-0.37	0.07	0.93	0.69	1.00

***Highlighted cells are significant (p<0.001)*

D. OBESITY AND THE BUILT ENVIRONMENT

A description of the study area has been noted above, however, one should further notice the differences across urban, urban cluster, and rural areas (see table 6). Obesity rates are lower in urban areas (29.68%; s_x=4.30), compared to urban clusters (33.42%; s_x=3.41) and rural areas (33.60%; s_x=2.17). The number of food stores within any road network service area decreases from urban areas to rural ones. This tends to also be true of healthy food stores, with the exception of the number of healthy stores within 800m. The number of healthy stores within 800 meters is highest in urban clusters, followed by urban areas, then rural ones. MRFEI averages tend to also be highest in urban areas with the exception of the MRFEI at 800 meters and 16 kilometers. The MRFEI at 800 meters is highest for urban clusters, followed by urban areas, and then rural ones. The MRFEI at 16 kilometers is highest on average in rural block groups then urban ones, and then urban clusters. The average distance to unhealthy food or food of any kind is lowest in urban clusters, then urban areas, and then rural ones. The average distance to healthy food is lowest in urban areas, then urban clusters, then rural areas. The average travel time to work increases from urban areas (19.23 minutes; s_x=4.92), to urban clusters (23.05 minutes; s_x=5.75), and lastly rural areas (28.35 minutes; s_x=5.78). The percentage that walk, bike or take public transportation to work also decreases from urban (7.47%; s_x=14.72) to rural (2.34%; s_x=3.04) areas. The average income is highest in rural areas ($22,223; s_x=5778.22), then urban clusters ($21,683; s_x=6995.56), and then urban areas ($21,441; s_x=8820.25). The percentage in poverty is highest in urban areas (25.60%; s_x=22.15) and lowest in rural ones (14.45%; s_x=8.09). The percentage of the population with a college degree or more is over double in urban areas (32.79%; s_x=17.11) what it is in urban clusters (15.32%; s_x=8.65) or rural areas (13.86%; s_x=7.90). Rural areas are whiter than urban clusters or urban areas. Urban clusters have the greatest proportion of women (51.21%; s_x=7.45). Lastly, the median age is highest in the rural block groups (43.39 years; s_x=5.92) and lowest in the urban areas (35.96 years; s_x=8.49).

TABLE 6. VARIABLE CENTRAL TENDENCIES

Mean (s_x)

	Entire Area (n=392)	Urban (n=76)	Urban Cluster (n=176)	Rural (n=140)
Obesity Rate	32.76 (3.57)	29.68 (4.30)	33.42 (3.41)	33.60 (2.17)
Food 800m	0.98 (1.95)	1.38 (2.72)	1.35 (2.05)	0.29 (0.83)
Healthy 800	0.18 (0.47)	0.20 (0.54)	0.27 (0.54)	0.07 (0.28)
MRFEI 800m	6.37 (18.70)	5.31 (17.94)	8.33 (19.05)	4.46 (18.55)
Food 1600m	3.41 (4.92)	6.20 (7.13)	4.47 (4.52)	0.58 (1.16)
Healthy 1600m	0.61 (1.06)	1.30 (1.68)	0.71 (0.90)	0.10 (0.32)
MRFEI 1600m	9.20 (16.71)	12.81 (18.10)	10.92 (14.25)	5.08 (18.03)
Food 3200m	8.30 (10.51)	17.47 (14.57)	10.16 (8.41)	0.98 (1.66)
Healthy 3200m	1.54 (2.08)	3.66 (3.05)	1.70 (1.46)	0.20 (0.44)
MRFEI 3200m	13.25 (18.42)	17.25 (13.41)	14.44 (12.66)	9.57 (25.13)
Food 8k	15.33 (16.95)	36.32 (17.05)	16.75 (13.30)	2.16 (3.20)
Healthy 8k	3.08 (3.69)	8.46 (4.32)	2.87 (1.93)	0.44 (0.77)
MRFEI 8k	16.35 (15.39)	21.89 (6.56)	17.25 (8.33)	12.22 (22.83)
Food 16k	35.82 (31.67)	83.21 (20.03)	32.93 (21.98)	13.73 (16.31)
Healthy 16k	6.88 (6.38)	16.89 (3.91)	5.84 (4.18)	2.74 (3.28)
MRFEI 16k	20.59 (13.57)	20.14 (2.68)	18.71 (6.17)	23.21 (21.33)
All Distance	3422.19 (4151.77)	2026.95 (2946.03)	1855.49 (2338.97)	6149.18 (5025.37)
Unhealthy Distance	3731.51 (4587.54)	2140.54 (3178.40)	1917.42 (2453.25)	6875.74 (5537.48)
Healthy Distance	5703.88 (6031.47)	2949.56 (3488.40)	3030.78 (3457.17)	10559.53 (6616.29)
Average Travel Time	24.20 (6.55)	19.23 (4.92)	23.05 (5.75)	28.35 (5.78)
Walk, Bike, take Public Transportation	3.78 (7.79)	7.47 (14.72)	3.33 (5.26)	2.34 (3.04)
Average Income	21829.13 (6984.95)	21440.88 (8820.25)	21683.45 (6995.56)	22223.02 (5778.22)
% in Poverty	18.50 (14.36)	25.60 (22.15)	18.65 (12.77)	14.45 (8.09)
% College Degree or More	18.19 (12.79)	32.79 (17.11)	15.32 (8.65)	13.86 (7.90)
% White	90.07 (15.51)	82.14 (19.59)	90.61 (14.85)	93.70 (12.01)
% Female	50.13 (7.51)	50.73 (8.38)	51.21 (7.45)	48.45 (6.80)
Median Age	41.31 (7.66)	35.96 (8.49)	41.96 (7.50)	43.39 (5.92)

The most significant predictors of obesity for the entire study area are the percentage of the population with a college degree or more (-0.724; p<0.001), the number of healthy food stores within 8 kilometers (-0.361; p<0.001), the percentage that walk, bike, or take public transportation to work (-0.339; p<0.001), the number of food stores within 8 kilometers (-0.300; p<0.001), and the number of healthy food stores within 16 kilometers (-0.286; p<0.001) (see table 7). In urban areas the most significant predictor is also the percentage with a college degree or more (-0.730; p<0.001), followed by the percentage that walk, bike, or take public transportation (-0.513; p<0.001), the average travel time to work (0.394; p<0.001), the median age (0.390; p<0.001), and the percentage of women (0.379; p<0.001). In urban clusters, the most significant predictor of obesity is the percentage with a college degree or more (-0.618;

$p<0.001$), followed by the average income (-0.388; $p<0.001$), the percentage white (-0.386; $p<0.001$), the median age (-0.347; $p<0.001$), and the percentage in poverty (0.288; $p<0.001$). Lastly, in rural areas the most significant predictor is percentage with a college education or more (-0.605; $p<0.001$), followed by average income (-0.417; $p<0.001$), poverty rate (0.379; $p<0.001$), median age (-0.216; $p<0.01$), and MRFEI at 16 kilometers (0.148; $p<0.1$).

The correlation between obesity and MRFEI is positive for all regions at 800 meters and 1600 meters. It is negative for the entire area and urban areas at 3200 meters, negative for every area except rural at 8 kilometers, and negative in urban areas and urban clusters at 16 kilometers. None of the Pearson correlations between obesity and MRFEI are significant. The only significant food distance measure is the distance to healthy food for the entire study area. In urban clusters and rural areas the relationship between obesity and distance to unhealthy food or food of any kind is negative. Both work commute variables are only significantly correlated with obesity for the entire study area and in urban regions. The relationship between average travel time and obesity is positive in every area, the relationship between the percentage that walk, bike, or take public transportation is negative in every area. Average income is significantly correlated with obesity in the entire study area, urban clusters, and rural areas. The poverty rate is significantly correlated with obesity in urban clusters and rural areas. The percentage with a college degree or more is significantly correlated in every area. The percentage white is significantly correlated with obesity in urban clusters. The percentage of women is significantly correlated in urban areas. Lastly median age is significantly correlated with obesity in urban areas and urban clusters. Interestingly the significant relationship is positive in urban areas but negative in urban clusters.

TABLE 7. PEARSON CORRELATIONS WITH OBESITY (P-VALUE)

	Entire Area (n=392)	Urban (n=76)	Urban Cluster (n=176)	Rural (n=140)
Food 800m	-0.080 (0.112)	-0.132 (0.256)	0.006 (0.939)	0.089 (0.296)
Healthy 800	-0.004 (0.937)	-0.141 (0.225)	0.070 (0.356)	0.073 (0.393)
MRFEI 800m	0.081 (0.111)	0.030 (0.797)	0.089 (0.240)	0.118 (0.165)
Food 1600m	-0.171 (0.001)**	-0.166 (0.151)	0.017 (0.819)	0.078 (0.357)
Healthy 1600m	-0.118 (0.019)	0.001 (0.996)	0.048 (0.523)	0.093 (0.277)
MRFEI 1600m	0.009 (0.865)	0.068 (0.557)	0.024 (0.747)	0.133 (0.118)
Food 3200m	-0.221 (0.000)**	-0.165 (0.154)	0.066 (0.387)	0.002 (0.980)
Healthy 3200m	-0.244 (0.000)**	-0.129 (0.267)	0.062 (0.415)	0.002 (0.980)
MRFEI 3200m	-0.055 (0.280)	-0.125 (0.281)	0.012 (0.877)	0.037 (0.667)
Food 8k	-0.300 (0.000)**	-0.219 (0.057)	0.053 (0.489)	-0.001 (0.989)
Healthy 8k	-0.361 (0.000)**	-0.219 (0.057)	0.041 (0.589)	-0.024 (0.775)
MRFEI 8k	-0.090 (0.075)	-0.142 (0.222)	-0.046 (0.540)	0.040 (0.638)
Food 16k	-0.278 (0.000)**	-0.030 (0.795)	0.187 (0.013)	-0.102 (0.229)
Healthy 16k	-0.286 (0.000)**	0.019 (0.870)	0.190 (0.012)	-0.093 (0.273)
MRFEI 16k	0.047 (0.351)	-0.075 (0.520)	-0.051 (0.502)	0.148 (0.082)
All Distance	0.085 (0.091)	0.184 (0.111)	-0.012 (0.873)	-0.085 (0.319)
Unhealthy Distance	0.096 (0.057)	0.189 (0.103)	-0.011 (0.890)	-0.056 (0.510)
Healthy Distance	0.169 (0.001)**	0.201 (0.081)	0.039 (0.609)	0.104 (0.220)
Average Travel Time	0.267 (0.000)**	0.394 (0.000)**	0.073 (0.333)	0.025 (0.770)
Walk, Bike, take Public Transportation	-0.339 (0.000)**	-0.513 (0.000)**	-0.014 (0.853)	-0.129 (0.130)
Average Income	-0.236 (0.000)**	-0.024 (0.834)	-0.388 (0.000)**	-0.417 (0.000)**
% in Poverty	-0.047 (0.351)	-0.273 (0.017)	0.288 (0.000)**	0.379 (0.000)**
% College Degree or More	0.724 (0.000)**	-0.730 (0.000)**	-0.618 (0.000)**	-0.605 (0.000)**
% White	-0.048 (0.340)	0.049 (0.674)	-0.386 (0.000)**	-0.077 (0.364)
% Female	0.051 (0.315)	0.379 (0.001)**	-0.81 (0.283)	0.040 (0.637)
Median Age	0.066 (0.193)	0.390 (0.001)**	-0.347 (0.000)**	-0.216 (0.011)

The best set of predictor variables for the entire study region is the number of food stores within 800 meters, the MRFEI within 1600 meters, the distance to healthy food, the percentage that walk, bike, or take public transportation to work, average income, percentage with a college degree or more, percentage white, percentage female, and median age[17]. The global OLS regression model has an adjusted r-squared of 0.606 for the entire study area – meaning these variables explain about 61% of the variation in obesity (see table 8). The model is a better fit in the urban block groups of the region (adjusted r-squared 0.705), and less of a fit in the urban clusters (adjusted r-squared 0.476) and rural areas (0.401). The number of food stores within 800 meters nears standard significance levels in the urban block groups (β=-0.133; p=0.103). The

[17] This was found using an iterative OLS and GWR method. Refer to the methods section for details.

variable is not significant in any other geographies. The standardized coefficient is negative (meaning as the number of stores go up obesity goes down) in the entire study area and urban block groups but positive in urban clusters and rural areas. Similarly, the MRFEI within a 1600 meter service area nears standard significance for the entire area (β=0.065; p=0.072), and is significant in urban block groups (β=0.184; p=0.010). The variable is fairly insignificant in urban clusters and rural block groups. Interestingly, the relationship is positive for every group except urban clusters where the significance is very low. This would suggest that as the density of healthy food goes up so does obesity. The distance to healthy food is not significant in any model. The percentage that walk, bike, or take public transportation to work is significant in every group ($p<0.05$), the relationship is negative in every case as well. This means that as the proportion of the population that commutes in this way goes up, obesity goes down. The coefficient for average income is significant in the model for the entire study area (β=0.126; p=0.011) and urban clusters (β=0.180; p=0.035) the variable is not significant in the other areas. In all models, except rural areas, average income is positively related to obesity. The percentage with a bachelor's degree or more is significantly negative in every model. The percentage that are white is also significantly negative in every model. The percentage of women is significant in all models but urban clusters and the relationship is positive in each case. Median age is significantly negative for the entire study area and urban clusters; the variable is not significant in urban or rural models. Significant spatial autocorrelation of the standardized residuals for the OLS model was not observed in any model.[18]

[18] The Global Moran's I test of spatial autocorrelation of the OLS standardized residuals p-value for the entire study region was 0.875, for the urban block groups it was 0.644, for the urban clusters it was 0.405, and for the rural block groups the p-value was 0.859.

TABLE 8. BEST MODEL FOR ENTIRE STUDY AREA, OLS RESULTS FOR ENTIRE AREA AND SUB-AREAS (ß (P-VALUE) | STANDARD ERROR)[19]

	Entire Area (n=392)		Urban (n=76)		Urban Cluster (n=176)		Rural (n=140)	
Intercept	*40.201 (0.000)***	*1.174*	*34.587 (0.000)***	*3.024*	*44.691 (0.000)***	*1.872*	*36.325 (0.000)***	*1.663*
Food 800m	-0.040 (0.284)	0.068	-0.133 (0.103)	0.127	0.012 (0.843)	0.104	0.025 (0.736)	0.192
MRFEI 1600m	0.065 (0.072)*	0.008	0.184 (0.010)**	0.017	-0.006 (0.928)	0.015	0.078 (0.325)	0.009
Distance to Nearest Healthy Food	0.053 (0.170)	0.000	0.079 (0.294)	0.000	0.068 (0.319)	0.000	0.093 (0.235)	0.000
% Walk, Bike or Take Public Transportation to Work	-0.207 (0.000)**	0.017	-0.236 (0.019)**	0.029	-0.113 (0.049)**	0.037	-0.192 (0.005)**	0.049
Average Income	0.126 (0.011)**	0.000	0.124 (0.316)	0.000	0.180 (0.035)**	0.000	-0.068 (0.502)	0.000
% with Bachelors Degree or More	-0.764 (0.000)**	0.012	-0.719 (0.000)**	0.025	-0.611 (0.000)**	0.027	-0.555 (0.000)**	0.025
% White	-0.234 (0.000)**	0.009	-0.241 (0.003)**	0.017	-0.315 (0.000)**	0.014	-0.160 (0.052)*	0.015
% Female	0.087 (0.10)*	0.016	0.131 (0.071)*	0.037	0.049 (0.386)	0.026	0.207 (0.009)**	0.025
Median Age	-0.108 (0.009)**	0.019	0.034 (0.738)	0.052	-0.213 (0.002)**	0.030	-0.037 (0.628)	0.028
R-Squared	*0.615*		*0.740*		*0.503*		*0.440*	
Adjusted r^2	*0.606*		*0.705*		*0.476*		*0.401*	

**p<0.05

*p<0.1

For the entire study area and urban clusters, the localized GWR model performs better. The adjusted r-squared for the entire model is 0.616, for urban areas it is 0.700, for urban clusters it is 0.513, and for rural areas it is 0.393 (see table 9). No significant patterns can be seen in the standardized residuals of the GWR models either (see figures 20, 21, 22, 23).

[19] Using the iterative process described in the methods section, this was the best GWR model. Reported here are the results of an OLS with that model.

TABLE 9. BEST GWR MODEL FOR ENTIRE STUDY AREA, LOCAL RESULTS FOR ENTIRE AREA AND SUB-AREAS (β Mean (s_x) | β Median | β Range)

	Entire Area (n=392)			Urban (n=76)			Urban Cluster (n=176)			Rural (n=140)		
Intercept	*40.801 (1.85)*	*40.619*	*37.251 - 45.018*	*34.642 (0.119)*	34.696	*34.240 - 34.775*	*46.246 (1.467)*	*46.481*	*44.296 - 49.409*	*35.880 (0.270)*	36.138	35.967 - 36.316
Food 800m	-0.044 (0.010)	-.043	-0.064 - -0.029	-0.135 (0.004)	-.134	-0.143 - -0.130	0.021 (0.10)	0.022	0.002 - 0.042	0.021 (0.005)	0.020	0.014 - 0.030
MRFEI 1600m	0.085 (0.020)	.079	0.058 - 0.135	0.185 (0.009)	.182	0.173 - 0.206	-0.004 (0.030)	-0.010	-0.052 - 0.062	0.084 (0.006)	0.086	0.072 - 0.091
Distance to Nearest Healthy Food	0.037 (0.039)	.041	-0.059 - 0.112	0.074 (0.006)	.072	0.067 - 0.091	0.066 (0.027)	0.058	0.023 - 0.114	0.092 (0.001)	0.091	0.091 - 0.094
% Walk, Bike or Take Public Transportation to Work	-0.214 (0.008)	-.210	-0.240 - -0.203	-0.235 (0.005)	-.237	-0.241 - -0.223	-0.128 (0.022)	-0.134	-0.165 - -0.091	-0.190 (0.002)	-0.191	-0.193 - -0.187
Average Income	0.142 (0.033)	.147	0.088 - 0.208	0.126 (0.001)	.125	0.125 - 0.127	0.191 (0.031)	0.207	0.140 - 0.242	-0.066 (0.000)	-0.066	-0.080 - -0.058
% with Bachelors Degree or More	-0.779 (0.037)	-.773	-0.861 - -0.724	-0.721 (0.002)	-.721	-0.725 - -0.719	-0.658 (0.055)	-0.668	-0.751 - -0.562	-0.553 (0.003)	-0.553	-0.558 - -0.546
% White	-0.249 (0.063)	-.230	-0.391 - -0.167	-0.241 (0.001)	-.241	-0.242 - -0.239	-0.379 (0.109)	-0.411	-0.620 - -0.224	-0.155 (0.008)	-0.155	-0.168 - -0.140
% Female	0.066 (0.040)	.056	-0.009 - 0.140	0.130 (0.002)	.129	0.128 - 0.138	0.035 (0.016)	0.035	0.006 - 0.067	0.208 (0.008)	0.207	0.192 - 0.222
Median Age	-0.106 (0.021)	-.099	-0.154 - -0.068	0.032 (0.000)	.032	0.032 - 0.035	-0.207 (0.054)	-0.205	-0.281 - -0.088	-0.034 (0.006)	-0.034	-0.044 - -0.026
R-Squared		*0.648*			*0.744*			*0.581*			*0.446*	
Adjusted R^2		*0.616*			*0.700*			*0.513*			*0.393*	

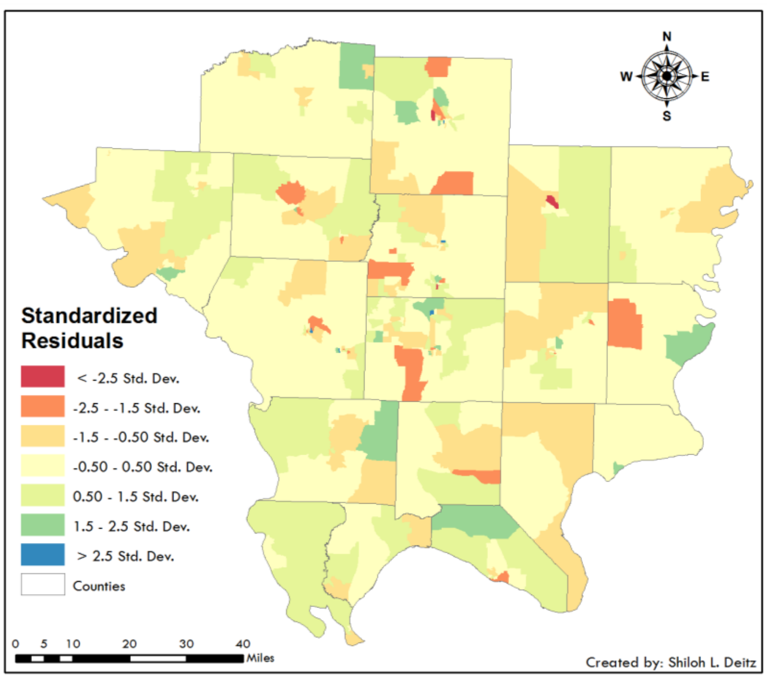

FIGURE 20. STANDARDIZED RESIDUALS FOR GWR MODEL, SOUTHERN IL

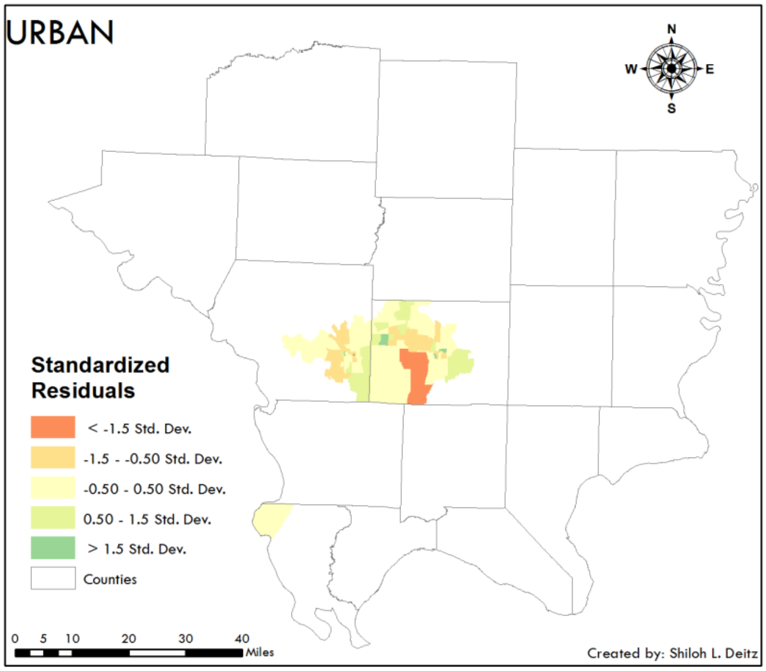

FIGURE 21. STANDARDIZED RESIDUALS FOR GWR MODEL, URBAN

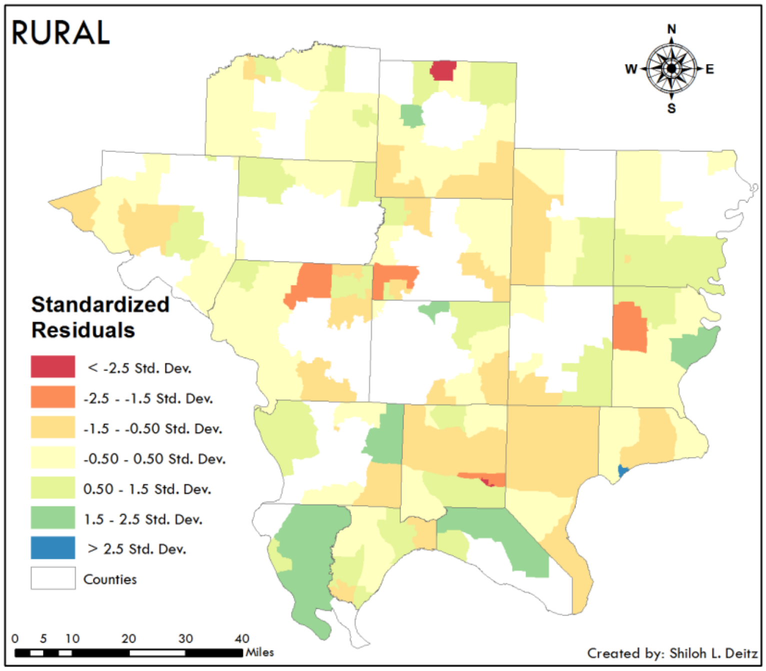

FIGURE 22. STANDARDIZED RESIDUALS FOR GWR MODEL, RURAL

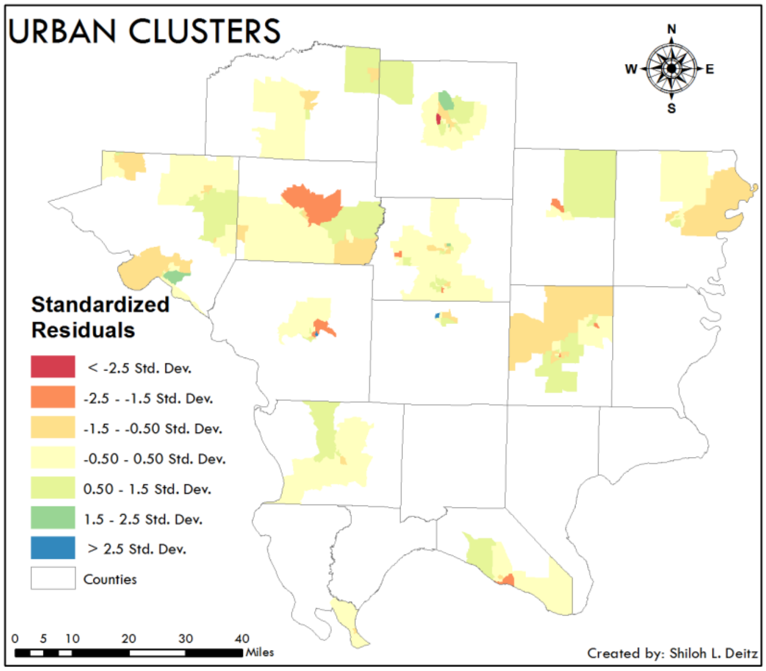

FIGURE 23. STANDARDIZED RESIDUALS FOR GWR MODEL, URBAN CLUSTERS

To further understand how the built environment and obesity were related, the best model from each subset area (rural, urban cluster, and urban) was found[20]. The best performing rural model had an adjusted r-squared of 0.439 for an unweighted OLS and 0.447 for GWR (see table 10). The variables included were: MRFEI at 800 meters, food within 8 kilometers, MRFEI at 16 kilometers, distance to nearest healthy food, percentage that walk, bike or take public transportation to work, percentage in poverty, percentage with a bachelor's degree or more, percent white, percent women, and median age. The coefficients for MRFEI at 800 meters and food within 8 kilometers are not significant. The MRFEI at 16 kilometers (about 10 miles) is significant and positive (β=0.174; p=0.023). The distance to healthy food is also significant and positive (β=0.187; p=0.038). The percentage who commute on foot, by bike, or with public

[20] Again, refer to the methods section. With rural (n=140), urban cluster (n=176), and urban (n=76) block groups only all models of 1 to 26 variables were run in OLS. Models were eliminated if they exhibited multicollinearity (VIF>7.5) or spatially autocorrelated residuals (p<0.05). Then each passing OLS was run in GWR and the best model was chosen based on adjusted r-squared and theoretical importance.

transportation has a significant negative effect (β=-0.172; p=0.011). The percentage in poverty is significantly positively correlated, the coefficient for the percentage with a college degree or more is significant and negative, and the percentage female is positive and nears standard significance levels (β=0.139; p=0.081). The percentage white and median age are not significant. The standardized residuals in the OLS model did not exhibit significant spatial autocorrelation (p=0.739) and appear to also be random in the GWR model (see figure 24).

TABLE 10. BEST REGIONAL MODEL, RURAL (N=140)

	OLS Results ß (p-value) | Standard Error		**Local GWR Results** ß Mean (s)	ß Median	ß Range
Intercept	*35.050 (0.000)***	*1.655*	*34.175 (0.902)*	*34.091*	*33.221 - 36.734*
MRFEI 800m	0.107 (0.137)	0.008	0.107 (0.009)	0.109	0.084 - 0.119
Food 8km	0.079 (0.309)	0.053	0.083 (0.014)	0.085	0.051 - 0.105
MRFEI 16km	0.174 (0.023)**	0.008	0.195 (0.095)	0.224	-0.001 - 0.329
Distance to Nearest Healthy Food	0.187 (0.038)**	0.000	0.188 (0.054)	0.201	0.064 - 0.265
% Walk, Bike or Take Public Transportation to Work	-0.172 (0.011)**	0.047	-0.166 (0.010)	-0.170	-0.181 - -0.146
% in Poverty	0.155 (0.037)**	0.020	0.130 (0.075)	0.155	-0.012 - 0.230
% with Bachelors Degree or More	-0.508 (0.000)**	0.020	-0.492 (0.063)	-0.482	-0.609 - -0.405
% White	-0.098 (0.231)	0.015	-0.057 (0.063)	-0.066	-0.156 - 0.110
% Female	0.139 (0.081)*	0.025	0.121 (0.035)	0.132	0.021 - 0.155
Median Age	-0.097 (0.182)	0.027	-0.090 (0.033)	-0.085	-0.152 - -0.041
R-Squared	*0.480*			*0.537*	
Adjusted R^2	*0.439*			*0.447*	

**p<0.05
*p<0.1

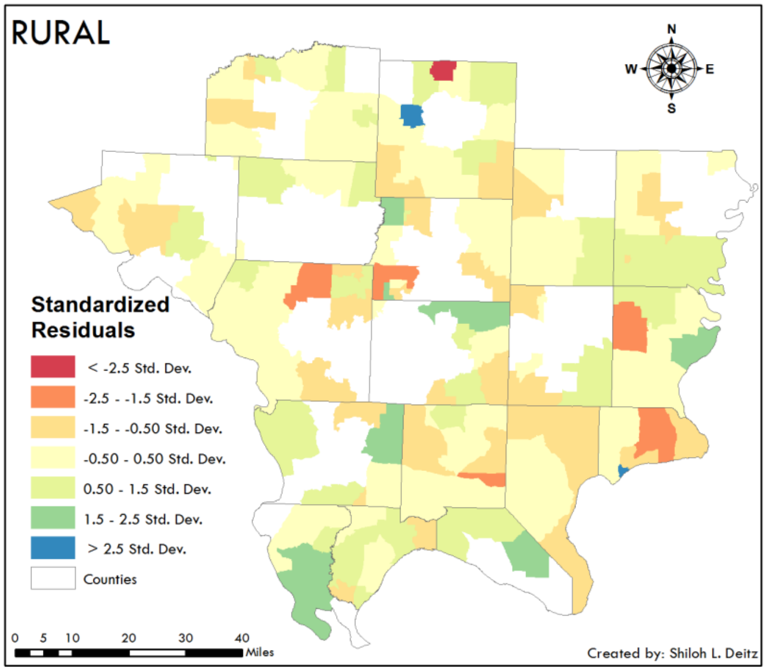

FIGURE 24. STANDARDIZED RESIDUALS FOR BEST REGIONAL MODEL, RURAL

The most significant model for the urban cluster block groups includes: the number of food stores within 16 kilometers, the MRFEI within the 16 kilometer service area, the distance to healthy food, the percentage that walk, bike, or take public transportation to work, the average travel time to work, the percentage with a college degree or more, the percentage white, the percentage female, and median age. The adjusted r-squared for the global OLS model is 0.488 and the r-squared for the GWR model is 0.539 (see table 11). The number of food stores within 16 kilometers exerts a significant positive effect on the model (β=0.138; p=0.021). The MRFEI at 16 kilometers is negative but insignificant. The percentage that walk, bike, or take public transportation to work nears standard significance levels (β=-0.092; p=0.105). The average travel time to work is insignificant. The percentage with a college degree or more are both negative and significant in the model. The percentage female is not significant. Finally, median age is significant and negative (β=-0.144; p=0.016). The OLS standardized residuals do not exhibit spatial clustering (p=0.120) and the GWR standardized residuals do not appear to follow any spatial pattern either (see figure 25).

TABLE 11. BEST REGIONAL MODEL, URBAN CLUSTER (N=176)

	OLS Results ß (p-value) | Standard Error		Local GWR Results		
			ß Mean (s_x)	ß Median	ß Range
Intercept	*42.642 (0.000)***	*2.101*	*44.837 (1.126)*	*44.498*	*42.294 - 46.876*
Food 16km	0.138 (0.021)**	0.009	0.157 (0.036)	0.175	0.093 - 0.200)
MRFEI 16km	-0.050 (0.375)	0.031	-0.028 (0.069)	-0.040	-0.117 - 0.160
Distance to Nearest Food (any kind)	0.117 (0.059)*	0.000	0.111 (0.019)	0.111	0.078 - 0.156
% Walk, Bike or Take Public Transportation to Work	-0.092 (0.105)	0.037	-0.121 (0.037)	-0.115	-0.198 - -0.055
Average Travel Time to Work	0.072 (0.245)	0.037	0.086 (0.056)	0.105	0.005 - 0.170
% with Bachelors Degree or More	-0.524 (0.000)**	0.023	-0.560 (0.046)	-0.563	-0.695 - -0.483
% White	-0.291 (0.000)**	0.013	-0.374 (0.180)	-0.409	-0.657 - -0.113
% Female	0.061 (0.275)	0.026	0.030 (0.037)	0.028	-0.061 (0.111)
Median Age	-0.144 (0.016)**	0.027	-0.145 (0.086)	-0.122	-0.278 - -0.019
R-Squared	*0.514*			*0.612*	
Adjusted R^2	*0.488*			*0.539*	
**p<0.05					
*p<0.1					

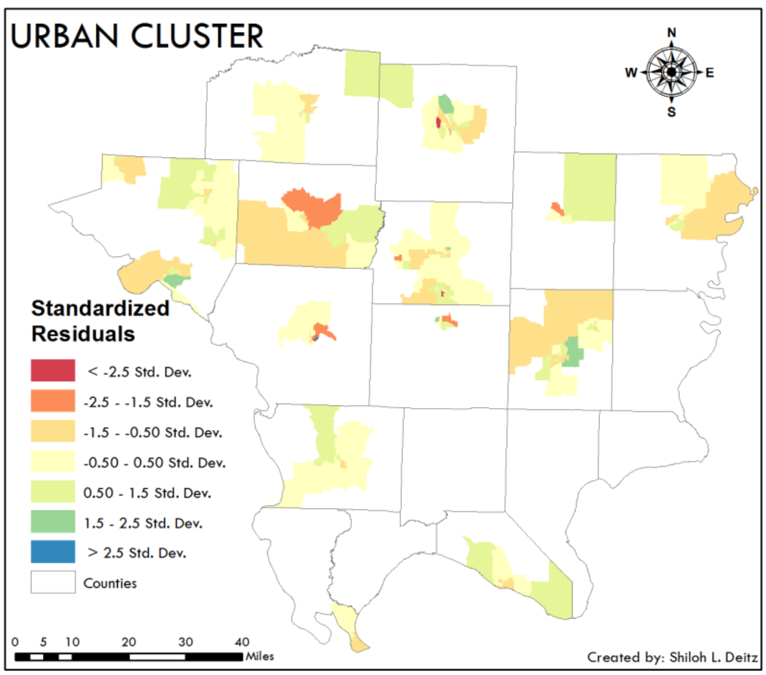

FIGURE 25. STANDARDIZED RESIDUALS FOR BEST REGIONAL MODEL, URBAN CLUSTER

The best model for the urban block groups has a global adjusted r-squared of 0.748 and geographically weighted regression adjusted r-squared or 0.736 (see table 12). The variables in this model are the MRFEI within 800 meters, the number of food stores within 1600 meters, the MRFEI at 1600 meters, the MRFEI at 8 kilometers, the distance to the nearest food of any kind, the percentage that walk, bike, or take public transportation to work, the percentage in poverty, the percentage with a college degree or more, the percentage white, and the percentage female. The MRFEI within an 800 meter service area has a significant negative coefficient (β=-0.206; p=0.004), the number of food stores within 1600 meters is also significant and negative (β=-0.157; p=0.069). The MRFEI for the 1600 meter service area is significant and positive (β=0.300; p<0.001), and the MRFEI for the 8 kilometer service area is not significant. The distance to food of any kind is positive and nears standard significance levels (β=0.107; p=0.150). The percentage the walk, bike, or take public transportation has a negative coefficient and is significant (β=-0.181; p=0.071). The percentage in poverty, with a college degree or more, and white all have significant negative coefficients. The percentage of women has a significant positive coefficient. The standardized residuals are not spatially autocorrelated in an OLS model (p=0.107). The standardized residuals for the GWR model do not exhibit any spatial pattern either (see figure 26).

TABLE 12. BEST REGIONAL MODEL, URBAN (N=76)

	OLS Results		Local GWR Results		
	ß (p-value)	**Standard Error**	**ß Mean (SD)**	**ß Median**	**ß Range**
Intercept	*36.481 (0.000**)*	*2.644*	*36.527 (0.221)*	*36.556*	*34.899 - 36.721*
MRFEI 800m	-0.206 (0.004)**	0.017	-0.195 (0.004)	-0.196	-0.198 - -0.173
Food 1600m	-0.157 (0.069)*	0.051	-0.172 (0.027)	-0.163	-0.267 - -0.138
MRFEI 1600m	0.300 (0.000)**	0.018	0.309 (0.034)	0.296	0.271 - 0.427
MRFEI 8km	0.083 (0.227)	0.045	0.079 (0.006)	0.078	0.072 - 0.120
Distance to Nearest Food (any kind)	0.107 (0.150)	0.000	0.125 (0.025)	0.114	0.097 - 0.197
% Walk, Bike or Take Public Transportation to Work	-0.181 (0.071)*	0.029	-0.171 (0.014)	-0.175	-0.192 - -0.141
% in Poverty	-0.200 (0.041)**	0.019	-0.198 (0.013)	-0.203	-0.212 - -0.140
% with Bachelors Degree or More	-0.737 (0.000)**	0.017	-0.742 (0.015)	-0.736	-0.793 - -0.725
% White	-0.285 (0.000)**	0.017	-0.283 (0.162)	-0.285	-0.292 - -0.246
% Female	0.162 (0.020)**	0.035	0.162 (0.003)	0.161	0.155 - 0.181
R-Squared	*0.782*			*0.789*	
Adjusted R^2	*0.748*			*0.736*	

**p<0.05; *p<0.1

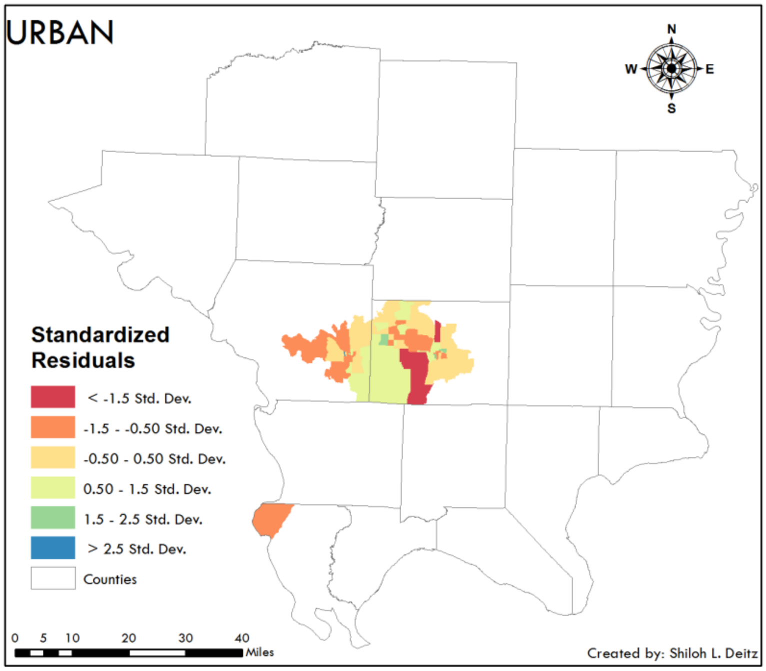

FIGURE 26. STANDARDIZED RESIDUALS FOR BEST REGIONAL MODEL, URBAN

V. DISCUSSION

In this section I will summarize the answers to the research questions and discuss broader contributions that this research may have.

A. HOW CAN SMALL AREA OBESITY ESTIMATES BE RELIABLY INTERPOLATED FROM DATA SETS AT LOWER SPATIAL RESOLUTIONS?

The steps involved in choosing the estimation model – theoretical decisions, exploratory data analysis, factor reduction through backward elimination OLS regression, all subsets OLS to remove multicollinearity and variable insignificance, geographically weighted regression to capture spatial patterns, analysis of accuracy through re-aggregation of predicted values and rRMSE calculation, and final model choice based on rRMSE values – could be precisely followed and reduced the error introduced into the process at every step. Particularly, the use of all subsets regression rather than solely relying on a stepwise removal technique, revealed patterns that would have otherwise been lost. In fact, the final model chosen looked nothing like the model that would have been chosen had stepwise regression alone been relied upon. Further, the use of rRMSE as a model selection criterion rather than just a test to check accuracy, improved the results and was a useful criterion for comparing models. In my initial rankings of the models the one that was chosen as having the lowest rRMSE ranked number twelve out of fourteen. While the model had the lowest root mean square error it would not have been chosen based on the GWR results only.

This analysis procedure also revealed the importance of accounting for geography in small area estimation. There is significant spatial autocorrelation of obesity in the United States ($p<0.001$). The model chosen was a *local* fit that captured the localized correlations between various socio-demographic variables and obesity. While the prediction model selected was a local fit, the methods used to find it could be replicated in other study areas. These methods could also be useful for small area estimation of other health indicators. Further, this process could be implemented by anyone with a basic knowledge of multivariate regression and spatial analysis.

B. HOW DOES OBESITY VARY GEOGRAPHICALLY IN SOUTHERN ILLINOIS AT THE BLOCK GROUP LEVEL?

It is well established in the literature and the analysis procedure described above confirmed that obesity is a geographically varying phenomenon and southern Illinois is not exceptional in that way. Obesity varies geographically in the region with significant spatial clustering in certain areas and the presence of significant outliers. The variance in obesity can be explained by multiple environmental factors such as culture, the built environment, the natural environment, political conditions, and economic conditions. Specifically, there are strong correlations between obesity and level of education, income, age, race, and population density. These factors interact together and with measures of the built environment such as walkability or food to explain a large part of the variation in obesity. The interactions and varying influences of these factors suggest something about culture – social influences meet built environment infrastructure in specific ways across different geographies.

C. HOW CAN WALKABILITY BE QUANTITATIVELY MEASURED IN NON-URBAN AREAS AND HOW WALKABLE IS SOUTHERN ILLINOIS?

Numerous measures have been proposed for quantifying walkability, however, these methods are expensive, time consuming, and outside urban areas the data is often unavailable. For these reasons it is necessary to consider other methods of quantifying walkability. In this study a known reliable walkability measure was used to test the effectiveness of other measures.

The known measure, *WalkScore*, is not available for all block groups in the study area and so was not used. The study found that the *WalkScore* and food density are highly correlated and that commuting behavior can also be a good indicator of overall walkability. It makes perfect logical sense that an increase in the number of food stores within 800 or 1600 meters of a person's residence would suggest a walkable environment and food location data is much easier to come by than other data such as intersection density, or the level of commercial/residential mixing.

Using these quantitative measures, the walkability of the study region was analyzed and significant spatial patterns were found. There are neighborhoods with high numbers of food stores within 800 meters in Benton, Carbondale, Carmi, Harrisburg, Johnston City, Metropolis, Mount Vernon, Murphysboro and Pinckneyville. There are higher proportions of the population walking, biking, or taking public transportation in Carbondale and Murphysboro. The average commute times are also lower in neighborhoods of Carbondale as well as Mount Vernon, Chester, and West Frankfort. However, just because the area looks to be walkable quantitatively does not necessarily mean that people perceive their environment as walkable. First, the neighborhoods that appear to have high walkability based on the number of food stores within 800 or 1600 meters are also clustered along the major interstate and state highway systems. Specifically, there are many clusters of neighborhoods with significantly high numbers of food stores along interstate highway 57, especially where a state highway intersects with interstate 57. Crossing the highway to walk to destinations is not a common behavior and thus, though the distance is small, the actual walkable infrastructure is bad. The same is true of the Carbondale-Marion urban areas – the variables suggest that these are *highly* walkable areas – particularly in Carbondale. However, a recent study conducted for the Southern Illinois Metropolitan Planning Organization found that although many amenities are within close distance the infrastructure that might encourage walking or biking is not present (Lochmueller Group and Alta 2014). For example, an inventory of sidewalks and bike paths found that many of the sidewalks are in poor condition and bike paths are sparse (see figures 27 and 28; Lochmueller Group 2014).

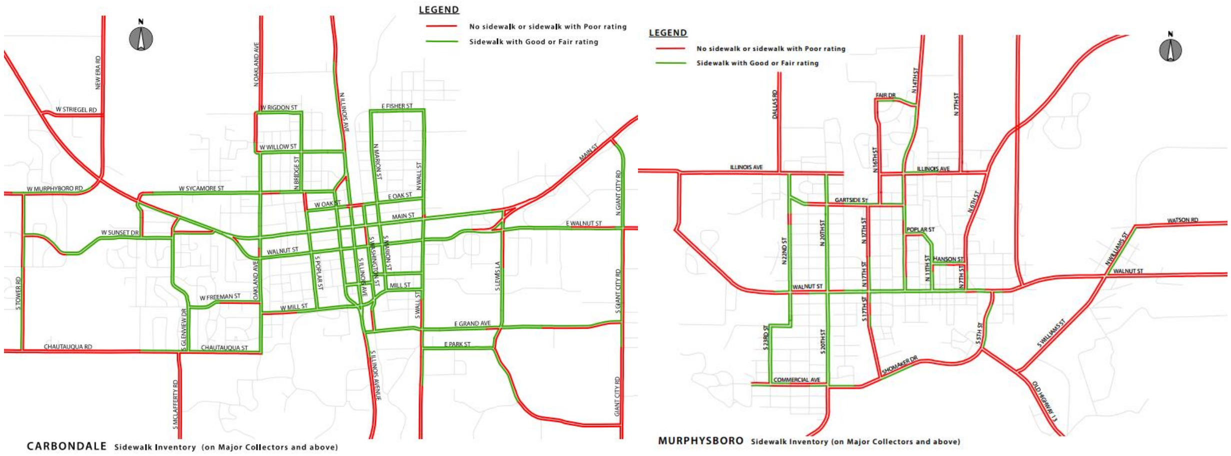

FIGURE 27. SIMPO SIDEWALK INVENTORY FOR CARBONDALE AND MURPHYSBORO

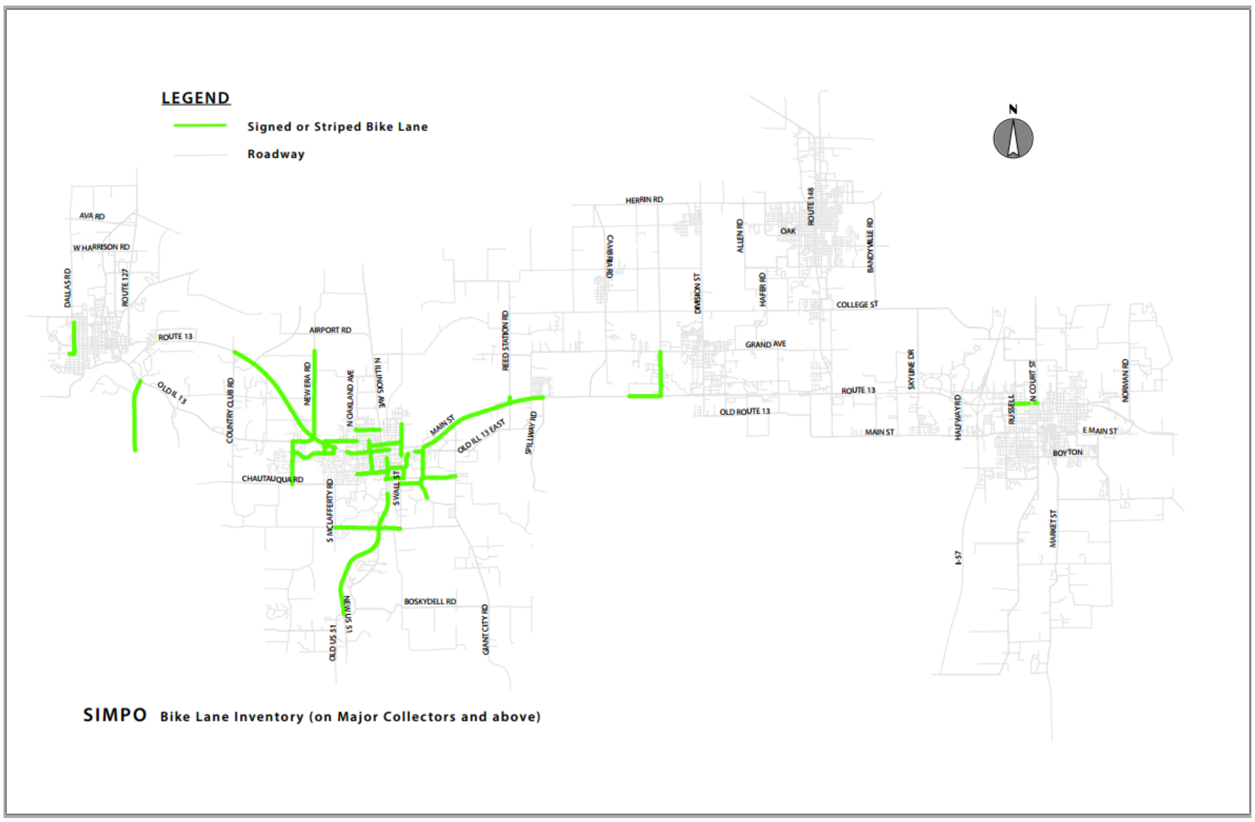

FIGURE 28. SIMPO BIKE LANE INVENTORY FOR CARBONDALE-MARION URBAN AREA

D. HOW CAN THE FOOD ENVIRONMENT BE QUANTIFIED IN NON-URBAN AREAS AND WHAT IS THE FOOD ENVIRONMENT IN SOUTHERN ILLINOIS?

Numerous measures have been proposed for quantifying the food environment. This study, with the inclusion of a wide variety of measures, revealed that quantifying food is a complex task. The first thing that was verified in this study was that the majority of food sources – especially in small towns – provide unhealthy food options (79% of the stores were classified as unhealthy). Unhealthy food is most likely to be nearest to a resident's home, and there are likely to be more options. In urban clusters (small towns) the distance to unhealthy food is lower even than it is in urban areas where the distance to healthy food is at its lowest. The density of unhealthy food options is in almost every case larger than the density of healthy food option.

There are two major concerns or lessons to be learned from this study. The first was dealt with within this project while the second was not. First, the messiness of food environments within various cultures and geographies makes it overly restrictive to impose binary measures of food access. Binary measures assume that the point at which the distance to food or the density of healthy food begins to negatively affect health or quality of life is known and can be generalized across multiple study areas. The results of this study suggest that the distance and density that makes a difference depends on many factors including sociodemographics (which may suggest culture) and geography. The contradictory results and disproved hypothesis of previous studies are further evidence of this fact. This study should encourage an open-minded approach towards measures of the food environment and further investigation of how the measures used in this study vary across different regions of the United States or in other countries.

The second issue that is implicit in this kind of quantitative study is how various food sources are used and what they offer. In this case a quantitative classification of 573 food sources as healthy or unhealthy required that assumptions were made based on what kinds of food stores or chains typically provided. It could be safely assumed that a McDonalds is unhealthy and a Kroger would provide healthy options, but there were a great many local grocers, convenience stores, dollar stores, and gas stations that fill a gray area. In this study, resources of both time and money were not available to visit each location in question and generalizations were made about chain dollar stores, gas stations, and convenience stores – local grocers were called or researched online. However, past research has suggested that in rural areas particularly, it is not uncommon for a gas station or dollar store to carry a full line of healthy groceries including fresh fruits and vegetables (Gustafson 2012). Further, these quantitative methods do not capture the healthy food that might be available through farms, trading, or sporadic trips to distant grocery stores to buy in bulk (McPhail et al. 2013; Morton and Blanchard 2007; Scarpello et al. 2009; Sharkey et al. 2010; Yousefian et al. 2011). Qualitative analysis of the food environment in southern Illinois would be a wonderful complement to this study and might reveal more about how certain food sources are used.

Finally, a recent study by Chen and Clark (2016) has suggested temporal considerations to food access. Their research found that the hours a store is open impact access and disadvantage certain groups. This is a relevant finding that may provide a more robust understand of the food environment but one that was not investigated in this study. Future research should take temporal access into account.

E. WHAT IS THE CORRELATION BETWEEN THE BUILT ENVIRONMENT AND OBESITY IN SOUTHERN ILLINOIS, AFTER CONTROLLING FOR SOCIO-DEMOGRAPHIC COVARIATES?

The final stage of this study aimed to investigate the relationship between multiple measures of the food environment, measures of walkability, and obesity. For the entire study area, the most significant correlations were the percentage with a college degree or more, the number of healthy food outlets within 8 kilometers, and the percentage that walk, bike, or take public transportation to work. The most significant bivariate relationships with obesity in urban areas was also the percentage with a college degree or more, followed by the percentage that walk, bike, or take public transportation to work, and the average commute time to work. In urban clusters the most significant correlations were with the percentage with a college degree or more, the average income, and the percentage white (all sociodemographic factors). In rural areas the most significant relationships were with the percentage with a college degree or more, average income, and percentage in poverty (again, all sociodemographic). Notably, for the entire study area and urban blocks in the study area the most significant relationships included measures of walkability and food measures but in urban clusters and rural areas the most significant relationships were all sociodemographic. This might suggest that culture is a more formidable barrier to healthy living in rural and small town areas. Meaning any food or walkability based interventions should be coupled with thorough educational campaigns

Focusing on food measures, the most significant Pearson correlation for the entire study area (n=392) was the number of healthy food stores within 8 kilometers, followed by the number of food stores of any kind within 8 kilometers, and the number of food stores within 16 kilometers. All of these measures are negative suggesting an increase in either healthy food stores or stores of any kind within about five or ten miles would decrease the prevalence of obesity. In urban areas the number of healthy food stores and stores of any kind within 8 kilometers have the most significant Pearson correlations with obesity and near standard significance levels. In urban clusters the most significant food correlation with obesity is between the number of healthy food stores within 16 kilometers and obesity, followed by the number of food stores within 16 kilometers. Interestingly, these measures are both positive suggesting a decrease in food stores would decrease obesity. In rural areas the correlation between obesity and MRFEI at 16 kilometers nears standard significance levels. Overall, these results suggest that the introduction of food stores within about 5 miles to 10 miles of resident homes would decrease obesity; however, in urban clusters and rural areas the relationship is the opposite – perhaps suggesting a misclassification of food stores.

Contrary to the results of other studies, there is no significant bivariate relationship between MRFEI and obesity. Also, the insignificant relationship between the two variables is positive at 800 meters and 1600 meters but negative at service area distance of 3200 meters and greater. As mentioned previously, these results could in part be attributed to the classification of stores. Research has shown that people outside urban areas tend to do more shopping at non-traditional locations such as convenience stores and dollar stores (Gustafson et al. 2012). Again, qualitative analysis of this phenomena might shed more light on these trends.

Regarding walkability, the percentage that walk, bike, or take public transportation to work, is consistently negatively correlated with obesity but only significantly so for the entire study area and the urban neighborhoods. The average travel time to work is consistently positive but only significantly so for the entire study area and urban neighborhoods as well. The number of food stores within 1600 meters (the most significant food based measure of walkability) exhibits a negative relationship with obesity for the entire study area and urban neighborhoods but a positive one for urban clusters and rural areas – the Pearson correlation is only significant for the entire study area. The Pearson correlation between obesity and the number of food stores within 800 meters is negative in the entire study area and urban block groups – this measure nears standard significance levels in the entire study area. These results overall point out that an increase in walkability – or walking behavior – would reduce obesity.

The most significant regression model for the entire study area includes food within 800 meters but the coefficient is not significant in any model. It also includes MRFEI at 1600 meters. The coefficient for this variable is only significant for the entire study area and urban neighborhoods and exerts a positive influence. The distance to healthy food is insignificant. The percentage that walk, bike, or take public transportation is negative and significant in every model but the largest coefficient is in urban neighborhoods.

The only model in which MRFEI exerts a significant negative relationship is the MRFEI at 800 meters in the urban model. The MRFEI at 1600 meters and 8 kilometers is also included in the urban model and exerts a positive influence. Past research suggests that the number of food stores should have a negative relationship with obesity due to its suggestion of walkability. This relationship is significantly negative for the number of food stores within 1600 meters in the urban neighborhood model, but in all other models it is either insignificant or positive or both. The coefficient for the percentage that walk, bike, or take public transportation to work is negative in every model it is included in. The coefficient for the distance to food (healthy or any) is positive in every case but not always significant.

Intervention-wise the model for the <u>entire study area</u> suggests that *decreasing the density of healthy food to all food*, and *increasing the percentage that walk, bike, or take public transportation to work* could decrease obesity. The <u>rural</u> model suggests that *decreasing the* *density of healthy food to all food within 16 kilometers*[21], *decreasing the distance to healthy food*, and *increasing the percentage that walk, bike, or take public transportation to work would decrease obesity*. The urban cluster model suggests that a decrease in obesity could be achieved with an *increase in the number of food stores within 16 kilometers*, and a *decrease of the distance to food of any kind*. Finally, the urban model suggests that *increasing the density of healthy food within 800 meters*, *increasing the number of food sources within 1600 meters, decreasing the density of healthy food to all food within 1600 meters*, and *increasing the percentage that walk, bike, or take public transportation to work* would decrease obesity.

The results of a qualitative study investigating perceptions of the environment for healthy eating and exercising in a rural northeastern community found that individuals perceive and use infrastructure in unique ways (Maley et al. 2010). In this study they find having unhealthy food in abundance at social gatherings is seen as a sign of good hospitality, the roads were seen as unsafe for exercise, and the weather was also perceived as a barrier to healthy living. This was despite the presence of a grocery store, walking trails, parks, and a community pool. This qualitative research supports the idea that specific sociocultural factors come together to promote obesity even if the infrastructure is in place that would be thought to reduce obesity. As noted above, other studies have found that the introduction of healthy infrastructure reduced obesity. Taken together, this would suggest that increases in walkability and healthy food options alongside culturally embedded education campaigns might be the most thorough obesity intervention.

[21] It makes no sense theoretically that decreasing the density of healthy food would decrease obesity. These models suggest that, analytically, MRFEI measures might not work in non-urban areas or particular attention should be paid to the classification of food stores and how stores are used in non-urban areas.

The regression models examining the relationship between the built environment and obesity are messy. The hypothesis that there is some relationship between the built environment and obesity is confirmed yet in many cases the relationship is not what was expected. Specifically, the MRFEI measure functioned in a surprising and erratic way in all of the models – perhaps due to improper classifications and generalizations about healthy and unhealthy food stores. The model with results that come closest to previous studies is for the urban neighborhoods of southern Illinois. This makes sense because the majority of quantitative studies on these topics have been conducted in urban areas – often ones larger than any present in this study area. These results do not suggest that these measures should be discarded but do suggest that investigative and flexible approaches should be taken.

VI. CONCLUSION

There are a few major takeaways from these analyses. First, localized geographically weighted small-area estimates provide more accuracy than those that are neither localized nor geographically weighted. Second, obesity is a serious health issue in the United States with multiple factors contributing to its prevalence. These factors vary across different communities for reasons both known and unknown. Certainly there is some room for further study and these analyses suggest that future research should continue to look at the issue with an open mind and a large toolkit of potential measures. In the final analyses of this study that looked at the relationship between the built environment and obesity in southern Illinois, few variables behaved in expected ways and new insights were garnered through an open-minded approach. Further, there appear to be both major regional differences in how obesity and the built environment interact and differences based on population density. This research suggests that studying urban, rural, and urban clusters together in one region may not produce the best results. Finally, the impact of sociodemographic factors on the models used in this study should not be discounted both for future analysis and as policy changes are considered. There is no doubt across multiple studies that there are striking correlations between obesity and race, income, and educational attainment. In most cases populations that already fill the most marginalized positions in society – poor minorities with low educational attainment – are most likely to be obese, and suffer from poor health in other ways. Any discussion of healthy communities must take these social factors into account and target disadvantaged communities for intervention.

In this study region, there is widespread public support for increasing downtown renewal, walkability, and an increase in healthy food options. In response to a question on the Paul Simon Public Policy Institute's fall 2015 Southern Illinois Poll, "how important [they] would say it is to improve the downtown of [their] community," 84.3% said it was either very or somewhat important. Local products – including food – are also important to southern Illinoisans. Over 3 in 4 (76.8%) southern Illinoisans claimed that they are more likely to purchase a product if it were local suggesting that an increase in local food products could be a good avenue for increasing healthy food options. Increases in local food have been shown to stimulate local economies in rural and small town areas as well (Hewitt 2009). In this way, the dual goal of economic revitalization and healthy communities could be advanced with the same program. Regarding walkability, although very few people seem to commute by modes of transportation other than cars, there is public support for not needing to rely on cars. In a statewide Simon Poll conducted by the Paul Simon Institute, about 1 in 4 downstate residents (24.3%) claimed they prefer to live in a place with transportation options while 63.0% prefer a place where they rely on cars, 6.0% claim it depends, and 6.7% don't know. While this isn't the same public support for transportation options observed in the city of Chicago (73.0%) it does suggest some public will for walkability. According to a spring 2014 poll of Jackson and Williamson Counties also conducted by the Paul Simon Public Policy Institute, 10.1% of the population claims they would support an increase in property taxes to support alternative transportation initiatives and 29.9% say they would support an increase in local sales tax for such initiatives.

Concretely, I recommend the following for southern Illinois:

- Dense urban cores, however small.
- Improved paths and sidewalks for cyclist and pedestrian safety.
- Emphasis on local food development. Particularly in areas with low food access.
- Highlight and preserve southern Illinois natural beauty and farmland.

These recommendations for the towns of southern Illinois would improve the health – and probably happiness – of residents, improve the natural environment, and provide economic stimulus.

There is very little harm that could come from increasing walkability, opportunities for physical activity, and healthy food options within an area. Perhaps this is why these concepts are so widely studied and prescribed. Even if there were no relationship between the built environment and obesity, why not redesign our built environments to be more health promoting? And perhaps as communities push for such changes, the positive relationship between healthy people and healthy environments will become clearer.

REFERENCES

Adu-Prah, Samuel, and Tonny Oyana. 2015. "Enabling Healthy Living: Spatiotemporal Patterns of Prevalence of Overweight and Obesity among Youths in the United States." International Journal of Applied Geospatial Research 6(2): 98-116.

Alviola IV, Pedro A., Rodolfo M. Nayga Jr., and Michael Thomsen. 2013. "Food Deserts and Childhood Obesity." Applied Economic Perspectives and Policy 35(1): 106-124.

Alviola IV, Pedro A., Rodolfo M. Nayga Jr., Michael R. Thomsen, and Zhongyi Wang. 2013. "Determinants of Food Deserts." *American Journal of Agricultural Economics* 95(5): 1259-1265.

Baile, Jose I., and Maria J. Gonzalez-Calderon. 2014. "Precision del Indice de Masa Corporal, Obtenido a Partir de Dato do Peso y Altura Autoinforados en una Muestra Infantil Espanola." Nutricion Hospitalaria 29(4): 829-831.

Befort, Christie A., Niaman Nazir, and Michael G. Perry. 2012. "Prevalence of Obesity among Adults from Rural and Urban Areas of the United States: Findings from NHANES (2005-2008)." Journal of Rural Health 28(4): 392-397.

Bell, Jennifer Denise. 2014. "BMI and POPs: Association Between Persistent Organic Pollutants and Increasing Obesity Prevalence in the United States." University of Georgia Thesis, M.S.

Bell, W.R., G.S. Datta, and M. Ghosh. 2013. "Benchmarking Small Area Estimators." Biometrika 100(1): 189-202.

Bellinger, William K., and Jue Wang. 2011. "Poverty, Place or Race: Causes of the Retail Gap in Smaller U.S. Cities." The Review of Black Political Economy 38: 253-270.

Berke, Ethan M., Thomas D. Koepsell, Anne Vernez Moudon, Richard E. Hoskin, and Eric B. Larson. 2007. "Association of the Built Environment with Physical Activity and Obesity in Older Persons." American Journal of Public Health 97(3): 486-492.

Bitter, Christopher, Gordon F. Mulligan, and Sandy Dall'erba. 2007. "Incorporating Spatial Variation in Housing Attribute Prices: A Comparison of Geographically Weighted Regression and the Spatial Expansion Method." Journal of Geographical Systems 9: 7-27.

Boone-Heinonen, Janne, Ana V. Diez-Roux, David C. Goff, Catherine M. Loria, Catarina I. Kiefe, Barry M. Popkin, and Penny Gordon-Larsen. 2013. "The Neighborhood Energy Balance Equation: Does Neighborhood Food Retail Environment + Physical Activity Environment = Obesity? The CARDIA Study." PLos ONE 8(12): 1-11.

Booth, Katie M., Megan M. Pinkston, and Walker S. Carlos Poston. 2005. "Current Research: Obesity and the Built Environment." Supplement to the Journal of the American Dietetic Association May: S110-S117.

Braun, Micheal T., and Frederick L. Oswald. 2011. "Exploratory Regression Analysis: A Tool for Selecting Models and Determining Predictor Importance." Behavior Research Methods 43:331-339.

BRFSS. 2014. Behavioral Risk Factor Surveillance System. May 16. Accessed December 11, 2015. http://www.cdc.gov/brfss/about/index.htm

Budzynska, Katarzyna, Patricia West, Ruth T. Savoy-Moore, Darlene Lindsey, Michael Winter, and PK Newby. 2013. "A Food Desert in Detroit: Associations with Food Shopping and Eating Behavior, Dietary Intakes and Obesity." Public Health Nutrition 16 (2): 2114-2123.

Carr, Lucas J., Shira I Dunsiger, and Bess H. Marcus. 2010. "Walk Score as a Global Estimate of Neighborhood Walkability." American Journal of Preventative Medicine 39(5): 460-463.

Carr, Lucas J., Shira I. Dunsiger, and Bess H. Marcus. 2011. "Validation of Walk Score for Estimating Access to Walkable Communities." British Journal of Sports Medicine 45(14): 1144-1148.

Casagrande, Sarah Stark, Joel Gittelsohn, Alan B. Zonderman, Michele K. Evans, and Tiffany L. Gary-Webb. 2011. "Association of Walkability with Obesity in Baltimore City, Maryland." American Journal of Public Health 101(S1): S318-S324.

Casey, Romain, Jean-Michel Oppert, Christiane Weber, Helene Charreire, Paul Salze, Dominique Badariotti, Arnaud Banos, Claude Fischler, Claudia Giacoman Hernandez, Basile Chaix, and Chantal Simon. 2014. "Determinants of Childhood Obesity: What Can we Learn from Built Environment Studies?" Food Quality and Preference 31: 164-172.

Caspi, Caitlin E., Ichiro Kawachi, S.V. Subramanian, Gary Adamkiewicz, and Gloria Sorensen. 2012. "The Relationship between Diet and Perceived Objective Access to Supermarkets among Low-Income Housing Residents." Social Science & Medicine 75: 1254-1262.

Cataife, Guido. 2014. "Small Area Estimation of Obesity Prevalence and Dietary Patterns: A Model Applied to Rio de Janeiro City, Brazil." Health & Place 26: 47-52.

Centers for Disease Control and Prevention (CDC). 2012. "Defining Overweight and Obesity." Last modified April 27. http://www.cdc.gov/obesity/adult/defining.html.

Centers for Disease Control and Prevention (CDC). 2014. "Fastats: Obesity and Overweight." Accessed 1 October. http://www.cdc.gov/nchs/fastats/obesity-overweight.htm.

Centers for Disease Control and Prevention (CDC). 2014a. "Healthier Food Retail: Beginning the Assessment Process in Your State or Community." Accessed 15 April. http://www.cdc.gov/obesity/downloads/hfrassessment.pdf.

Centers for Disease Control and Prevention (CDC). 2014b. "Census Tract Level State Maps of the Modified Retail Food Environment Index (mRFEI)." Accessed 7 October. ftp://ftp.cdc.gov/pub/Publications/dnpao/census-tract-level-state-maps-mrfei_TAG508.pdf.

Centers for Disease Control and Prevention (CDC). 2014c. "About BMI for Adults." Last modified July 11. http://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html.

Cervero, Robert, and Kara Kockelman. 1997. "Travel Demand and the 3Ds: Density, Diversity, and Design." Transportation Research Part D 2(3): 199-219.

Chen, Susan, Raymond J.G.M. Florax, Samantha Snyder, and Christopher C. Miller. 2010. "Obesity and Access to Chain Grocers." Economic Geography 84: 431-452.

Chen, Xiang, and Jill Clark. 2016. "Measuring Space-Time Access to Food Retailers: A Case of Temporal Access Disparity in Franklin County, Ohio." *Professional Geographer* 68(2): 175-188.

Choi, Yuran, and Tsutomu Suzuki. 2013. "Food Deserts, Activity Patterns, & Social Exclusion: The Case of Tokyo, Japan." Applied Geography 43: 87-98.

Datta, G.S., M. Ghosh, R. Steorts, and J.J. Maples. 2011. "Bayesian Benchmarking with Aplications to Small Area Estimation." *Test* 20: 574-588.

Dietz, William H., and Mary C. Bellizzi. 1999. "Introduction: the Use of Body Mass Index to Assess Obesity in Children." American Journal of Clinical Nutrition 70(1): 123-125.

Economic Research Service (ERS). 2014. "Food Access Research Atlas." *U.S. Department of Agriculture (USDA).* Accessed May. http://www.ers.usda.gov/data-products/food-access-research-atlas.aspx.

Economic Research Service (ERS). 2015. "SNAP-authorized Stores." *U.S. Department of Agriculture (USDA).* www.ers.usda.gov/datafiles/Food_Environment_Atlas/Data_Access_and_Documentation_Downloads/Current_Version/DataDownload.xls

Edwards, Kimberley L., G.P. Clarke, J.K. Ransley, and J. Cade. 2009. "The Neighborhood Matters: Studying Exposures Relevant to Childhood Obesity and the Policy Implications in Leeds, UK." Journal of Epidemiology & Community Health 64: 194-201.

ESRI. 2016. “Interpreting GWR Results.” http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/spatial_statistics_tools/interpreting_gwr_results.htm.

Felton, Gwen M., Marsha Dowda, Dianne S. Ward, Rod K. Dishman, Stewart G. Trost, Ruth Saunders, and Russel R. Pate. 2002. “Differences in Physical Activity Between Black and White Girls Living in Rural and Urban Areas.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* West Sussex, England: John Wiley & Sons Ltd.

Frank, Lawrence D., Brian E. Saelens, James Chapman, James F. Sallis, Jacqueline Kerr, Karen Glanz, Sarah C. Couch, Vincent Learnihan, Chuan Zhou, Trina Colburn, and Kelli L. Cain. 2012. “Objective Assessment of Obesogenic Environments in Youth: Geographic Information System Methods and Spatial Findings from the Neighborhood Impact on Kids Study.” American Journal of Preventative Medicine 42(5): 47-55.

Frank, Lawrence D., Jacqueline Kerr, James F. Sallis, Rebecca Miles, and Jim Chapman. 2008. “A Hierarchy of Sociodemographic and Environmental Correlates of Walking and Obesity.” Preventative Medicine 47: 172-178.

Frank, Lawrence D., Martin A. Andresen, and Thomas L. Schmid. 2004. “Obesity Relationships with Community Design, Physical Activity, and Time Spent in Cars.” American Journal of Preventative Medicine 27(2): 87-96.

Frank, Lawrence D., Thomas L. Schmid, James F. Sallis, James Chapman, and Brian E Saelens. 2005. “Linking Objectively Measured Physical Activity with Objectively Measured Urban Form: Findings from SMARTRAQ.” American Journal of Preventative Medicine 28(2S2): 117-125.

Freeman, Lance, Kathryn Neckerman, Ofira Schwartz-Soicher, James Quinn, Catherine Richards, Michael D. M. Bader, Gina Lovaski, Darby Jack, Christopher Weiss, Kevin Konty, Peter Arno, Deborah Viola, Bonnie Kerker, and Andrew G. Rundle. 2012. “Neighborhood Walkability and Active Travel (Walking and Cycling) in New York City.” Journal of Urban Health: Bulletin of the New York Academy of Medicine 90(4): 575-585.

Gao, Xiaolu, Yasushi Asami, Chang-Jo F. Chung. 2006. “An Empirical Evaluation of Spatial Regression Models.” *Computers & Geosciences* 32: 1040-1051.

Giskes, K., Frank van Lenthe, M. Avedano-Pabon, and J. Brug. 2011. “A Systematic Review of Environmental Factors and Obesogenic Dietary Intakes among Adults: Are we Getting Closer to Understanding Obesogenic Environments?” Obesity Reviews 12: 95-106.

Glazier, Richard H., Maria I. Creatore, Jonathan T. Weyman, Ghazal Fazli, Flora I. Matheson, Peter Gozdyra, Rahim Moineddin, Vered Kaufman Shriqui, and Gillian L. Booth. 2014. "Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada." PLOS One 9(1): 1-9.

Goovaerts, Pierre. 2006. "Geostatistical Analysis of Disease Data: Visualization and Propagation of Spatial Uncertainty in Cancer Mortality Risk Using Poisson Kriging and P-Field Simulation." International Journal of Health Geographics 5: 7-33.

Goovaerts, Pierre. 2010. "Geostatistical Analysis of County-Level Lung Cancer Mortality Rates in the Southeastern United States." Geographic Analysis 42: 32-52.

Groves, Robert M. 2011. "Urban Area Criteria for the 2010 Census, Notice." Federal Register 76(164): 53029-53043.

Gustafson, Alison A., Sarah Lewis, Corey Wilson, and Stephanie Jilcott-Pitts. 2012. "Validation of Food Store Environment Secondary Data Source and the Role of Neighborhood Deprivation in Appalachia, Kentucky." BMC Public Health 12: 688-699.

Hartley, David. 2004. "Rural Health Disparities, Population Health, and Rural Culture." American Journal of Public Health 94(10): 1675-1678.

Hennessy, Erin, Vivica I. Kraak, Raymond R. Hyatt, Julia Bloom, Mark Fenton, Colby Wagoner, and Christina D. Economos. 2010. "Active Living for Rural Children: Community Perspectives Using PhotoVOICE." American Journal of Preventative Medicine 39(6): 537-545.

Hewitt, Ben. 2009. *The Town that Food Saved: How One Community Found Vitality in Local Food.* New York: Rodale Inc.

Hilbrecht, Margo, Bryan Smale, and Steven E. Mock. 2014. "Highway to Health? Commute Time and Well-Being Among Canadian Adults." *World Leisure Journal* 56(2): 151-163.

Hill, James O., and John C. Peters. 1998. "Environmental Contributions to the Obesity Epidemic." Science 280: 1371-1374.

Hill, James O., Holly R. Wyatt, and John C. Peters. 2013. "The Importance of Energy Balance." European Endocrinology 9(2): 111-115.

Hilmers, Anglea, David C. Hilmers, and Jayna Dave. 2012. "Neighborhood Disparities in Access to Healthy Foods and Their Effects on Environmental Justice." American Journal of Public Health 102(9): 1644-1654.

Hinde, Sarah, and Jane Dixon. 2005. “Changing the Obesogenic Environment: Insights from a Cultural Economy of Car Reliance.” Transportation Research Part D 10: 31-53

Huffman, Fatma G., Sankarabharan Kanikireddy, and Manthan Patel. 2010. “Parenthood—A Contributing Factor to Childhood Obesity.” International Journal of Environmental Research and Public Health 7(7): 2800-2810.

Humpel, Nancy, Neville Owen, and Eva Leslie. 2002. “Environmental factors Associated with Adults’ Participation in Physical Activity.” American Journal of Preventative Medicine 22(3): 188-199.

Humpel, Nancy, Neville Owen, Don Iverson, Eva Leslie, and Adrian Bauman. 2004. “Perceived Environment Attributes, Residential Location, and Walking for Particular Purposes.” American Journal of Preventative Medicine 26(2): 119-125.

Hutchinson, Paul L., J. Nicholas Bodor, Chris M. Swalm, Janet C. Rice, and Donald Rose. 2012. “Neighbourhood Food Environments and Obesity in Southeast Louisiana.” Health & Place 18: 854-860.

Inagami, Sanae, Deborah A. Cohen, Brian Karl Finch, and Steven M. Asch. 2006. “You Are Where You Shop: Grocery Store Locations, Weight, and Neighborhoods.” American Journal of Preventative Medicine 31(1): 10-17.

InfoUSA. 2015. “SIC and Franchise for Select Counties.” Papillion, NE.

Joens-Matre, Roxane R., Gregory J. Welk, Miguel A. Calabro, Daniel W. Russell, Elizabeth Nicklay, and Larry D. Hensley. 2008. “Rural-Urban Differences in Physical Activity, Physical Fitness, and Overweight Prevalence of Children.” *The Journal of Rural Health* 24(1): 49-54.

Larson, Nicole I., Mary T. Story, and Melissa Nelson. 2009. “Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S.” American Journal of Preventative Medicine 36(1): 74-81.

Lee, Jay, Mohammad Alnasrallah, David Wong, Heather Baird, and Everett Logue. 2014. “Impacts of Scale on Geographic Analysis of Health Data: An Example of Obesity Prevalence.” ISPRS International Journal of Geo-Information 3: 1198-1210.

Li, Wenjun, Jennifer L. Kelsey, Zi Zhang, Stephanie C. Lemon, Solomon Mezgebu, Cynthia Boddie-Willis, and George W. Reed. 2009. “Small-Area Estimation and Prioritizing Communities for Obesity Control in Massachusetts.” American Journal of Public Health 99(3): 511-519.

Li, Wenjun, Thomas Land, Zi Zhang, Lois Keithly, and Jennifer L. Kelsey. 2009a. "Small-Area Estimation and Prioritizing Communities for Tobacco Control Efforts in Massachusetts." *American Journal of Public Health* 99(3): 470-479.

Liese, Angela D., Timothy L. Barnes, Archana P. Lamichhane, James D. Hibbert, Natalie Colabianchi, and Andrew B. Lawson. 2013. "Characterizing the Food Retail Environment: Impact of Count, Type, and Geospatial Error in 2 Secondary Data Sources." Journal of Nutrition Education and Behavior 45(5): 435-442.

Liu, Jihong, Kevin J. Bennett, Nusrat Harun, and Janice C. Probst. 2008. "Urban-Rural Differences in Overweight Status and Physical Inactivity Among US Children Aged 10-17 Years." *The Journal of Rural Health* 24(4): 407-415.

Lochmueller Group and Alta 2014. "Final Report: Multimodal Transportation System Assessment." *Southern Illinois Metropolitan Planning Organization* (June). http://www.greateregypt.org/SIMPO/simpo-planning-documents/

Lochmueller Group. 2014. "Summary of Pedstrian and Bicycle Facilities Inventory." *Southern Illinois Metropolitan Planning Organization* (June). http://www.greateregypt.org/SIMPO/simpo-planning-documents/

Lopez, Russ P. 2007. "Neighborhood Risk Factors for Obesity." Obesity 15: 2111-2119.

Mackenbach, Joreintje D., Harry Rutter, Sofie Compernolle, Ketevan Glonti, Jean-Michel Oppert, Helene Charreire, Ilse De Bourdeaudhuij, Johannes Brug, Giel Nijpels, and Jeroen Lakerveld. 2014. "Obesogenic Environments: A Systematic Review of the Association between the Physical Environment and Adult Weight Status, the SPOTLIGHT Project." BMC Public Health 14: 233-260.

Malec, Donald, William W. Davis, and Xin Cao. 1999. "Model-Based Small Area Estimate of Overweight Prevalence Using Sample Selection Adjustment." Statistics in Medicine 18: 3189-3200.

Maley, Mary, Barbour S. Warren, and Carol M. Devine. 2010. "Perceptions of the Environment for Eating and Exercise in a Rural Community." Journal of Nutrition Education and Behavior 42(3): 185-191.

Manzoli, L., P. Vallari, G.M. Pirone, and A. Boccia. 2007. Marital Status and Mortality in the Elderly: A Systematic Review and Meta-Analysis." Social Science and Medicine 64(1): 77-94.

McDermott, Andrew J., and Mark B. Stephens. 2010. "Cost of Eating: Whole Foods Versus Convenience Foods in a Low-Income Model." *Family Medicine* 42(4): 280-284.

McEntee, Jesse, and Julian Agyeman. 2010. “Towards the Development of a GIS Method for Identifying Rural Food Deserts: Geographic Access in Vermont, USA.” Applied Geography 30: 165-176

McMurray, Robert G., Joanne S. Harrell, Shrikant I. Bangdiwala, and Shibing Deng. 1999. “Cardiovascular Disease Risk Factors and Obesity of Rural and Urban Elementary School Children.” *The Journal of Rural Health* 15(4): 365-374.

McPhail. 2013. “The Rural and the Rotund? A Critical Interpretation of Food Desert and Rural Adolescent Obesity in the Candian Context.” Health & Place 22: 132-139.

Meit, Michael, Alana Knudson, Tess Gilbert, Amanda Tzy-Chyi Yu, Erin Tanenbaum, Elizabeth Ormson, Shannon TenBroeck, Alycia Bayne, Shena Popat, NORC Walsh Center for Rural Health Analysis. 2014. The 2014 Update of the Rural-Urban Chartbook. Grand Forks, North Dakota: Rural Health Reform Policy Research Center.

Merchant, Anwar T., Patrick DeLuca, Mamdouh M. Shubair, Julie Emili, and Pavlos Kanaroglou. 2011. “Local Data for Obesity Prevention: Using National Data Sets.” The Professional Geographer 63(4): 550-559.

Morland, Kimberly B., and Kelly R. Evenson. 2009. “Obesity Prevalence and the Local Food Environment.” Health & Place 15: 491-495.

Morland, Kimberly, Ana V. Diez Roux, and Steve Wing. 2006. “Supermarkets, Other Food Stores, and Obesity: The Atherosclerosis Risk in Communities Study.” American Journal of Preventative Medicine 30(4): 333-9.

Morland, Kimberly, Steve Wing, and Ana Diez Roux. 2002. “The Contextual Effect of the Local Food Environment on Residents’ Diets: The Atherosclerosis Risk in Communities Study.” Health & Place 18: 46-54.

Morton, Lois Wright and Troy C. Blanchard. 2007. “Starved for Access: Life in Rural America’s Food Deserts” The Rural Sociologist: A Publication of the Rural Sociological Society 27(1): 20-29.

National Heart, Lung, and Blood Institute (NHLBI). 2012. “What Causes Overweight and Obesity.” National Institutes of Health. Last modified July 13. www.nhlbi.nih.gov/health/health-topics/obe/causes.

Nelson, Melissa C., Penny Gordon-Larsen, Yan Song, and Barry M. Popkin. 2006. “Built and Social Environments: Associations with Adolescent Overweight and Activity.” *American Journal of Preventative Medicine* 31(2): 109-117.

Paul Simon Public Policy Institute. 2014. “Jackson/Williamson County Poll.” *Southern Illinois University*: Carbondale.

Paul Simon Public Policy Institute. 2015. "The Southern Illinois Poll Poll." *Southern Illinois University*: Carbondale.

Pfeffermann, Danny, and Charles Barnard. 1991. "Some New Estimators for Small-Area Means with Applications to the Assessment of Farmland Values." *Journal of Business & Economic Statistics* 9(1): 73-84.

Pfeffermann, Danny, and Richard Tiller. 2006. "Small-Area Estimation with State-Space Models Subject to Benchmark Constraints." *Journal of American Statistical Association* 101: 1387-1397.

Pienta, Amy M., Mark D. Hayward, and Kristi R. Jenkins. 2000. "Health Consequences of Marriage for the Retirement Years." Journal of Family Issues 21(5): 559-586.

Powell, Lisa M., and Yanjun Bao. 2009. "Food Prices, Access to Food Outlets and Child Weight." Economics and Human Biology 7(1): 64-72.

Powell, Lisa M., M. Christopher Auld, Frank J. Chaloupka, Patrick M. O'Malley, and Lloyd D. Johnston. 2007. "Associations between Access to Food Stores and Adolescent Body Mass Index." American Journal of Preventative Medicine 33(4): 301-307.

Rahman, Tamanna, Rachel A. Cushing, and Richard J. Jackson. 2011. "Contributions of Built Environment to Childhood Obesity." Mount Sinai Journal of Medicine 78: 49-57.

Rundle, Andrew, Kathryn M. Neckerman, Lance Freeman, Gina S. Lovasi, Marnie Purciel, James Quinn, Catherine Richards, Neelanjan Sircar, and Christopher Weiss. 2009. "Neighborhood Food Environment and Walkability Predict Obesity in New York City." Environmental Health Perspectives 117(3): 442-447.

Saelens, Brian E., James F. Sallis, Lawrence D. Frank, Sarah C. Couch, Chuan Zhou, Trina Coburn, Kelli L. Cain, James Chapman, and Karen Glanz. 2012. "Obesogenic Neighborhood Environments, Child and Parent Obesity: The Neighborhood Impact on Kids Study." American Journal of Preventative Medicine 42(5): e57-e64.

Sallis, James F. Brian E. Saelens, Lawrence D. Frank, Terry L. Conway, Donald J. Slymen, Kelli L. Cain, James E. Chapman, and Jacqueline Kerr. 2009. "Neighborhood Built Environment and Income: Examining Multiple Health Outcomes." Social Science & Medicine 68: 1285-1293.

Sandercock, Gavin, Caroline Angus, and Joanna Barton. 2010. "Physical Activity Levels of Children Living in Different Built Environments." Preventative Medicine 50: 193-198.

Scarpello, T., F. Poland, N. Lambert, and T. Wakeman. 2009. "A Qualitative Study of the Food Related Experiences of Rural Village Shop Customers." Journal of Human Nutrition and Dietetics 22: 108-115

Schafft, Kai A., Eric B. Jensen, and C. Clare Hinrichs. 2009. "Food Deserts and Overweight Schoolchildren: Evidence from Pennsylvania." Rural Sociology 74(2): 153-177.

Sharkey, Joseph R., and Scott Horel. 2008. "Neighborhood Socioeconomic Deprivation and Minority Composition Are Associated with Better Potential Spatial Access to the Ground-Truthed Food Environment in a Large Rural Area." The Journal of Nutrition 138: 620-627.

Sharkey, Joseph R., Scott Horel, and Wesley R. Dean. 2010. "Neighborhood Deprivation, Vehicle Ownership, and Potential Spatial Access to a Variety of Fruits and Vegetables in a Large Rural Area in Texas." International Journal of Health Geographics 9: 26-52.

Sohi, Inderbir, Bethany A. Bell, Jihong Liu, Sarah E. Battersby, and Angeal D. Liese. 2014. "Differences in Food Environment Perception and Spatial Attributes of Food Shopping Between Residents of Low and High Food Access Areas." Journal of Nutrition Education and Behavior: 1-9.

Springer, Andrew E. Deanna M. Hoelscher, and Steven H. Kelder. 2006. "Prevalence of Physical Activity and Sedentary Behaviors in US High School Students by Metropolitan Status and Geographic Region. *Journal of Physical Activity & Health* 3(4): 365-380.

Springer, Andrew E., Deanna M. Hoelscher, Steven H. Kelder, Brian Castrucci, and Adriana Perez. 2009. "Prevalence of Physical Activity and Sedentary Behaviors by Metropolitan Status in 4^{th}-, 8^{th}-, and 11^{th}-Grade Students in Texas, 2004-2005." *Preventing Chronic Disease: Public Health Research, Practice, and Policy* 6(1): A21.

Suminski, Richard R., Walker S. Carlos Poston, Rick L. Petosa, Emily Stevens, and Laura M. Katzenmoyer. 2005. "Features of the Neighborhood Environment and Walking by U.S. Adults." American Journal of Preventative Medicine 28(2): 149-155.

Swinburn, Boyd, Garry Egger, and Fezeela Raza. 1999. "Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity." Preventative Medicine 29: 563-570.

United State Census Bureau. 2013a. "Five Year Estimates." *American Community Survey*. Washington, DC: U.S. Census Bureau's American Community Survey Office.

United States Census Bureau. 2013. "Block Group Boundaries." *TIGER/Line® Shapefiles.* Washington, DC: U.S. Census Bureau.

---"County Boundaries." *TIGER/Line® Shapefiles.* Washington, DC: U.S. Census Bureau.

---"Roads." *TIGER/Line® Shapefiles.* Washington, DC: U.S. Census Bureau.

United States Department of Agriculture (USDA). 2009. "Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and their Consequences, Report to Congress." Economic Research Service. June. Accessed 10, 2015. http://www.ers.usda.gov/media/242675/ap036_1_.pdf.

Walk Score. 2014. "Walk Score Methodology." Accessed on 7 October at http://www.walkscore.com/methodology.shtml.

Wang, Chih-Hao, and Na Chen. 2015. "A GIS-Based Spatial Statistical Approach to Modeling Job Accessibility by Transportation Mode: Case Study of Columbus, Ohio." *Journal of Transport geography* 45: 1-11.

Wang, Fahui, Ming Wen, and Yanqing Xu. 2013. "Population-Adjusted Street Connectivity, Urbanicity and Risk of Obesity in the U.S." Applied Geography 41: 1-14.

Wang, Junyuan, Wayne A. Fuller, and Yongming Qu. 2008. "Small Area Estimation Under a Restriction." *Survey Methodology* 34: 29-36.

Wang, Youfa, and May A. Beydoun. 2007. "The Obesity Epidemic in the United States—Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis." Epidemiologic Reviews 29: 6-28.

Wheeler, David, and Michael Tiefelsdorf. 2005. "Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression." *Journal of Geographical Systems* 7: 161-187.

Yousefian, Anush, A. Leighton, K. Fox, and D. Hartley. 2011. "Understanding the Rural Food Environment – Perspectives of Low-Income Parents." Rural and Remote Health 11: 1631-1641

Zhang, Lianjun, Jeffrey H. Gove, and Linda S. Health. 2005. "Spatial Residual Analysis of Six Modeling Techniques." *Ecological Modelling* 186: 154-177.

Zhang, Qi, and Youfa Wang. 2004. "Socioeconomic Inequality of Obesity in the United States: Do Gender, Age, and Ethnicity Matter?" *Social Science & Medicine* 58: 1171-1180.

Zhang, Xingyou, James B. Hold, Hua Lu, Stephen Onufrak, Jiawen Yang, Steven P. French,and Daniel Z. Sui. 2014. "Neighborhood Commuting Environment and Obesity in the United States: An Urban-Rural Stratified Multilevel Analysis." Preventative Medicine 59: 31-36.

Zick, Cathleen D., Ken R. Smith, Jessie X. Fan, Barbara B. Brown, Ikuho Yamada, and Lori Kawleski-Jones. 2009. "Running to the Store? The Relationship between Neighborhood Environments and the Risk of Obesity." Social Science & Medicine 69: 1493-1500.

VITA

Graduate School
Southern Illinois University

Shiloh L. Deitz

shiloh.leah.deitz@gmail.com

Whitworth University
Bachelor of Arts, Peace Studies, May 2009

Southern Illinois University Carbondale
Masters of Arts, Sociology, May 2014

Special Honors and Awards:

Presidential Scholarship, Whitworth University, 2006

Masters Fellowship Award, SIUC, 2012

Celia Howard Fellowship, SIUC, 2014

Ben Dziegielewski Graduate Scholarship, SIUC, 2015

Thesis Title:

A Spatial Analysis of the Relationship between Obesity and the Built Environment in Southern Illinois

Major Professor: Leslie A. Duram

Publications:

Deitz, Shiloh, and Kristen Barber. 2015. "Geographies of Inequality: Urban Renewal and the Race, Gender, and Class of Post-Katrina New Orleans." Race, Gender & Class (December): New Orleans for Whom and What?

Deitz, Shiloh. 2015. "Envisioning Carbondale's Creative Community: Current Challenges and Opportunities." Paul Simon Public Policy Institute, November.

Deitz, Shiloh. 2015. "Internet Access in Southern Illinois: Results from the Jackson/Williamson County Survey." Paul Simon Public Policy Institute, July.

Deitz, Shiloh. 2015. "Gender Sensitive Disaster Policy in Southern Illinois." The Simon Review. Carbondale, IL: Paul Simon Public Policy Institute.

Barber, Kristen, and Shiloh Deitz. 2015. "Missing in the Storm: The Gender Gap in Hurricane Katrina Research and Disaster Management Efforts." In Rethinking

Disaster Recovery: A Hurricane Katrina Retrospective, edited by Jeannie Haubert. Lanham, MD: Lexington Books.

Deitz, Shiloh. 2014. “Alumni Community Survey.” Paul Simon Public Policy Institute, November.

Deitz, Shiloh, and Kent Dolezal. 2014. “Simon Poll 2008-2014.” Open SIU; Odum Institute, Open ICPSR [Codebook], September.

Deitz, Shiloh, and Kent Dolezal. 2014. “The Southern Illinois Poll 2010-2013.” Open SIU; Odum Institute, Open ICPSR [Codebook], September.

Deitz, Shiloh. 2014. “Religiosity and Attitudes toward Immigrants and Immigration.” SIU Online Research Papers: OpenSIUC.