

# An Ecological Approach to Understanding Adult Obesity Prevalence in the United States: A County-level Analysis using Geographically Weighted Regression

Nyesha C. Black

Received: 15 February 2013 / Accepted: 10 June 2014 /  
Published online: 22 June 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** This study examines ecological influences on adult obesity prevalence in the coterminous United States. Several secondary data sources are used in this study to construct a rich dataset of county-level demographic, socioeconomic, and environmental variables. This study uses a spatially explicit approach by using traditional regression methods (i.e., ordinary least squared regression (OLS)), along with geographic weighted regression (GWR) to explore non-stationarity in the relationships between obesity and selected covariates. OLS results reveal that there is a positive relationship between adult obesity and poverty, black residents, Native American residents, and adult physical inactivity at the county level. There is a negative relationship between the percentage of residents who are rural, Hispanic, and college educated. Furthermore, GWR results confirm that place matters and the relationship between ecological influences and obesity prevalence varies substantially across place. GWR provides an empirical basis to design interventions that effectively target obesity at a more local level.

**Keywords** Geographically weight regression · Obesity · Health

## Introduction

Obesity is a pressing public health concern in the United States, and internationally. In the U.S., the obesity epidemic affects more than one-third of adults, and almost 17 % of adolescents (Flegal et al. 2012; Ogden et al. 2012). Obesity has a number of medical consequences that diminishes metabolic, anesthetic, endocrine, and psychological health and well-being (Lean 2010). The economic costs are also substantial. (Finkelstein et al. 2009) analysis of the monetary cost of obesity reveals that medical expenditures associated with obesity in the United State increased from \$78.5 billion in

---

N. C. Black (✉)

Department of Sociology Population Research Institute, The Pennsylvania State University,  
University Park, PA, USA  
e-mail: ncb146@psu.edu

1998 to an estimated \$147 billion in 2008. Government programs such as Medicare and Medicaid absorbed a significant portion of the cost. Furthermore, it is estimated that obese workers cost private employers more than \$73 billion a year in lost productivity (Finkelstein et al. 2009).

A number of factors that are associated with obesity, also spatially vary across place. For example, physical inactivity is positively associated with obesity risk throughout the life course (Hu et al. 2003). Americans are also consuming more calorie-dense foods (Ledikwe et al. 2006), such as fast foods (Bowman and Vinyard 2004), and sugared-sweetened beverages (Ludwig et al. 2001; Schulze et al. 2004; Popkin 2008). While personal dietary behaviors are linked to increased trends in obesity, emerging evidence implicates the importance of the local ecology on health behaviors and health outcomes (Lovasi et al. 2009). Quality of life and health are associated with not only the social context, but also the physical environment (Pacione 1984), including the natural and built environment (Kearns and Gesler 1998; Macintyre et al. 2002). The built environment has been recognized as a crucial marker for access to resources, such as food stores, which may mitigate the risk of poor health outcomes such as obesity. Poor access to nutritious food options can be a barrier for changing poor personal dietary behaviors that increase the risk of obesity (Baker et al. 2006). In the U.S., it has been well documented that there is a positive association between food access, a nutritious diet, and lower rates of obesity and other diet-related diseases (Larson et al. 2009); and differences in food retail access may contribute to and perpetuate health disparities (Diez-Roux and Mair 2010). Empirical evidence shows a negative relationship between the consumption of nutritious foods and limited access to supermarkets (Morland et al. 2002). Convenience stores are positively associated with a less nutritious diet, and higher rates of obesity at the local level (Morland et al. 2002; Bodor et al. 2010). Furthermore, features of the built environment also play a role in physical activity levels of local residents (Sallis and Glanz 2009). For example, the presence of sidewalks and adequate lighting, and population density as opposed to heavy traffic, pedestrian unfriendly land use, and vacant land may promote or inhibit physical activity (Berrigan and Troiano 2002).

Obesity rates remain high, especially among minorities and low-income communities; and there are persistent spatial disparities in obesity rates across residential context in the United States. There is much debate in the health disparities literature about the role of individual and contextual influences on health. Macintyre, Maciver, and Sooman (1993) argue that the social and built environment, at the local level, foster or inhibit health at the individual level. Thus, public health interventions should focus on improvements in the local infrastructure to promote better health outcomes. Individual and contextual effects are often considered competing influences on health; but people shape places, just as places shape people (Dorling 2001; Tunstall, Shaw, and Dorling 2004). In turn, both material infrastructure and collective social functioning may influence health (Macintyre et al. 2002). However, it is difficult to uncover individual and place effects on health (Mitchell 2001), but it is imperative to understand the role of place on health (Cummins et al. 2007), especially in a context of stark spatial inequalities (Shaw, Dorling, and Mitchell 2002).

Social and economic inequalities are salient determinants of population health (Lynch and Kaplan 1997), and the spatial patterning of health inequalities varies across various health outcomes (Arcaya et al. 2012), including obesity. It is well documented

that these protective factors and disease prevalence are not evenly distributed across people or places (Geronimus et al. 1996; Murray et al. 2006; Cossman et al. 2007; Ezzat et al. 2008). Inequalities in health are pervasive across racial and ethnic groups, socioeconomic status, gender status, and metropolitan status. A host of health outcomes, including obesity, vary across place (Murray et al. 2006). Higher prevalence of obesity has been reported for rural populations compared with urban populations (Jackson et al. 2005). Rural–urban variation in obesity rates and health may reflect socioeconomic differences, with rural areas being more characterized by local poverty and a lower educated population (Ramsey and Glenn 2002). The demography of obesity is partly attributed to the unequal distribution of resources, as well as social and demographic characteristics. Thus, a more nuanced understanding of the geographic variation in obesity and the determinants of disease are imperative to reduce obesity rates and to eliminate disparities across the United States. Several studies have documented the prevalence and trends of obesity across the United States (Wang and Beydoun 2007; Jackson et al. 2005), and while other studies have accounted for the importance of social context on obesity by using multilevel modeling (Kim et al. 2006), limited studies have examined geographic variation in obesity rates (Wen et al. 2010), and few studies have examined this relationship using spatial analysis in the United States. Some studies found supportive evidence of spatial variation in the relationship between obesity and covariates (Procter et al. 2008; Wen et al. 2010), while other studies found no evidence of geographic variation in the relationship (Fraser et al. 2012; Schuurman, Peters and Oliver 2012).

This objective of this study is to explore the relationship between aggregate level variables and obesity among coterminous counties in the United States. This study contributes to the obesity literature, by moving beyond traditional aspatial methods such as ordinary least square regression by using geographically weighted regression (GWR), an exploratory technique that accounts for the heterogeneous relationships between each predictor and the outcome variable (i.e., adult obesity). Traditional regression methods assume that the relationship between the dependent variable and covariates are stationary or the same across place, thereby potentially concealing variations in relationships across contexts. More concerning, traditional methods, which fail to account for spatially varying relationships across place may lead to biased parameter estimates, misleading conclusions, and ineffective interventions (Ali et al. 2007). A spatially explicit approach, however, can provide insights into where and how the relationship between the independent variables and obesity vary by context. Few studies have employed GWR to examine contextual influences on obesity prevalence, at any aggregate level (Procter et al. 2008; Chalkias et al. 2013; Wen, Chen and Tsai 2010). This empirical analysis provides a useful demonstration of how GWR may be used to examine how context shapes obesity.

## Methods

### Data and Data Sources

Several secondary county-level data sources detailed in Table 1 were used in this study. The main source of data came from the 2010 United States Department of Agriculture

**Table 1** Description of Data Sources

Variable	Source	Original Source	Year
% Obesity	Food Environment Atlas	Behavioral Risk Factor Surveillance System	2008
% Rural	U.S. Census Bureau	U.S. Census Bureau	2010
% Population Density	U.S. Census Bureau	U.S. Census Bureau	2010
% Black	Food Environment Atlas	U.S. Census Bureau	2008
% Hispanic	Food Environment Atlas	U.S. Census Bureau	2008
% Asian	Food Environment Atlas	U.S. Census Bureau	2008
% Native American	Food Environment Atlas	U.S. Census Bureau	2008
% Poverty	Food Environment Atlas	U.S. Census Bureau	2008
% College Education	U.S. Census Bureau	U.S. Census Bureau	2010
% HH with no car, >1 mile from grocery store	Food Environment Atlas	U.S. Department of Agriculture	2006
Food Retail Landscape	Food Environment Atlas	U.S. Census Bureau	1999
Amenity Scale	Food Environment Atlas	U.S. Department of Agriculture	2008
% Adult Physical Inactivity	Food Environment Atlas	Behavioral Risk Factor Surveillance System	

(USDA) Food Environment Atlas. The Food Environment Atlas has over 160 county-level indicators of food access, health and well-being, and community and social characteristics. For this study, the indicators that were the most relevant to understanding variation in adult obesity rates in the United States were included. The Food Environment Atlas is a compilation of county-level data from various government sources (i.e., U.S. Census Bureau, the Centers for Disease Control and Prevention, and the U.S. Department of Agriculture). All coterminous counties or county-equivalents ( $n=3108$ ) were included in the analysis. Supplemental data was gathered from the Behavioral Risk Factor Surveillance System (BRFSS), a national random-digit-dialing health questionnaire sponsored by the Centers for Disease Control and Prevention (CDC). The BRFSS includes data from all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam. The BRFSS is the largest telephone health survey in the world, with more than 350,000 adults surveyed each year. The data provides state-specific data on risk behaviors related to personal behaviors and preventive health practices. The 2008 BRFSS estimates 3-year (2007, 2008, and 2009) county-level statistics for health risk factors such as physical inactivity and obesity. 3 years of data were used to improve the precision of the year-specific county-level estimates of selected risk factors.

The 2008 age-adjusted, county-level adult obesity rate in the United States is the dependent variable in this study. County-level age-adjusted obesity prevalence is a continuous variable collected from the BRFSS, and is determined by the age-adjusted percentage of person age 20 and older who are obese. BRFSS collects self-reports of height and weight, and individuals are considered obese if their body mass index  $\geq 30$ . The residential context of each county is measured by population density and the percentage of residents living in rural areas. Population density is measured by the total

county population per land area. The percentage of total county residents who live in rural areas of the county is included as a covariate to account for rural and urban differences in obesity rates. The percentage of Blacks, Hispanics, Asians, and Native Americans is secondary data reported in the Food Environment Atlas. The percentage of the white population was not included in the analysis to avoid issues of multicollinearity. The acceptance and adherence to public health interventions is shaped by a community's ability to understand the importance of obesity prevention (House et al. 2009). Therefore, the social and economic context of each county is measured by the percent of county residents who have completed at least 4 years of post-secondary education. The percentage of county residents who live below the federal poverty threshold is also included as a covariate to capture the social and economic context at the county level. The income threshold in 2010 was \$22, 314 for a family of four.

Environmental measures for both the built and natural environment are included in this study. Food store access in the built environment is measured by the percentage of county households without a car and who must travel more than one mile to a large grocery store or supermarket in urban areas, or more than ten miles in rural areas. In an effort to address potential errors of multicollinearity between the food retail variables and population density, an entropy score was also calculated to measure food store access. The entropy score measures the food retail mix of supermarket, grocery stores, convenience stores, and specialty food stores at the county level (Thiele and Weiss 2003). The entropy score, or diversity of the food retail landscape can be defined as:

$$E_i = \sum_{r=1}^r (\pi_{ri}) \ln[1/\pi_{ri}]$$

where  $\pi_{ri}$  refers to a particular food retail store (i.e., supermarket, grocery store, convenience store, or specialized food store) proportion of the sum of food retail stores in county  $i$  (Iceland 2004). An entropy index ranges in values from 0 to 1; an entropy score closer to 1 indicates that a county has a diverse food retail landscape, while an entropy score of 0 indicates that a county has only one type of retail store. The natural amenity resources scale was used to assess the effect of the natural environment on obesity rates in the United States. The natural amenity resources scale is a composite of access to natural landscape forms such as waterfronts, mountains, and temperate climates. The index of natural amenity index is constructed by the USDA Economic Research Service, and reflects the sum of six standardized measures: average January temperature, average January days of sun, low January to July average temperature gain, low average July humidity, topographical variation, and the ratio of water area to land area. A county with an amenity scale above 3 has high natural amenities, while a county with a negative score on the amenities scale has low access to natural landscapes. County-level reports of physical inactivity are also included as an independent variable to examine how aggregate level health behaviors predict obesity rates in the United States. The percentage of adults who are physically inactive reflects BRFSS estimates of adults who did not self-report at least 150 min of moderate physical activity per week, or 75 min of vigorous activity per week, which meets U.S. public health guidelines for physical activity.

## Global Regression

Traditional regression methods yield a single or ‘global’ parameter estimate across the entire study area, and can be expressed as:Eq 1

$$\hat{Y}_i = \beta_0 + \sum \beta_k x_{ik} + \varepsilon_i \quad (1)$$

where  $\hat{y}$  is the predicted value of the dependent variable for observation  $i$ ,  $\beta_0$  is the intercept,  $\beta_k$  is the parameter estimate for the independent variable,  $x_{ik}$  represents the  $i_{th}$  observation of the  $k_{th}$  independent variable, and  $\varepsilon_i$  represents the error term (Allison 1999). However, neighboring geographies (i.e., counties) are likely to share similar compositional and contextual characteristics that are spatially correlated. More specifically, a set of conditions, in this case obesity and several predictors of obesity, of one county may affect obesity outcomes in neighboring counties (Tobler 2004). The inherent feature of geographical or ecological units sharing similar characteristics violates the assumption of normal distribution and uncorrelated error terms endorsed when using aspatial models. A violation of this assumption will yield biased parameter estimates, and generate low estimates of the confidence intervals and real variance (Anselin 1988; Ward and Gleditsch 2008). Geographically weighted regression (GWR) corrects for both spatial heterogeneity and spatial dependence of the data by estimating local, more robust parameters that capture the distinctiveness of place, and the spatial variations between dependent and independent variables (Brunsdon et al. 1998). This spatially explicit approach helps identify relationships that are likely masked by using traditional regression methods, which assumes that the relationship between variables are the same throughout the entire study area.

## Geographically Weighted Regression

The GWR model can be expressed with the following function: Eq.2

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (2)$$

$y_i$  is the percentage of the adult population reporting obesity for county  $i$ ,  $(u_i, v_i)$  denotes the coordinates of the centroid of county  $i$ ,  $\beta_{0i}$  and  $\beta_{ni}$  represents the local estimates intercept and the effect of variable  $k$  on county  $i$ . Each county centroid,  $i$ , is a regression point, and observations closer in proximity to  $i$  are weighted more than counties that are farther away (Fotheringham et al. 2002). Hence, neighboring counties have more influence than distant counties. The geographic center of each county area, or the geographic-weighted centroid, instead of population-weighted centroid, is used in this study since all the data is aggregated at the county-level, and the specific location of individuals that make up the county data is unknown in the dataset. There is little variation between population and geographic-weighted centroids for counties with medium to high population density (Rushton et al. 2010); however, the geographic-weighted centroids may be less robust among counties with a low population density. An adaptive weighting scheme, as used in the analysis reported here, accounts for the variation in the size of counties, and the sparseness of counties in parts of the United States. The adaptive weighting function permits a larger bandwidth when the data are sparse and a small bandwidth when data are more dense (Fotheringham et al. 2002, p.

46). The bandwidth influences the weighting scheme, and is expressed in the same number of units included in the analysis. The large bandwidth approaches the parameter achieved in a global model, while a small bandwidth reveals local relationships (Charlton, Fotheringham and Brunson 2003). The adaptive spatial kernel also ensures that an equal number of observations receive a non-zero weight value at each regression point (Bitter et al. 2007, p.15). The Akaike Information Criterion ( $AIC_c$ ), is not defined by the kernel function, but is used to determine the optimal kernel bandwidth size (Fotheringham et al. 2002). The  $AIC_c$  uses a bi-square decaying function based on nearest neighbors used (Fotheringham et al. 2002). To test for spatial non-stationarity, I relied on the Monte Carlo significance test by using GWR 3.0 software (Charlton et al. 2003). The Monte Carlo test provides a randomized null hypothesis model, in which the location of each observation is arbitrarily assigned to the predictor and response variables. The null hypothesis assumes there is no significant difference in patterns of parameters estimates across place (Brunson et al. 1998). Hence, a significant Monte Carlo test reveals spatial variation in local response variables (Fotheringham et al. 2002). The model with the lowest  $AIC_c$  value indicates a better model fit. Finally, all GWR regression output was transferred to ArcGIS 10, and coefficient estimates were mapped to better interpret the results.

## Results

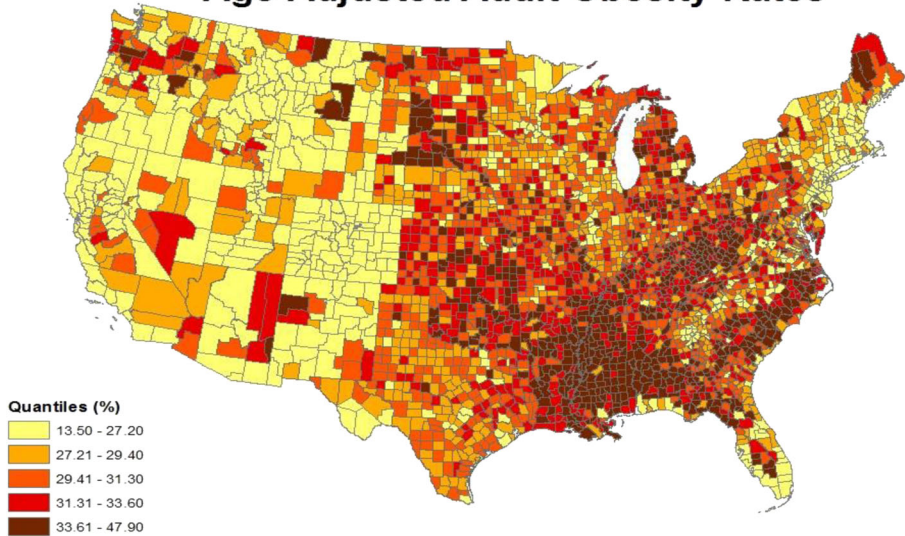
Table 2 details the sample of counties across the continental United States. The average rate of age-adjusted adult obesity was about 30 %, and the lowest rate of obesity was reported in Routt County, Colorado (13.5 %), while Greene County, Alabama reported the highest prevalence of adult obesity (47.9 %). Fig. 1 confirms that obesity rates vary substantially across the United States, and the epidemic is mostly concentrated in counties throughout the Southeast, and many of the counties in the Appalachian

**Table 2** Descriptive Statistics for variables in the study by county ( $N=3,108$ )

	Mean	S.D.	Min	Max.
% Obesity	30.32	3.61	13.5	47.90
% Rural	59.85	30.91	0.00	100.00
Population Density	259.56	1764.24	0.06	71,467.32
% Black	9.00	14.32	0.00	85.50
% Hispanic	7.71	12.83	0.10	97.30
% Asian	1.01	1.92	0.00	31.93
% Native Americans	1.59	6.20	0.00	83.40
% Poverty	15.27	6.05	3.10	54.40
% College education	16.51	7.80	4.92	63.75
% HH with no car, > 1 miles to grocery store	3.98	2.60	0.00	27.91
Food retail landscape	0.57	0.16	0.00	1.00
Amenity Scale	0.05	2.28	-6.40	11.17
% Adult Physical Inactivity	26.94	4.94	10.60	42.80



## Age-Adjusted Adult Obesity Rates



**Fig. 1** Adult obesity prevalence in the United States (2009) (Black and White Map) Adult obesity prevalence in the United States (2009)

Mountain Region. There are noticeably lower rates of adult obesity in the West. County-level population and land area varies substantially across the United States. According to the 2010 U.S. Census, Los Angeles County, California, the most populated county, had approximately 10 million residents, while Loving County, Texas had less than 100 residents. The largest counties by land area are in the West, while counties among the smallest in geographic area are along the East coast. Thus, both population density and the proportion of residents in rural areas vary widely across the United States (U.S. Census 2010). Similarly, the racial composition of each county varies greatly. African-Americans, on average make up about 9 % of the total population, while Hispanics make up less than 8 % of the total population at the county-level. In contrast, Asian-Americans and Native-Americans each make-up less than 2 % of the total population at the county-level. The average county-level poverty rate is about 15 %, while about 16 % of residents have a college education. Less than 4 % of county residents in the United States do not own a car, and have to travel more than one mile to a grocery store, and there is slightly more than a medium diversity of food store options (0.57). Most counties boast some attractive landscape areas (i.e., the mean amenity scale score is 0.05). County-level physical inactivity among adults range from less than 5 % to more than 42 %; adults in Teton County, Wyoming are the most active, compared to adults in Carter County, Kentucky, who are the least physically active in the United States.

General linear regression models calculate one parameter estimate for each variable, and the single measure of model fit. These global estimates for the entire study area are reported in Table 3. For every one percentage increase of rural residents, the county level adult obesity rate decreases by less the .01 %, and the relationship is similar for population density. The effect of racial and ethnic composition on county-level adult obesity varies by group. As the percentage of African-Americans and Native-American



**Table 3** OLS (Global) regression model predicting age-adjusted adult obesity ( $N=3,108$ )

	Coefficient	Std. Error	<i>p</i> -value	VIF
Intercept	21.605	0.580	***	
% Rural	- 0.007	0.002	**	2.21
Population Density	- 0.001	0.000	***	1.19
% Black	0.059	0.005	***	1.84
% Hispanic	- 0.022	0.004	***	1.46
% Asian	0.059	0.033		1.78
% Native Americans	0.075	0.001	***	1.18
% Poverty	0.044	0.012	***	2.60
% College education	-0.118	0.010	***	2.45
% HH with no car, > 1 miles to grocery store	0.042	0.029		2.46
Food retail landscape	0.707	0.351	*	1.42
Amenity Scale	-0.354	0.025	***	1.36
% Adult Physical Inactivity	0.348	0.014	***	2.23
Adjusted R-square	0.600			
Akaike information criterion	14945.360			

\* $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\* $p \leq 0.001$

increases by one unit, the adult obesity rate increases by about .06 % and .07 %, respectively. In contrast, a one-percentage increase in the number of Hispanic residents in a county is associated with a .02 % decrease in the obesity rate. There is no significant relationship between the percentage of Asian residents and adult obesity.

There is a positive relationship between county level poverty rates and obesity rates. In contrast, there is a negative relationship between college educated residents and obesity rates.

The built environment measures show conflicting results. There is no significant relationship between county level adult obesity and the percentage of residents with no car, and more than one mile from a grocery store, when controlling for other variables. In contrast with the literature, a diverse food retail landscape is associated with higher adult obesity rates, after controlling for other variable. A county with more natural amenities is associated with a .03% decrease in adult obesity rates. As the percentage of adults who are not physically active increases by one unit, the obesity rate increases by .35 %. The percentage of physically inactive adults had the largest effect on obesity prevalence at the county-level.

The variance inflation factor (VIF) was used to determine test for multicollinearity between the independent variables. The VIFs did not exceed the common threshold of 10, therefore there is no indication that multicollinearity is biasing the results (Menard 2001). Overall, the model is well-specified, and explains 60 % of the variance in the obesity rates among adults at the county level. The AIC for the global regression model is 14,945.36.

Further analysis is needed to determine spatial heterogeneity in the outcome variables, and whether a local, nuanced model is more robust compared to the OLS model. GWR 3.0 software was used to compute the results for the local regression models. The

GWR 5-number summary and Monte Carlo significance tests for spatial heterogeneity of parameter test results are featured in Table 4. The Monte Carlo significance test in the model indicates that most of the parameters in the model, with the exception of both of the food access variables, are non-stationary across space. In other words, the effect of food access does not significantly vary across the United States. The AIC in the GWR model is 14,230.50 and the adjusted r-square indicates that the model explains 71 % of the variance, both of which confirms that the GWR model is a better fit than the global model. The AIC determined bandwidth in the GWR model is 257 counties. That is, each local GWR model is estimated on data from 257 counties, which is less than 10 % of the 3,108 counties in the global analysis.

Figure 2 maps the results of local r-square values. The global model in the linear regression analysis reported a single r-square value of .60. The GWR results show that local r-square range from .20 to .86. Hence, the model fits well in some parts of the country, while the model explains substantially less of the variance in other areas. For example, the model fits best in Central and South Florida, parts of South Carolina, Georgia, and Alabama, and a band of counties in the West. Comparatively, the covariates used in the analysis fails to adequately predict obesity prevalence in much of the Midwest and in many counties in Appalachia. This indicates that additional covariates are needed to better understand determinants of county-level adult obesity rates in these counties.

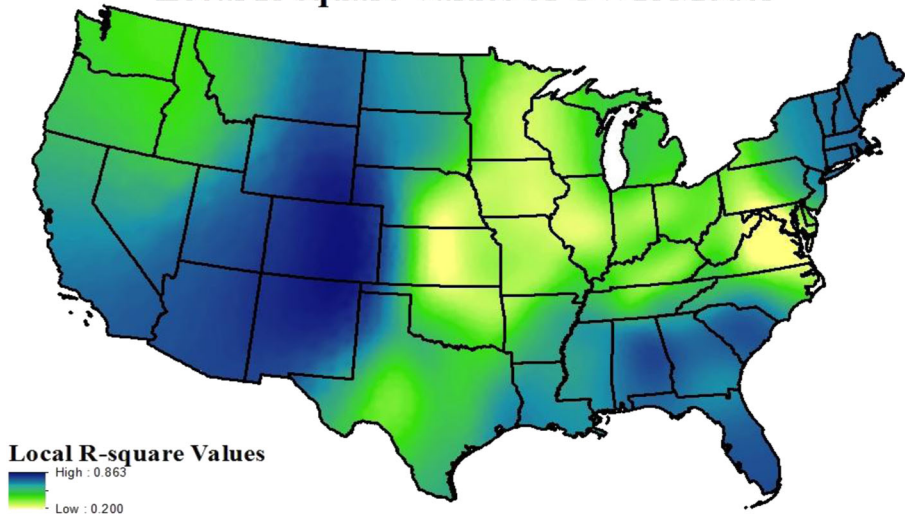
Table 4 shows that the effect of the covariates varies greatly across the study area. However, it is essential to map the local parameter estimated to observe where there is

**Table 4** Geographically Weighted Regression 5 number parameter summary results and Monte Carlo significance test for spatial variability of parameters ( $N=3,108$ )

	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	Monte Carlo Testing Status
Intercept	7.601	21.204	23.987	27.390	35.745	Non-stationary***
% Rural	-0.051	-0.015	-0.006	0.007	0.033	Non-stationary***
Population Density	-0.005	-0.001	-0.001	-0.001	0.046	Non-stationary***
% Black	-0.920	0.004	0.053	0.107	0.694	Non-stationary***
% Hispanic	-0.278	-0.028	0.001	0.027	0.265	Non-stationary***
% Asian	-1.788	-0.206	-0.065	0.158	1.539	Non-stationary***
% Native Americans	-1.613	0.054	0.128	0.197	4.319	Non-stationary*
% Poverty	-0.216	-0.031	-0.008	-0.060	0.176	Non-stationary***
% College Education	-0.290	-0.128	-0.080	-0.037	0.105	Non-stationary***
% HH with no car, > 1 miles to grocery store	-0.242	-0.069	-0.007	0.107	0.523	Stationary
Food retail landscape	-6.865	-0.513	0.404	1.082	6.446	Stationary
Amenity Scale	-1.182	-0.359	-0.168	-0.007	0.596	Non-stationary***
% Adult Physical Inactivity	-0.009	0.174	0.288	0.343	0.678	Non-stationary***
Adjusted R-Square	0.715					
Akaike Information Criterion	14230.50					

\* $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$

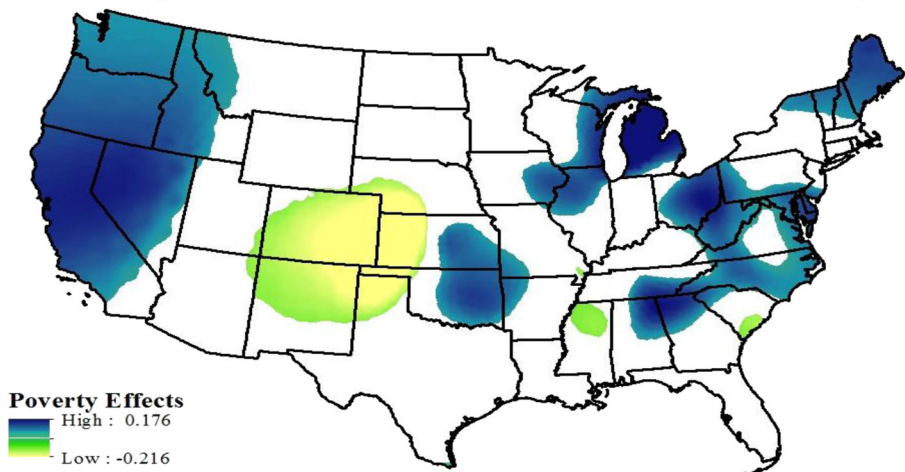
### Local R-square Values of GWR Model



**Fig. 2** (Color Map) Local r-square of GWR model (Black and White Map) Local r-square of GWR model

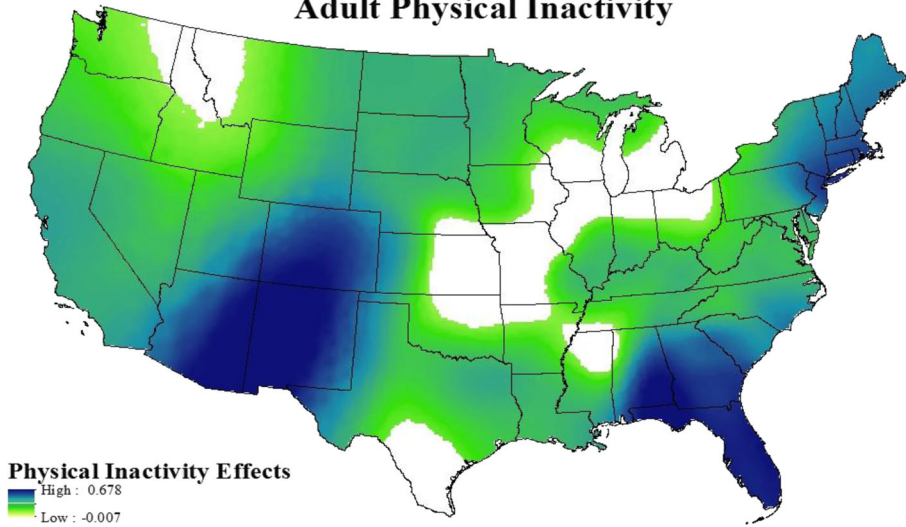
significant spatial heterogeneity between the independent and dependent variables. Figure 3 and Fig. 4 reveal the GWR results of poverty and adult physical inactivity, respectively (maps of additional non-stationary covariates are available upon request). Fig. 4 shows the effect of poverty on obesity prevalence varies across the study area. The solid white areas indicate that the parameter estimated do not significant vary across the study area. The shaded areas display where the spatially varying effects are significant (Matthews and Yang 2012). For example, the magnitude of the effect of poverty on adult obesity is positive and larger in parts of the Pacific Northwest, the

### Significant Local Parameter Estimates for Poverty



**Fig. 3** (Color Map) Significant local parameter estimates of the effect of poverty on adult obesity prevalence (Black and White Map) Significant local parameter estimates of the effect of poverty on adult obesity prevalence

## Significant GWR Estimates of Adult Physical Inactivity



**Fig. 4** (Color Map) Significant local parameter estimates of the effect of physical inactivity on adult obesity prevalence (Black and White Map) Significant local parameter estimates of the effect of physical inactivity on adult obesity prevalence

D.C. Metro area; West Virginia; much of Ohio, North Carolina, Michigan; and in several New England States. However, the effect of poverty on obesity is negative in the area surrounding the Texas Panhandle to Southwest Nebraska and much of Colorado and New Mexico. These results suggest that targeting poverty as a social policy tool to combat obesity will be more effective in some areas than others. Figure 4 reports the effects of adult physical inactivity across the study area. There is significant geographic variation in the relationship between physical inactivity and obesity across most of the study area. GWR results show that increasing physical activity among residents will yield positive, but varying, results across the study area. The effect is particularly pronounced in counties in parts of Georgia and Alabama, and much of Florida, Arizona, Colorado, and New Mexico.

### Discussion

This study extends upon traditional aspatial regression analyses by employing GWR to uncover geographic variation in obesity in the United States. This approach accounts for the unique attributes of counties and the spatially varying relationships between the outcome and predictor variables across the continental United States. The results from this study yield several interesting findings, many of which capture the salience of place. There are clear social disparities in obesity rates across the United States, and this study confirms that there is also explicit spatial patterning of the epidemic. This study reveals that traditional regression results (i.e., OLS) does not completely capture the complete picture of the relationship between obesity and ecological variables at the county-level. The OLS model explained 60 % of the variance in obesity rates between

counties in the United States. Thus, the OLS model was well-specified, and given the specificity of the OLS model, it is reasonable to conclude that the GWR results are also robust (i.e., the explained variance ranged from 20 to 83 %).

To effectively address high obesity rate in the United States, policymakers must have an understanding of context-specific determinants of obesity at the local level. Geographically weighted regression provides an augmented awareness of what variables should be targeted to effectively and efficiently combat the epidemic. This study reveals several insights and potential policy prescriptions to combat the epidemic. For example, most socio-demographic and environmental variables in the study are salient predictors of obesity in the United States, and many of these relationships varied across the study area – particularly poverty. Beyond the budgetary constraints that individuals may face when making food choices (Drewnowski and Specter 2004), the local context characterized by high rates of poverty may perpetuate and reinforce norms around poor dietary behaviors and sedentary lifestyles. Addressing community norms circumventing physical activity are perhaps the most salient factor to achieve decreased obesity rates, especially in the Northeast, the Southeast, and the Southwest. Effective community-based interventions to improve physical activity among adults could promote population-wide health and substantially reduce the burden of obesity in the United States.

While the OLS model yields a single r-square value for the entire study area, the range of r-square values calibrated by the GWR results confirms that “one-size fits all” policy strategy is not the most appropriate approach to address obesity at the national level. A subnational or regional strategy to address obesity is likely more efficient where stimuli may provoke a similar response. The GWR model reveals that a similar intervention to combat obesity will be virtually equally effective in extreme Northeast, many bounded counties across the Plains and Southwest, and the Deep South. The results also highlight the need for a deeper understanding of what is driving the obesity epidemic in parts of Appalachia and the Midwest. Perhaps ethnographic studies can better contextualize what is contributing to high rates of adult obesity in these regions.

While this study has several strengths, it is not without limitations. The analysis is cross-sectional, and does not contribute to the understanding of how macro-level processes are responsible for increased prevalence of obesity overtime. There are also potential concerns about using aggregate level data to make generalizations about individuals or places. This study does not account for the possible selection of people into counties. Thus, the spatial patterns observed may be an artifact of the preferences of individuals. Also, the modifiable area unit problem (MAUP) is well-known in the spatial analysis literature (Fotheringham and Wong 1991; Jelinski and Wu 1996). It is possible that different parameters and inferences may be drawn depending on the spatial scale used in the unit of analysis, or when aggregate units are divided into different zonal areas (Openshaw and Taylor 1981). However, counties, the unit of analysis used in this study, are less arbitrary than other spatial units (e.g., census tracts and block groups). Counties are meaningful in the sense that they serve as political units, and many public health services and policy interventions can feasibly be administered at the county level. County-level public health services have become more pertinent for policy reasons in the United States as government services have become more decentralized since the 1980s (Lobao and Kraybill 2005). Furthermore, health data are commonly collected at the county-level. Therefore smaller scales of

analysis, such as neighborhood-level health data is less available, and is more difficult to replicate at a large scale across time and place (McLaughlin et al. 2007). However, the size and shape of counties vary across the United States; therefore, a county-level analysis may be less meaningful in some regions of the United States, than in other regions (Goodchild 2011). Nevertheless, county-level boundaries are relatively stable compared to boundaries at smaller scales, and county jurisdictions capture daily activities, as well as economic and social ties (Matthews 2011). Other limitations include multicollinearity among local estimates, which can be problematic when using GWR (Wheeler and Tiefelsdorf 2005). GWR output can also report a wide range of parameter estimates that can be cumbersome to interpret and spatial heterogeneity can be overstated (Farber and Páez 2007). Additional variables such as local taxes, local policies, and access to recreational facilities that capture the social and economic context, as well as the built environment, may also increase our understanding of spatial variation in obesity rates.

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

## References

- Ali, K., Partridge, M. D., & Olfert, M. R. (2007). Can geographically weighted regressions improve regional analysis and policy making? *International Regional Science Review*, *30*, 300–329.
- Allison, P. D. (1999). *Multiple regressions: a primer*. Thousand Oaks, CA: Pine Forge Press.
- Anselin, L. (1988). Do spatial effects really matter in regression analysis? *Papers of the Regional Science Association*, *65*, 11–34.
- Arcaya, M., Brewster, M., Zigler, C. M., & Subramanian, S. V. (2012). Area variations in health: a spatial multilevel modeling approach. *Health Place*, *18*(4), 824–831.
- Baker, E. A., Schootman, M., Barnidge, E., & Kelly, C. (2006). The role of race and poverty in access to foods that enable individuals to adhere to dietary guidelines. *Preventing Chronic Disease*, *3*(3), A76.
- Berrigan, D., & Troiano, R. P. (2002). The association between urban form and physical activity in U.S. adults. *American Journal of Preventive Medicine*, *23*, 74–79.
- Bitter, C., Mulligan, G., & Dallerba, S. (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.
- Bodor, J., Rice, J., Farley, T., Swalm, C., & Rose, D. (2010). The association between obesity and urban food environments. *Journal of Urban Health*, *87*, 771–781.
- Bowman, S. A., & Vinyard, B. T. (2004). Fast food consumption of U.S. Adults: impact on energy and nutrient intakes and overweight status. *Journal of the American College of Nutrition*, *23*, 163–168.
- Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression-modelling spatial Non-stationarity. *Journal of the Royal Statistical Society Series D (The Statistician)*, *47*, 431–443.



- Chalkias, C., Papadopoulos, A. G., Kalogeropoulos, K., Tambalis, K., Psarra, G., & Sidossis, L. (2013). Geographical heterogeneity of the relationship between childhood obesity and socio-environmental status: empirical evidence from Athens, Greece. *Applied Geography*, 37, 34–43.
- Charlton, M., Fotheringham, S., Brunsdon, C., 2003. GWR 3. Software for geographically weighted regression. *Spatial Analysis Research Group, Department of Geography, University of Newcastle upon Tyne, England*.
- Cossman, J. S., Cossman, R. E., James, W. L., Campbell, C. R., Blanchard, T. C., & Cosby, A. G. (2007). Persistent clusters of mortality in the United States. *American Journal of Public Health*, 97, 2148–2150.
- Cummins, S., Curtis, S., Diez-Roux, A. V., & Macintyre, S. (2007). Understanding and representing 'place' in health research: a relational approach. *Social Science & Medicine*, 65(9), 1825–1838.
- Diez-Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, 1186, 125–145.
- Dorling, D. (2001). How much does place matter. *Environment and Planning A*, 33(8), 1335–1369.
- Drewnowski, A., & Specter, S. E. (2004). Poverty and obesity: the role of energy density and energy costs. *The American Journal of Clinical Nutrition*, 79(1), 6–16.
- Ezzat, M., Friedman, A. B., Kulkarni, S. C., & Murray, C. J. L. (2008). The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States. *PLoS Medicine*, 5(4), 66.
- Farber, S., & Páez, A. (2007). A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems*, 9(4), 371–396.
- Finkelstein, E. A., Trogon, J. G., Cohen, J. W., & Diez, W. (2009). Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Affairs*, 28(5), 822–831.
- Flegal, K. M., Carroll, M. D., Ogden, C. L., & Johnson, C. L. (2002). Prevalence and trends in obesity among US adults, 1999–2000. *JAMA, the Journal of the American Medical Association*, 288, 1723–1727.
- Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA, the Journal of the American Medical Association*, 307(5), 491–497.
- Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in statistical analysis. *Environment and Planning*, 23, 1025–1044.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.
- Fraser, L. K., Clarke, G. P., Cade, J. E., & Edwards, K. L. (2012). Fast food and obesity: a spatial analysis in a large United Kingdom population of children aged 13–15. *American Journal of Preventive Medicine*, 42(5), 77–85.
- Geronimus, A. T., Bound, J., Waidmann, T. A., Hillemeier, M. A., & Burns, P. B. (1996). Excess mortality among blacks and whites in the United States. *New England Journal of Medicine*, 335, 1552–1558.
- Goodchild, M. F. (2011). Formalizing place in geographic information systems communities, neighborhoods, and health. In L. M. M. Burton, S. A. P. Matthews, M. Leung, S. P. A. Kemp, & D. T. T. Takeuchi (Eds.), *Social disparities in health and health care* (pp. 21–33). New York: Springer.
- House, J. S., Schoeni, R. F., Kaplan, G. A., & Pollack, H. (2009). *The health effects of social and economic policy: the promise and challenge for research and policy*. Ann Arbor, Michigan: The National Poverty Center.
- Hu, F. B., Li, T. Y., Colditz, G. A., Willett, W. C., & Manson, J. E. (2003). Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *JAMA, the Journal of the American Medical Association*, 289, 1785–1791.
- Iceland, J. 2004. The multigroup entropy index (also known as Theil's H or the information theory index). US Census Bureau. [http://www.census.gov/hhes/www/housing/housing\\_patterns/multigroup\\_entropy.pdf](http://www.census.gov/hhes/www/housing/housing_patterns/multigroup_entropy.pdf).
- Jackson, J., Doescher, M. P., Jerant, A. F., & Hart, L. G. (2005). A national study of obesity prevalence and trends by type of rural county. *The Journal of Rural Health*, 21, 140–148.
- Jelinski, D., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology* 11, 129–140
- Kearns, R. A., & Gesler, W. M. (1998). *Putting health into place: landscape, identity, and well-being*. Syracuse, New York: Syracuse University Press.
- Kim, D., Subramanian, S. V., Gortmaker, S. L., & Kawachi, I. (2006). US state- and county-level social capital in relation to obesity and physical inactivity: a multilevel, multivariable analysis. *Social Science & Medicine*, 63(4), 1045–1059.
- Larson, N. I., Story, M. T., & Nelson, M. C. (2009). Neighborhood environments: disparities in access to healthy foods in the U.S. *American Journal of Preventive Medicine*, 36, 74–81.

- Lean, M. (2010). Health consequences of overweight and obesity in adults. In D. Crawford, R. W. Jeffery, K. Ball, & J. Brug (Eds.), *Obesity epidemiology: from Aetiology to public health* (pp. 43–58). Oxford: Oxford University Press.
- Ledikwe, J. H., Blanck, H. M., Khan, L. K., Serdula, M. K., Seymour, J. D., Tohill, B. C., et al. (2006). Dietary energy density is associated with energy intake and weight status in US adults. *The American Journal of Clinical Nutrition*, *83*, 1362–1368.
- Lobao, L., & Kraybill, D. S. (2005). The emerging roles of county governments in metropolitan and nonmetropolitan areas: findings from a national survey. *Economic Development Quarterly*, *19*(3), 245–259.
- Lovasi, G. S., Hutson, M. A., Guerra, M., & Neckerman, K. M. (2009). Built environments and obesity in disadvantaged populations. *Epidemiologic Reviews*, *31*, 7–20.
- Ludwig, D. S., Peterson, K. E., & Gortmaker, S. L. (2001). Relation between consumption of sugar-sweetened drinks and childhood obesity: a prospective, observational analysis. *The Lancet*, *357*, 505–508.
- Lynch, J. W., & Kaplan, G. A. (1997). Understanding How inequality in the distribution of income affects health. *Journal of Health Psychology*, *2*, 297–314.
- Macintyre, S., Maciver, S., & Sooman, A. (1993). Area, class and health: should we be focusing on places or people? *Journal of Social Policy*, *22*(02), 213–234.
- Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: how can we conceptualise, operationalise and measure them? *Social Science & Medicine*, *55*, 125–139.
- Matthews SA. Spatial polygamy and the heterogeneity of place: studying people and place via egocentric methods. In: Burton LM, Kemp SP, Leung M, Matthews SA, Takeuchi DT, editors. *Communities, neighborhoods, and health: Expanding the boundaries of place*. Springer; 2011. pp. 35–55.
- Matthews, S. A., & Yang, T. C. (2012). Mapping the results of local statistics. *Demographic Research*, *26*(6), 151–166.
- McLaughlin, D. K., Stokes, C. S., Smith, P. J., & Nonoyama, A. (2007). Differential mortality across the United States: the influence of place-based inequality. In L. M. Lobao, G. Hooks, & A. R. Tickamyer (Eds.), *The sociology of spatial inequality*. Albany: State University of New York Press.
- Menard, S. (2001). *Sage publications*. Incorporated.: Thousand Oaks, CA. Applied logistic regression analysis.
- Mitchell, R. (2001). Multilevel modeling might not be the answer. *Environment and planning A*, *33*(8), 1357–1360.
- Morland, K., Wing, S., Diez-Roux, A., & Poole, C. (2002). Neighborhood characteristics associated with the location of food stores and food service places. *American Journal of Preventive Medicine*, *22*, 23–29.
- Murray, C. J. L., Sandeep, C. K., Michaud, C., Tomijima, N., Bulzacchelli, M., Iandiorio, T. J., et al. (2006). Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, *3*, 1513–1524.
- Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among US children and adolescents 1999–2010. *JAMA, the Journal of the American Medical Association*, *307*(5), 483–490.
- Openshaw S, Taylor PJ. 1981. The modifiable areal unit problem. In: Wrigley, N., Bennett, R., Kegan, P. (Eds), *Quantitative Geography: A British View*. London. pp. 60–69.
- Pacione, M. (1984). Evaluating the quality of the residential environment in a high-rise public housing development. *Applied Geography*, *4*, 59–70.
- Popkin, B. M. (2008). *The world is Fat—the fads, trends, policies, and products that Are fattening the human race*. New York, NY: Avery-Penguin Group.
- Procter, K. L., Clarke, G. P., Ransley, J. K., & Cade, J. (2008). Micro-level analysis of childhood obesity, diet, physical activity, residential socioeconomic and social capital variables: where are the obesogenic environments in Leeds? *Area*, *40*(3), 323–340.
- Ramsey, P. W., & Glenn, L. L. (2002). Obesity and health status in rural, urban, and suburban southern women. *Southern Medical Journal*, *95*, 666–671.
- Rushton, G., Armstrong, M. P., Gittler, J., Greene, B. R., Pavlik, C. E., West, M. M., & Zimmerman, D. L. (Eds.). (2010). *Geocoding health data: the use of geographic codes in cancer prevention and control, research and practice*. CRC Press
- Sallis, J. F., & Glanz, K. (2009). Physical activity and food environments: solutions to the obesity epidemic. *Milbank Quarterly*, *87*, 123–154.
- Schulze, M. B., Manson, J. E., Ludwig, D. S., Colditz, G. A., Stampfer, M. J., Willett, W. C., et al. (2004). Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA, the Journal of the American Medical Association*, *292*, 927–934.
- Schuurman, N., Peters, P. A., & Oliver, L. N. (2012). Are obesity and physical activity clustered? a spatial analysis linked to residential density. *Obesity*, *17*(12), 2202–2209.

- Shaw, M., Dorling, D., & Mitchell, R. (2002). *Health, place, and society*. Harlow: Pearson Education.
- Thiele, S., & Weiss, C. (2003). Consumer demand for food diversity: evidence for Germany. *Food Policy*, 28(2), 99–115.
- Tobler, W. (2004). On the first law of geography: a reply. *Annals of the Association of American Geographers*, 94(2), 304–310.
- Tunstall, H. V. Z., Shaw, M., & Dorling, D. (2004). Places and health. *Journal of Epidemiology and Community Health*, 58(1), 6–10.
- Wang, Y., & Beydoun, M. A. (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.
- Ward, M. D., & Gleditsch, K. (2008). *Spatial regression models*. London: Sage.
- Wen, T. H., Chen, D. R., & Tsai, M. J. (2010). Identifying geographical variations in poverty-obesity relationships: empirical evidence from Taiwan. *Geospatial Health*, 4(2), 257–265.
- Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, 7(2), 161–187.