THE ROLE OF RESIDENTIAL SEGREGATION AND VULNERABILITY ON POTENTIAL ACCESSIBILITY TO PRIMARY CARE SERVICES IN TEXAS

Introduction

Spatial accessibility can affect health outcomes such as breast cancer diagnosis (Dai, 2010), health care services (Gaskin, Dinwiddie, Chan, & McCleary, 2012a; Gaskin, Price, D.T., & LaVeist 2009), primary care (Gaskin, Dinwiddie, Chan, & McCleary, 2012b; Luo & Qi, 2009), surgical care (Hayanga et al., 2009), and pediatric providers (Guagliardo, Ronzio, Cheung, Chacko, & Joseph, 2004). Scholars have concluded that reduced number of physicians and health clinics increases populations' rates of late disease diagnoses. Traditionally, spatial accessibility to health care has been measured based on the relationship between *realized* care and the *potential* for health care use -- when supply (health care system) and demand (population) coexist in space and time. Five dimensions of barriers between potential and realized care have been identified: availability, accessibility, affordability, acceptability, and accommodation (Guagliardo, 2004; Guagliardo et al., 2004; Penchansky & Thomas, 1981). Spatial accessibility to health services is influenced by factors such as the number of physicians, number of people living in the region, urban or rural community type, as well as socio-demographic characteristics of the place and the population inhabiting it.

Previously, scholars have measured accessibility to primary care using either a distance-based measure or the ratio of physicians to population. Usually, when considering a distance-based measure (e.g., travel distance, gravity models, or 2 Step Floating Catchment Area – 2SFCA methods), the results inform us about physician shortages within an area; however, other covariates are often not considered (Langford & Higgs, 2006; McGrail, 2009; Polzin, Borges, & Coelho, 2014). When the outcome is defined as the ratio of physician to population, researchers usually control for other socioeconomic and demographic characteristics, which enables them to assess the relative effects of these characteristics on potential accessibility to primary care (Guagliardo et al., 2004), but more often than not, analyses using a gravity or 2SFCA methodology often do not attempt to go beyond basic mapping or spatial cluster analysis of the accessibility measure. This is seen as a shortfall of many of these analyses, and given the context of the present analysis, the state of Texas; it is particularly interesting to consider the factors shaping the spatial access to healthcare.

As the demographics of Texas shift and grow increasingly diverse, the unique characteristics and different needs that minority groups experience should become more salient. In 2013 in Texas, Hispanic or Latino groups represented 41% of the population, non-Hispanic blacks12%, whites 41%, and other racial/ethnic groups 6% compared to a national level of 17% Hispanics, 12% non-Hispanic black, 62% whites, and 8% other racial/ethnic groups (Kaiser Family Foundation, 2013). From a health policy perspective, understanding the characteristics and needs of the increasingly vulnerable Hispanic groups would help inform and design adequate public policies that might better serve the needs of these groups. This study focuses on primary care physicians because of their key role in the success of preventive care and because accessibility to these providers represents an important indicator of overall health (Gautam, Li, & Johnson, 2014; Luo & Wang, 2003)

This study will employ a spatial measure of potential accessibility to primary health care physicians using the 2SFCA method. Then, Ordinary Least Squares (OLS), weighted OLS, and spatial regression models (spatial lag and error specifications) will be used to determine the associations between the spatial access, residential segregation, poverty, education, unemployment, and type of place on potential spatial accessibility to primary care in the state of Texas.

Literature Review

Residential segregation has been extensively studied by demographers, sociologists, and geographers when looking at a variety of health outcomes. Several studies examine residential segregation and potential spatial accessibility measured by the physician-to-population ratio and concentration of zip codes (Gaskin et al., 2012a, 2012b; Gaskin et al., 2009). Residential segregation was also associated with concentrated poverty (Acevedo-Garcia, Lochner, Osypuk, & Subramanian, 2003; Bell, Zimmerman, Almgren, Mayer, & Huebner, 2006; Dai, 2010; White & Borrell, 2011; Williams & Collins 2001). Also, residential segregation limits access to education and job opportunities (Williams & Jackson, 2005) increasing the risk of high school dropout rates and low job availability, which in turn translate into lower income levels and ability to purchase market resources. Fewer jobs and concentrated poverty create unsafe neighborhoods by increasing crime, particularly violent crime, in an area.

For example, researchers found that predominantly African American zip codes were negatively associated with the availability of physician services (Gaskin et al., 2012a), and that while Asian and Hispanic zip codes had a positive relationship with residential segregation, that the Hispanics and African American populations experience poorer access to health services compared with Whites (Williams & Collins 2001). Furthermore, an increase in the Hispanic and African American population in already highly segregated counties was associated with a decline in the availability and use of surgical services (Gaskin et al., 2012b; Hayanga et al., 2009).

However, Kwan suggested that when studying residential segregation, health outcomes, and accessibility, considering the place and space where individuals spend most of their time (e.g., work and daily interactions) is more important than place of residence (Kwan, 2013).

Potential spatial accessibility and multi-group residential segregation are described from the *risk regulators* perspective (Glass & McAtee, 2006). The risk regulators perspective asserts that social conditions are not disease causes as previously asserted (Link & Phelan, 1995; Williams & Collins 2001) but are instead "relatively stable feature[s] of a particular patch of the social and built environments" that shape, motivate, and induce behavioral risk factors that further lead to disease (Glass & McAtee, 2006, p. 1659). The main idea of the risk regulators theory is to highlight that social conditions are variables that influence health outcomes; however this influence is second-order and not causal. Contrary to causal risk factors, risk regulators operate through different pathways over time and space. Space – or geography – is an important factor, and therefore should be included in analyses in which the outcome tends to vary over space. When examining access to primary care, space is important because racial/ethnic minorities as well as individuals in lower SES groups are spatially segregated (Jargowksy, 1997; Kwan, 2013; Massey & Fong, 1990; Massey & Shibuya, 1995). Highly segregated places are characterized by significantly more contextual risk factors, which usually put individuals in vulnerable positions leading to risky behaviors, isolation from socioeconomic support and health services, and barriers to upward mobility (Do et al., 2008; Fitzpatrick & LaGory, 2003).

The *community determinants* model presented by Shi and Stevens (2010) highlights that certain place characteristics create higher vulnerability among populations living in those places compared to populations that do not present these characteristics. Vulnerable populations register higher rates of morbidity and mortality, and are at greater risk of poor mental, physical, and social health than their less-vulnerable counterparts. Aggregate-level characteristics such as socio-demographic characteristics and health needs are likely to increase the health risks of particular segments of the population (Shi & Stevens, 2010). In particular, characteristics such as socioeconomic status (lower educational attainment and concentrated poverty) and type of place (metro/non-metro) can act as barriers to accessing primary care (Shi & Stevens, 2010).

According to the *risk regulators* perspective, residential segregation does not directly influence spatial accessibility, but rather has a second-order effect. For this study, specifically, counties in Texas that are least exposed to residential segregation (i.e., are most integrated) may have a different average level of educational attainment compared with more segregated counties. Residential segregation can further affect spatial accessibility of primary care resources. Similarly, urban residents have more opportunities, especially with respect to primary care physicians, compared to rural residents that are not only more spatially isolated but may also have limited resources.

Being in a vulnerable position (e.g., living in a neighborhood with concentrated poverty) and living in a non-metropolitan area increases the burden on the populations seeking out medical care. Rural populations encounter greater spatial limitations for supporting adequate health due to the lack of infrastructure and the relative distance to health care providers. For example, Probst and colleagues found that distance to healthcare facilities did not vary by race, but by location in rural/urban environment. Rural residents and African Americans populations experience greater travel burdens than urban residents or whites respectively when seeking medical care (Probst, Laditka, Wang, and Johnson, 2007). This issue is more pronounced for minority populations since they have greater difficulties finding culturally adequate doctors in areas that are already medically underserved (Landrine & Corral, 2009). Also, residents of rural populations tend to have lower socioeconomic status (primarily due to lower levels of educational attainment) when compared with their urban counterparts. It is therefore important to consider type of place, when examining spatial accessibility to primary care services.

Fewer studies have explored access to primary care in Texas counties and the negative or protective effect of socio-demographic characteristics and how it might affect residents' accessibility to primary care physicians. One of the studies found that care-related Emergency Department (ED) visit rates were strongly correlated with the proportion of the population below the poverty line and the uninsured rate, even though the authors encourage using other covariates when looking at access to healthcare that they did not consider in their study (Begley, Vojvodic, Seo, & Burau, 2006). Various studies have analyzed the relationship between residential segregation and access to primary care, however in these studies residential segregation was never measured by the multi-group segregation index (Gaskin et al., 2012a, 2012b; Gaskin et al., 2009; Williams & Collins 2001). Demographic shifts and increased diversity in Texas requires researchers to go beyond the two-group segregation measures as defined by Douglas and Massey to identify the unique characteristics and special needs of minority groups (Massey & Denton 1988). So far, no study examined the association between residential segregation measured by the Theil index (a multi-group segregation index) and spatial accessibility to primary care measured with the 2SFCA method.

This research contributes to the existing literature in two ways. First, it considers the effects of residential segregation, SES (education and poverty), unemployment, and rural/urban status along with a spatial accessibility index computed by the 2SFCA. Few studies have considered these covariates when looking at spatial accessibility to primary care. Second, it will employ a spatial econometric perspective and use spatial regression models for the analysis. The risk regulators perspective is used to present a possible mechanism through which neighborhood and population characteristics affect potential spatial accessibility to primary health care services. Also, to explain why for this particular study spatial error models might be a better choice than spatial lag models.

Based on the above discussion, this research will address the following research questions:

1. Does the effect of multi-group residential segregation in areas of poor socioeconomic status and in rural areas has a negative effect on potential spatial accessibility?
2. Do residential segregation, education, poverty, unemployment, and type of place have similar effects on spatial accessibility to primary care indices? Does this effect vary according to the distance considered?

Data and Methods

The data for this study come from several sources. First is the Licensed Physician database for 2011, purchased from the Texas Medical Board (TMB) store. The 2011 Licensed Physician database contains information about the demographic characteristics of physicians as well as information about their practice such as license number, mailing and practice address, medical school attended, graduation year, and primary and secondary specialty, among others. The primary physician address is used for geocoding in ArcGIS 10.3. To provide more accurate information as well as account for possible errors in primary care physicians' addresses, any missing or incomplete practice address is replaced with their mailing address. The next source is the 2010 Summary File 1 (SF1) at the census tract level, and data from American Fact Finder (DP2 – social characteristics- and DP3 – economic characteristics) summary tables at the tract level for the state of Texas from the 2005-2009 5 year ACS summary file (Census Bureau, 2010). No studies looked at residential segregation as measured by a multi-group index in conjunction with the effects of socio-demographic characteristics on accessibility to primary care as measured by the 2SFCA method. Increasing diversity and demographic shifts in Texas should make us alert to the unique characteristics and special needs of minority groups. To calculate the multi-group residential segregation, the Theil Entropy index, I will aggregate the 2010 SF1 data at the block group level up to the census tract level. Having the primary care physician locations and the weighted population centroids at the Census tract level (Census, 2010) allows for the calculation of the potential spatial accessibility using the 2SFCA method.

Measures of Access to Primary Care

An index of potential accessibility to primary care for 2011 is calculated using the 2SFCA method. The primary care specialties include: family practice, family medicine, general preventive medicine, gynecology, internal medicine (preventive medicine), and mammography. The 2SFCA method was initially used for determining physician-shortage area designations by Luo (2003), but subsequent studies have demonstrated the applicability of this method to measuring potential health care accessibility (Luo & Whippo, 2012). The 2SFCA, a special case of the gravity model, has most of the advantages of a gravity model yet is simple to interpret since it generates a special form of physician-to-population ratio (Luo, 2003; Luo & Qi, 2009; Luo & Whippo, 2012). The 2SFCA method reveals detailed spatial variation within larger administrative areas (counties) and considers potential interactions between individuals and physicians across administrative borders (Luo, 2003).

The first step of this method requires that for each physician location j, all population locations k (the sum of each population centroids) that are within the distance d_0 from location j

(physician locations), to compute the physician-to-population ratios R_j (by dividing the sum of physicians S_j by P_k the sum of population centroids) within the catchment area (equation 1):

$$R_j = \frac{S_j}{\sum_{k \in \{d_{kj} \leq d_0\}} P_k} \qquad (1)$$

In the second step, for each population location *i*, (P_k - population centroids) search all physician locations *j* within the distance d_0 and sum up the physician-to-population ratios (from previous step), R_j:

$$A_i^F = \sum_{j \in \{d_{ij} \leq d_0\}} R_j = \sum_{j \in \{d_{ij} \leq d_0\}} \frac{S_j}{\sum_{k \in \{d_{kj} \leq d_0\}} P_k} \qquad (2)$$

Also, $\sum_{j \in \{d_{ij} \leq d_0\}} R_j$ - reads as the sum within *j* (physician locations) where the distance between population location *i* and physician locations *j* d_{ij} is within the desired distance d_0, then sum up the physician-to-population ratios calculated in step (1) - R_j.

The above formulas consider the physician location *j* (practice or mailing address) at the census tract level while the centroid of the census tract measures the populations k. After primary care physicians are geocoded in ArcGIS 10.3, different catchments (distances from each physician location to population Census centroids) are calculated in Python (spatial accessibility indices are calculated for 10 miles, 15 miles, 30 miles, 45 miles, and 60 miles). Given the high computational operations and memory needed that ArcGIS could not handle, Python was chosen to calculate the spatial accessibility index for the entire state of Texas. Unlike ArcGIS, which uses a straight-line formula to calculate the distance between two points (Euclidian distance), while calculating the spatial accessibility index using Python, the Haversine approach was used. This approach takes the Earth's spherical shape into consideration when calculating the distance (Westra, 2013).

Socio-demographic Characteristics

Residential segregation is the primary independent variable for this aim. Though Massey and Denton (1988) encourage the use of "a battery of indices rather than one single index", due to the multidimensional phenomenon of segregation, previously mentioned dimensions (evenness, exposure, concentration, centralization, and clustering) are designed to measure segregation among two population groups (usually comparing Non-Hispanic whites with blacks or Hispanics (Massey & Denton 1988; Reardon & Firebaugh, 2002). Given the high racial diversity in the state of Texas, a multi-group index is more suitable for an adequate description of the complex patterns of racial segregation and integration. As previously mentioned, multi-group residential segregation as measured by the Theil Entropy index will be calculated at the Census tract level. Calculation of Theil index is done in two steps. First, the Entropy (E) index will be calculated in which E is seen as a measure of population "diversity" (Reardon & Firebaugh, 2002). Then, the multi-group index will be calculated using the Entropy index calculated in Step 1. Overall, Theil information index (H) measures how evenly groups are distributed among organizational units (Reardon & Firebaugh, 2002).

$$E = \sum_{k=0}^{n} \pi_m \ \ln \frac{1}{\pi_m} \qquad (3)$$

$$H = \sum_{m=1}^{M} \sum_{j=1}^{J} \frac{t_j}{TE} \pi_{jm} \ ln \frac{\pi_{jm}}{\pi_m} \qquad (4)$$

First, the Entropy Index (E) computed according to equation (3) can be identified as a measure of diversity. If all individuals are members of a single group, we can conclude that there is no diversity among populations. If individuals are evenly distributed among the M groups (π_m = 1/M for all *m*), then the index is maximized. The Information Theory Index is calculated based on equation (4). This index captures "evenness" to the extent that groups are evenly distributed among organizational units (Reardon & Firebaugh, 2002).

Demographic characteristics will be measured at the Census tract level by the following variables: the poverty rate, proportion of the population age 25 and over without a high school education, unemployment rate, and whether the Census tract is located in a motropolitan area or not. A metro/non-metro variable will be created based on the U.S. Department of Agriculture Economic Research Service (ERS) Rural-Urban Continuum Codes classification based on the population size and degree of urbanization (USDA, 2011). Metropolitan counties are counties with a population between 250,000 and/more 1 million inhabitants, while non-metro counties have populations between 2,500 and/more 20,000 or more inhabitants

Measures of Spatial Autocorrelation

First, an exploratory analysis will be conducted to determine possible clustering in the data. Moran's I statistic will be used to detect possible global autocorrelation for all dependent and independent variables. The interpretation of the Moran's I is similar to the interpretation of a correlation coefficient. For example, if the observed value is greater than the expected value, then the specific observation is surrounded by neighbors with similar values; otherwise if the observed value is smaller than the expected value, that specific observation is surrounded by neighbors with different values.

To determine whether global autocorrelation is present for the residuals of the OLS and weighted OLS models, Global Moran's I statistics is used. Usually the Global Moran's I tool offers a single summary spatial correlation measure that can be interpreted as a standard correlation. For example, if the Moran's I value is 0.5 that would indicate a highly significant degree of global autocorrelation in the OLS or weighted OLS models residuals. Based on the Moran's I's values further modeling including spatial lag and error regressions will be used

(Anselin, 1988, 1995; Anselin & Bera, 1998; Anselin, Bera, Florax, & Yoon, 1996; Schabenberger & Gotway, 2005).

Statistical Models

Second, OLS and weighted OLS models for potential spatial accessibility to primary care physicians at the tract level will be analyzed. For non-spatial data, OLS seems the most appropriate approach for prediction if assumptions like normality of residuals and equal variance (homoskedasticity) are met. If the data display spatial heterogeneity, then several OLS assumptions do not hold; specifically, autocorrelation of the residuals from the model and heteroskedasticity in the residuals from the model. Tobler's First Law of Geography states that everything is related to everything, but things that are closer together are more related (Tobler, 1970). Since the outcome, spatial accessibility to primary care physicians, varies by census tracts (e.g., some tracts have a high number of primary care physicians and a high number of population in the tract, while other tracts have fewer physicians or none and fewer population) weighted least squares models will be used to evaluate the effect of tract population on spatial accessibility to primary care. Three different models for spatial accessibility at distances of 15, 30 and 60 miles, measured as the distance between population centroids at the tract level and primary care physician locations will be considered, each including the same five covariates: residential segregation, % in poverty, % without high school education, unemployment rate, and type of place. Each outcome will be fit using OLS and weighted least squares with the following model specification,

$$Access_i = \alpha + \beta_1 segregation_i + \beta_2 \%poverty_i + \beta_3 unemployment_i + \beta_4 \%no\ high\ school_i + \beta_5 metro_i + \varepsilon \quad (5)$$

where $segregation_i$, $\%poverty_i$, $unemployment_i$, $\%no\ high\ school_i$, are the z scored variables described above, $metro_i$ is a dummy variable for metro/non-metro status, α represents the intercept, while ε the model error term. However, a visualization of the Q-Q plot and the Shapiro Wilk test indicate non-normality in the standardized residuals, a violation of the OLS normality assumption. After using White's correction, the non-normality in residuals is still present.

Spatially Autoregressive Models

After evaluating the residuals of the OLS and weighted OLS models for autocorrelation and using the specification tests to determine which model would be appropriate for each outcome, spatial lag and spatial error models will be used. Expected spatial associations between spatial accessibility to primary care and each of the covariates is measured by the spatial lag model as it accounts for local structure in the dependent variable, while the spatial error model accounts for autocorrelation in the model residuals.

The spatial lag model has the following specification,

$$y = \rho W y + x\beta + \varepsilon \qquad (6)$$

where ρ is the spatially parameter, Wy is the spatially lagged potential accessibility to primary care services for row-standardized spatial matrix W, x is a matrix of observations on the explanatory variables, and ε is a vector error term.

Also, the spatial error model is specified below,

$$y = x\beta + \varepsilon \qquad (7)$$

$$\varepsilon = \varphi W\varepsilon + \gamma$$

where the first equation is the baseline equation presented for the OLS model, ε is the vector of error terms that are spatially weighted by the weight matrix (W), φ is the spatial error coefficient (if no spatial correlation between errors is present then $\varphi = 0$), and γ is a vector of uncorrelated error terms.

The same covariates used for the OLS and weighted OLS models will be considered for these models as well. To model the spatial structure of the Census tracts, a row-standardized spatial connectivity matrix is used following a distance-based algorithm of classifying cases as neighbors. This approach selects k nearest neighbors (where k is the number of neighbors; for this study, k=4).

Results

Table 1 highlights descriptive of the spatial accessibility indices represented by the mean, median, and standard deviation. For example, the mean and median values for accessibility indices that were measured by the distance between population centroids at the Census tract level and primary care physician locations look slightly different. There is a notable variation, highlighted by the standard deviation of these indices presented in Table 1 for these spatial accessibility indices. For better visualization 3 maps are presented in Figure 1, 2, and 3 each describing spatial accessibility index within 15, 30, and 60 miles for the state of Texas, San Antonio (up right corner), Houston (down left corner), and Dallas.

Table 1 Descriptive for Spatial Accessibility Index Within Different Distances and Predictors and Moran's I

	Mean	Median	Standard deviation	Moran's I
15 miles distance	0.47	0.19	0.52	0.56
30 miles distance	0.83	0.63	0.77	0.60
60 miles distance	1.21	0.95	0.90	0.55
Multi-group residential segregation	0.04	0.02	0.08	0.12
% in poverty	12.50	15.57	12.36	0.09
% without high school education	23.72	21.20	15.75	0.15
Unemployment rate	7.40	6.70	4.43	0.07
Metro/non-metro	0.17	-	0.38	0.26

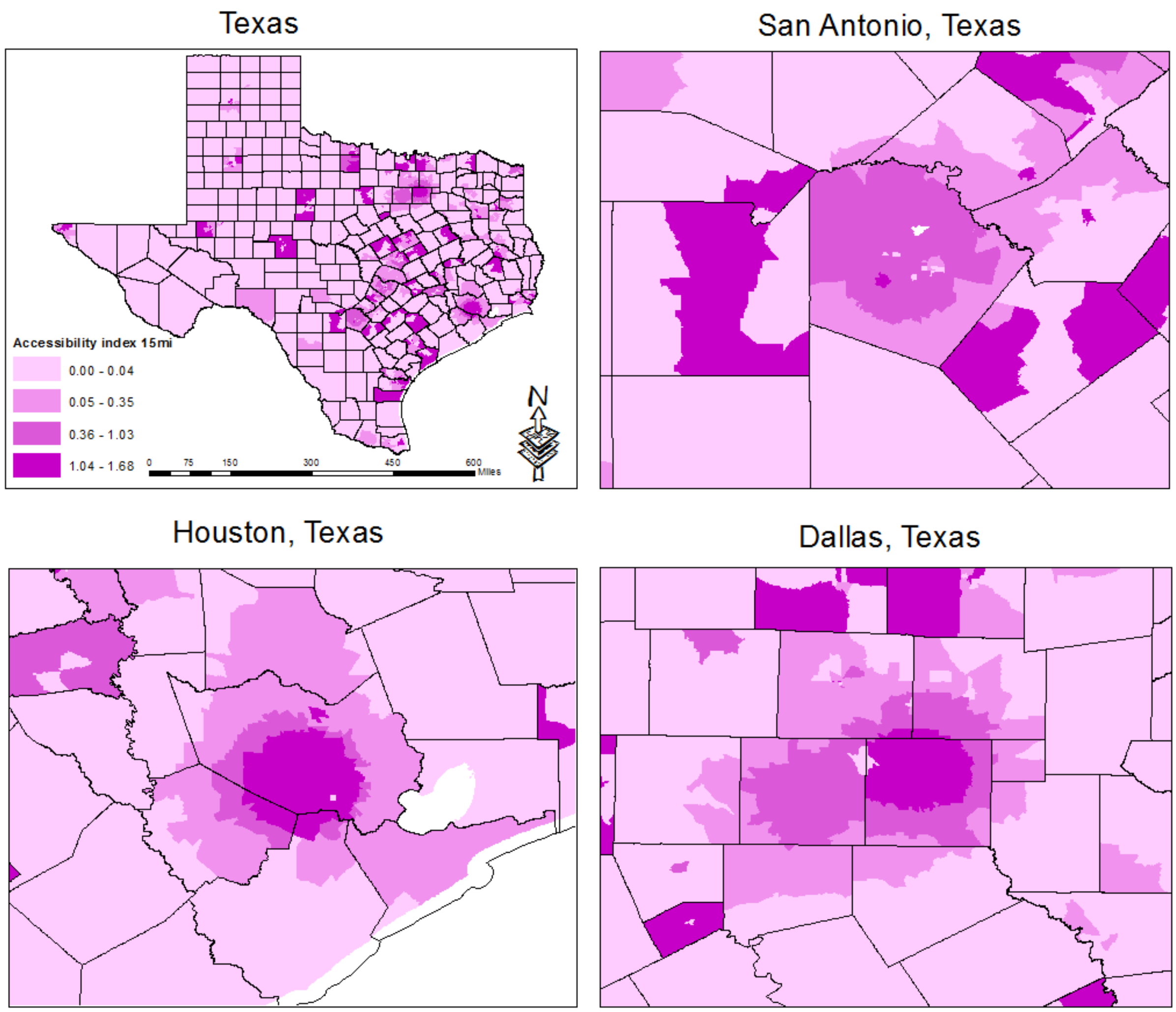

Figure 1 Spatial Accessibility index within 15 miles for Texas, San Antonio, Houston, and Dallas Census tracts

Patterns emerge and can be identified in each of the three figures 1, 2, and 3. Clusters of higher values of spatial accessibility can be seen in the San Antonio – Austin – Dallas corridor and the area around Houston. For each index within the maps for San Antonio, Houston, and Dallas it can be observed that the spatial accessibility to primary care physicians vary based on the location to the center of the city. The other clusters of spatial accessibility are somehow consistent for each index.

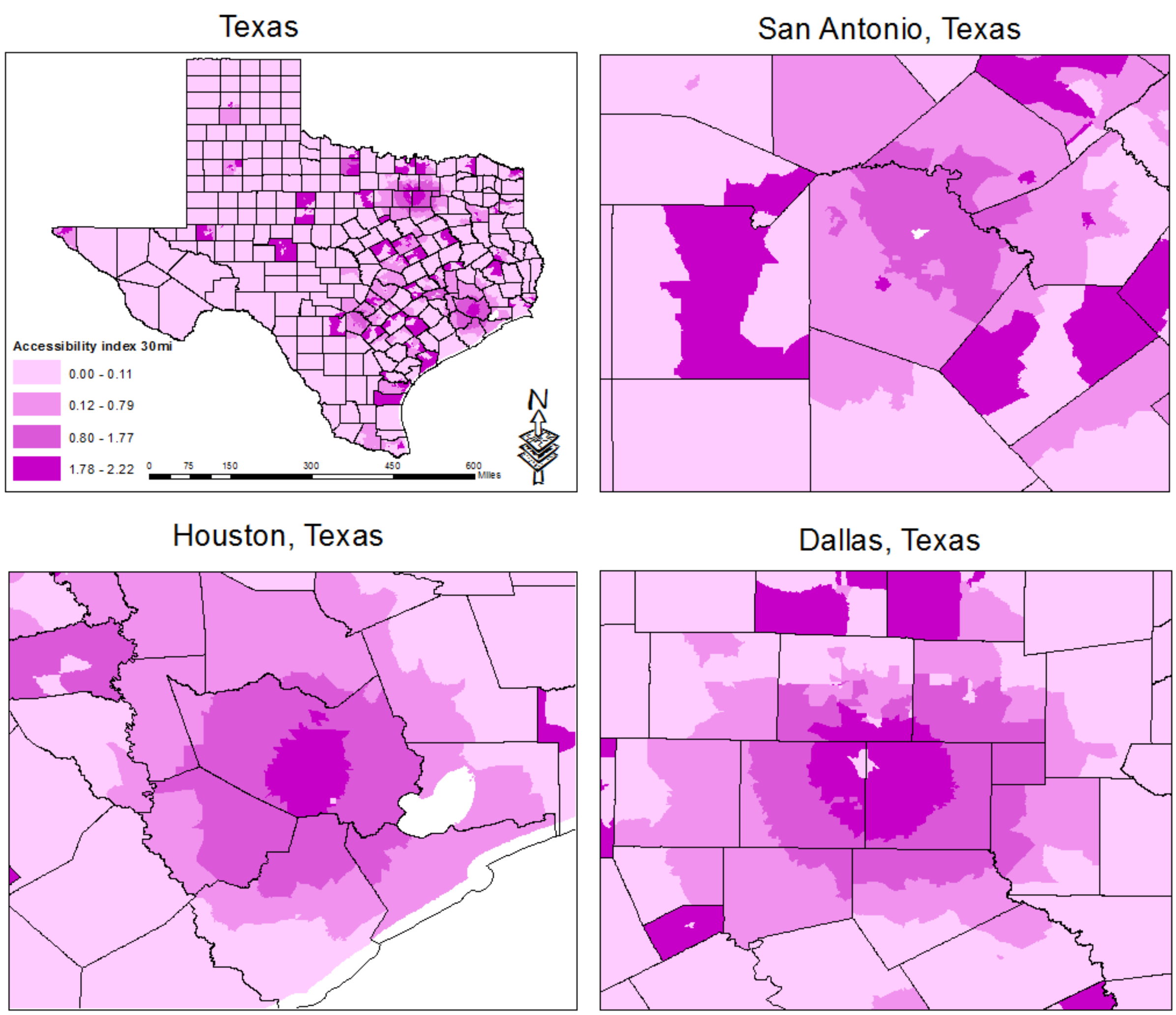

Figure 2 Spatial Accessibility index within 30 miles for Texas, San Antonio, Houston, and Dallas Census tracts

Table 2 presents results from the OLS and weighted OLS models for the association between potential spatial accessibility and socio-demographic variables as well as type of place. Surprisingly OLS models indicate that spatial accessibility rates increase as the % without a high school education and the unemployment rate increases for the three outcomes (15, 30, and 60 mile spatial accessibility indices). However, for the spatial accessibility index within 15 miles unemployment rate is significant only at 90% level.

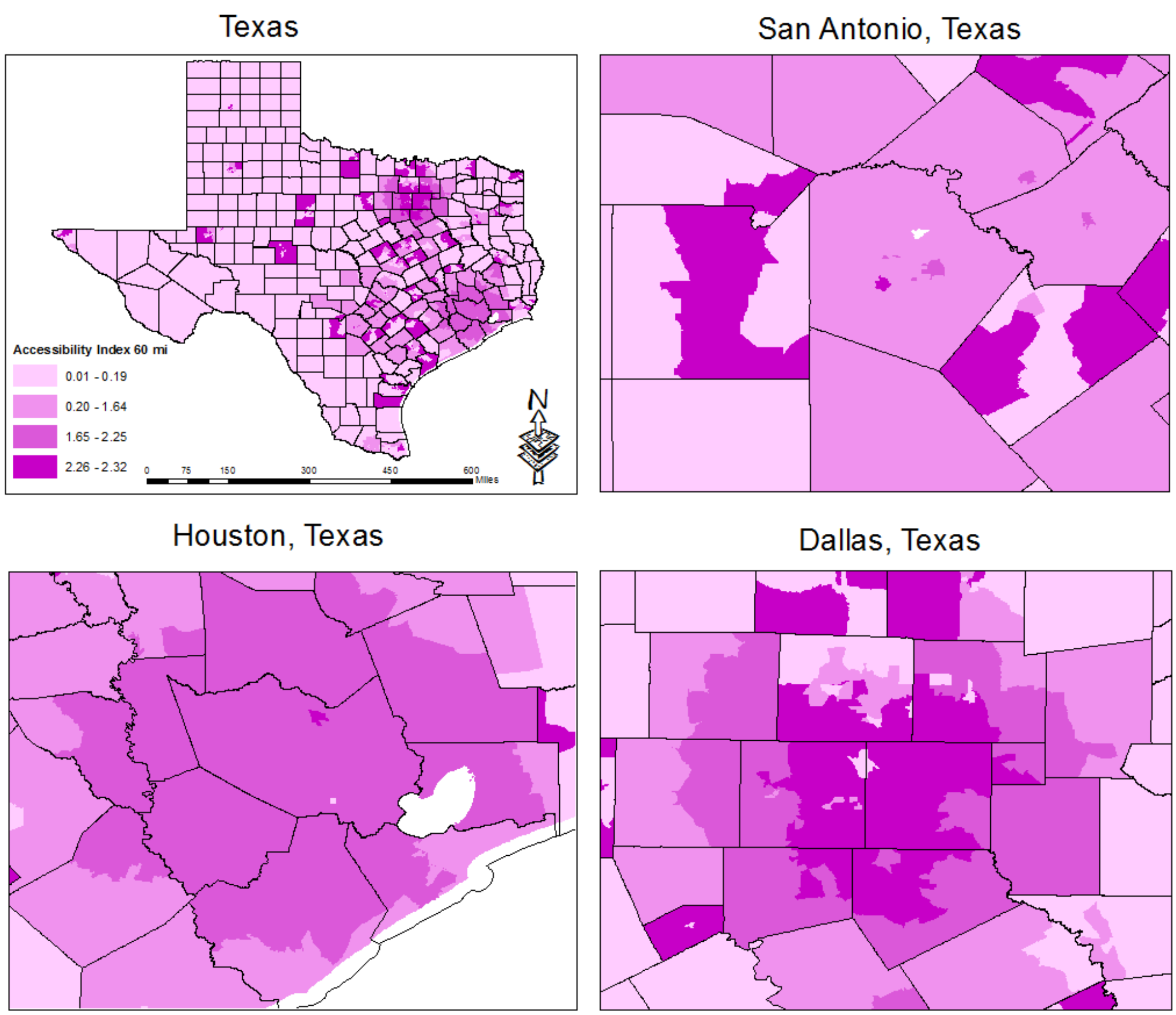

Figure 3 Spatial Accessibility index within 60 miles for Texas, San Antonio, Houston, and Dallas Census tracts

Table 2 Associations Between Residential Segregation, SES, Education, Unemployment, Type of Place and Spatial Accessibility Using OLS and Weighted OLS Models, 2011

Models	Estimate	Robust SE	Pr(>|t|)	Observed Moran's I
OLS model (15 miles)				**0.52**
Segregation (Theil Index)	-0.08	0.01	<0.001***	
% in Poverty	-0.12	0.02	<0.001***	
% ≤ High School	0.18	0.02	<0.001***	
Unemployment rate	0.04	0.02	0.028*	
Metro/non-metro	-1.09	0.02	<0.001***	
Intercept	0.19	0.02	<0.001***	
OLS model (30 miles)				**0.57**
Segregation (Theil Index)	-0.10	0.01	<0.001***	
% in Poverty	-0.23	0.02	<0.001***	
% ≤ High School	0.17	0.02	<0.001***	
Unemployment rate	0.09	0.02	<0.001***	
Metro/non-metro	-1.22	0.02	<0.001***	
Intercept	0.21	0.02	<0.001***	
OLS model (60 miles)				**0.52**
Segregation (Theil Index)	-0.10	0.01	<0.001***	
% in Poverty	-0.30	0.02	<0.001***	
% ≤ High School	0.14	0.02	<0.001***	
Unemployment rate	0.13	0.02	<0.001***	
Metro/non-metro	-1.18	0.03	<0.001***	
Intercept	0.20	0.02	<0.001***	
Weighted OLS model (15 mi)				**0.50**
Segregation (Theil Index)	-0.10	0.01	<0.001***	
% in Poverty	-0.14	0.03	<0.001***	
% ≤ High School	0.22	0.03	<0.001***	
Unemployment rate	0.04	0.02	0.026*	
Metro/non-metro	-1.10	0.02	<0.001***	
Intercept	0.19	0.02	<0.001***	
Weighted OLS model (30 mi)				**0.54**
Segregation (Theil Index)	-0.12	0.01	<0.001***	
% in Poverty	-0.28	0.02	<0.001***	
% ≤ High School	0.21	0.02	<0.001***	
Unemployment rate	0.10	0.02	<0.001***	
Metro/non-metro	-1.26	0.02	<0.001***	
Intercept	0.25	0.02	<0.001***	
Weighted OLS model (60 mi)				**0.49**
Segregation (Theil Index)	-0.11	0.01	<0.001***	
% in Poverty	-0.36	0.02	<0.001***	
% ≤ High School	0.18	0.02	<0.001***	
Unemployment rate	0.15	0.02	<0.001***	
Metro/non-metro	-1.21	0.03	<0.001***	
Intercept	0.26	0.02	0.0010***	

Significance level: ***p≤0.0001, **p≤0.05, *p≤0.01, Population Weight at Census tract level

Also, as residential segregation, % in poverty, and living in a non-metropolitan area increase, potential spatial accessibility to primary care physicians decreases. Surprisingly, increase in unemployment rate increases potential spatial accessibility to primary care physicians.

When the variability in the population size is controlled for, the effects of the estimates have a slightly higher effect on spatial accessibility than in the OLS models. The directionality of the relationship between the outcome and estimates is similar to the one discussed for OLS models. We can see a potential spatial accessibility advantage for census tracts with higher unemployment rates and with % of population with less than a high school education. Also, a disadvantage in potential spatial accessibility to primary care services with increases in residential segregation, % in poverty, and whether populations reside in a rural area as seen in Table 2.

Further, Moran's I is calculated to evaluate the correlation in the model residuals. Also in table 4.2 the global observed Moran's I for each of the OLS and weighted OLS models considered – the 15-mile, 30-mile, and 60-mile catchments, respectively is presented. All of the observed values are greater than the expected values indicating a highly and significant degree of global autocorrelation in the OLS and weighted OLS model residuals. Based on these results, spatial regression models are further estimated. Since the results from the OLS and weighted OLS models are almost similar for 15, 30, and 60 miles only the results from 15 and 60 miles will be illustrated in table 3 for the spatial lag and error models.

Table 3 Associations Between Residential Segregation, SES, Education, Unemployment, Type of Place and Spatial Accessibility Using Spatial Regression Models, 2011

Parameters	Estimate	Robust standard error	Pr(>|z|)
Spatial lag model (15 miles)			
Segregation (Theil Index)	-0.05	0.01	0.0002***
% in Poverty	-0.05	0.02	0.002**
% ≤ High school	0.09	0.02	<0.001***
Unemployment rate	0.01	0.01	0.3717
Metro/non-metro	-0.76	0.02	<0.001***
Intercept	0.13	0.01	<0.001***
ρ	**0.64**		
LR test vs. OLS	1665.8		
AIC	**7504.8**		
Spatial lag model (60 miles)			
Segregation (Theil Index)	-0.06	0.01	<0.001***
% in Poverty	-0.19	0.02	<0.001***
% ≤ High school	0.06	0.02	0.0002**
Unemployment rate	0.08	0.01	<0.001***
Metro/non-metro	-0.86	0.03	<0.001***
Intercept	0.15	0.01	<0.001***
ρ	**0.63**		
LR test vs. OLS	1640.9		
AIC	**7235.9**		
Spatial error model (15 miles)			
Segregation (Theil Index)	-0.0294	0.01	0.0135*
% in Poverty	-0.0260	0.02	0.1154
% ≤ High school	0.0706	0.02	<0.001***
Unemployment rate	0.0092	0.01	0.4856
Metro/non-metro	-0.8126	0.03	<0.001***
Intercept	0.1363	0.04	0.0002***
ρ	**0.6819**		
LR test vs. OLS	1656.9		
AIC	**7513.6**		
Spatial error model (60 miles)			
Segregation (Theil Index)	-0.05	0.01	<0.001***
% in Poverty	-0.15	0.02	<0.001***
% ≤ High school	0.04	0.02	0.0144*
Unemployment rate	0.07	0.01	<0.001***
Metro/non-metro	-0.93	0.03	<0.001***
Intercept	0.17	0.04	<0.001***
ρ	**0.68**		
LR test vs. OLS	1618.4		
AIC	**7258.4**		

Significance level: ***p≤0.0001, **p≤0.05, *p≤0.01

The results of spatial models are presented in Table 3. Results for spatial lag and spatial error models look similar to the ones presented from the OLS regression in Table 2. Since spatial accessibility indices for 15, 30, and 60 miles had similar results, I only present results for spatial indices calculated for 15 and 60 miles. General and Robust Lagrange multiplier statistics are tested for the presence of residual autocorrelation in the OLS models, indicating that the error model is better in these circumstances (Anselin et al., 1996).

The spatial error model controls for the autocorrelation in the model error term. The ρ parameter presented for each model (15 and 60 miles) is over 0.6 indicating a significant autocorrelation in the error term. When autocorrelation in the error term is controlled for, an overall decrease can be seen in the regression coefficients for residential segregation, % in poverty, % without a high school education, and living in rural areas in predicting Census tracts spatial accessibility rates (when comparing the OLS estimates from table 2 with the estimates from table 3 for the error model). Also, the spatial error term shows a significant improvement in total fit of the OLS model (a drop in AIC over 1,500 compared to the OLS model).

The spatial lag models indicate slightly different results from the error and OLS models. We see similar decrease in the overall regression coefficients, however not as much as in the spatial error models. When the outcome is the spatial accessibility within 15 miles, the regression estimate for unemployment becomes insignificant (similar for the spatial error model). Other than the unemployment estimate, all other regression estimates experience an overall decrease for the spatial lag model (with the outcome spatial accessibility index within 15 and 60 miles). When the spatial lag model is compared with the OLS model, as in the case of spatial error model we observe a significant improvement in the overall fit (the drop in AIC is over 1,500 compared to the OLS model). The spatial autoregressive parameter, ρ, indicates a significant autocorrelation in the lagged spatial accessibility rates. The regression coefficients for the spatial lag and error models are almost similar; however the spatial error models (for spatial accessibility within15 and 60 miles) present lower estimates compared to the lag model. In terms of model fit the spatial lag models perform better than the spatial error models for both outcomes, indicating that the spatial process that generates the data might work more on the dependent variable rather than the error component of the data. However, the basic Lagrange multiplier statistics indicates that the spatial error model is a slightly better model than the spatial lag model and that the spatial error model might provide a better representation of the process.

Even though, statistically, the spatial lag model is the best model, conceptually, the spatial error model is more likely to explain the spatial pattern present in the spatial accessibility index and that the result of unobserved independent variables might be omitted from the model.

Discussion and Limitations

The purpose of this research was to determine whether multi-group residential segregation in areas of poor socioeconomic status and in rural areas has a negative effect on potential spatial accessibility. Also, this research addressed the possibility that estimates of the covariates considered might have differential effects on spatial accessibility indices calculated based on 15, 30, and 60 miles.

Results from this study indicate that indeed residential segregation has a significant negative effect on potential spatial accessibility to primary care and that any increase in residential segregation index would lower spatial accessibility to primary care physicians in the state of Texas. On a similar note, Gaskin and colleagues (2012) found that African American zip codes were negatively associated with the availability of physician services after controlling for socioeconomic and other factors. In contrast, Hispanic and Asian zip codes were positively affected by segregation; however, residential segregation was computed as a two-group measure instead of a multi-group measure. Probst and colleagues also found that rural residents and African Americans populations experience higher travel burdens than urban residents or whites, respectively, when seeking medical care (Probst, Laditka, Wang, & Johnson, 2007). The "risk regulators" theoretical framework treats residential segregation's effect on spatial accessibility to primary care as a second-order rather than causal, as had been previously theorized (Williams & Collins 2001). Also, % in poverty and living in a rural area have demonstrable negative impacts on spatial accessibility to primary care. For example, the estimate for '% in poverty' has a consistent negative association with spatial accessibility in the OLS and weighted OLS models. This result highlights the idea presented by Shi and Stevens (2010) in which certain place characteristics create higher vulnerability among populations living in those places compared with populations that do not present those place characteristics. This can be seen as a "vicious circle", where communities with higher levels of residential segregation face higher levels of concentrated poverty (Acevedo-Garcia & Lochner, 2003). Other consequences could be that residents of these communities are more likely to have lower levels of educational attainment, have access to fewer job opportunities, and experience higher unemployment. For this study, even if unemployment rate and % without a high school education were statistically significant their estimates seemed to have a positive effect on spatial accessibility index. Therefore this can further translate into fewer minority populations competing for high-earning jobs or earn competitive salaries when compared to populations living in more affluent neighborhoods. Fewer economic resources and lack of health knowledge can in turn act as a barrier to accessing primary care for individuals living in more segregated neighborhoods.

The second part of this study tested whether the effect of the independent variables considered varied by the distance considered when calculating spatial accessibility indices. Each statistical analysis considered three spatial accessibility indices, which were calculated for 15 miles, 30 miles, and 60 miles distance between population centroid at the tract level and primary care physician locations. Indeed, the independent variables considered did vary by distance. Each of the three models with spatial accessibility (15, 30 and 60 miles) presented slightly different estimates, indicating that, for this particular study, potential spatial accessibility to primary care physicians in Texas did vary by distance, however the results were slightly different.

Also, since the Global Moran's I values indicated the presence of spatial autocorrelation further spatial lag and error models were tested. Even though, the general and robust Lagrange multiplier statistics that tests for the presence of residual autocorrelation in the OLS models indicated the error model as the best model, the results (AIC's) between spatial lag, error, and OLS indicated the spatial lag as having a slightly better fit (Anselin et al., 1996). The spatial lag model indicates a diffusive process in the spatial accessibility index. This implies that the value of the spatial accessibility index in one particular location is influenced by the value of its neighbors. However, the "risk regulators" framework shows support for the theory behind the spatial error model, that there might be unobserved independent variables that are omitted from the model that could explain these processes.

Although this research may provide valuable knowledge regarding the influence of multi-group residential segregation and other factors on spatial accessibility to primary care, the following limitations should be acknowledged. Usually these types of studies analyze smaller areas (rather than an entire state) because they are computationally expensive and the level of error increases commensurately. Choosing the catchment (threshold) to calculate the spatial accessibility to primary care services is important. Previous studies used driving distance (which accounted for driving speed limits) from the population centroids to physician locations. For this analysis, I choose various catchments to calculate the spatial accessibility index: 10, 15, 30, 45, and 60 miles; however, only results for the 15, 30, and 60 miles catchments were presented given the similarities in outcomes. One study found that sicker individuals are more likely to travel longer distances if hospitalization is involved (McGrail & Humphreys, 2014). This study cannot determine whether rural populations are more likely to do that since the measure examines potential spatial accessibility without knowing if access was realized (i.e., actually seeing a physician). Additionally, Phibbs and Luft (1995) conducted a study in upstate New York underscoring the fact that straight-line distance can be a very reasonable proxy for travel time, unless the study focuses on specific hospitals or dense urban areas with high congestion and reliance on surface streets.

Furthermore, for this particular study, not all population-weighted centroids at the tract level were able to identify a physician location since they may have been outside of the 60-mile range. In addition, spatial accessibility index was computed at the tract level, and therefore some values for the index are zero due to smaller populations in a tract, the lack of primary care physicians, or simply the absence of primary care physicians for particular tracts, especially when considering smaller distances (e.g., 10 and 15 miles). Regarding residential segregation there might be a small disadvantage when computing the multi-group index. Usually this index considers all 4 racial/ethnic groups (non-Hispanic whites, non-Hispanic blacks, Hispanics, and non-Hispanic other), however this mix of racial/ethnic groups can be usually found in metropolitan areas (e.g., San Antonio, Austin, Houston, Dallas, etc.). If the presence of one racial/ethnic group is missing, the final index computed would have a smaller value.

Despite these limitations, this research demonstrates that even after controlling for factors such as residential segregation, percent in poverty, percent without a high school education, unemployment, and type of place, the potential spatial accessibility to primary health care services was lower in higher-poverty, segregated, and rural areas. Further studies should focus on smaller areas (for example, one or more counties, even Metropolitan Statistical Areas – MSA's), which will reduce the intense computational processes and potential for error. Also, future studies should consider additional covariates such as the percentage of population with medical insurance or the percentage of linguistically isolated communities.

References

Acevedo-Garcia, D., & Lochner, A. K. (2003). Residential Segregation and Health. In I. Kawachi & F. L. Berkman (Eds.), *Neighborhoods and Health*. New York, NY: Oxford University Press, Inc.

Acevedo-Garcia, D., Lochner, A. K., Osypuk, L. T., & Subramanian, V. A. (2003). Future Directions in Residential Segregation and Health Research: A Multilevel Approach. *American Journal of Public Health, 93*(2).

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht, London.

Anselin, L. (1995). Local indicators of spatial association LISA. *Geographical Analysis, 27*, 93-115.

Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah & D. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237-289). New York.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*, 77-104.

Begley, E. C., Vojvodic, R. W., Seo, M., & Burau, K. (2006). Emergency Room Use and Access to Primary Care: Evidence from Houston, Texas. *Health Care for the Poor and Underserved, 17*, 610-624.

Bell, J. F., Zimmerman, F. J., Almgren, G. R., Mayer, J. D., & Huebner, C. E. (2006). Birth outcomes among urban African-American women: a multilevel analysis of the role of racial residential segregation. *Social Science & Medicine, 63*, 3030-3045.

Census Bureau, C. A. F. (2010). from http://www.census.gov/geo/reference/county-adjacency.html

Dai, D. (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Health & Place, 16*, 1038-1052. doi: 10.1016/j.healthplace.2010.06.012

Do, D. P., Finch, K. B., Basurto-Davila, R., Bird, C., Escarce, J., & Lurie, N. (2008). Does place explain racial health disparities? Quantifying the contribution of residential context to the Black/white health gap in the United States. *Social Science & Medicine, 67*, 1258-1268.

Fitzpatrick, K. M., & LaGory, M. (2003). Placing' health in an urban sociology: cities and mosaics of risk and protection. *City and Community, 2*, 33-46.

Gaskin, J. D., Dinwiddie, Y. G., Chan, S. K., & McCleary, R. R. (2012a). Residential Segregation and Disparities in Health Care Services Utilization. *Medical Care Research and Review*, *69*(2), 158-175.

Gaskin, J. D., Dinwiddie, Y. G., Chan, S. K., & McCleary, R. R. (2012b). Residential Segregation and the Availability of Primary Care Physicians. *Health Services Research, 47*(6), 2353-2376.

Gaskin, J. D., Price, A., D.T., B., & LaVeist , T. A. (2009). Segregation and disparities in health service use. *Medical Care Research and Review, 66*(5), 578-589.

Gautam, S., Li, Y., & Johnson, G. T. (2014). Do alternative spatial healthcare access measures tell the same story? *GeoJournal, 79*(2), 223-235.

Glass, A. T., & McAtee, J. M. (2006). Behavioral science at the crossroads in public health: Extending horizons, envisioning the future. *Social Science & Medicine*, *62*, 1650-1671.

Guagliardo, F. M. (2004). Spatial accesibility of primary care:concepts, methods and challenges. [Review]. *International Journal of Health Geographics*, *3*, 1-13.

Guagliardo, F. M., Ronzio, R. C., Cheung, I., Chacko, E., & Joseph, G. J. (2004). Physician accessibility: an urban case study of pediatric providers. *Health & Place, 10*, 273-283.

Hayanga, A. J., Kaiser, E. H., Sinha, R., Berenholtz, M. S., Makary, M., & Chang, D. (2009). Residential Segregation and Access to Surgical Care by Minority Populations in US Counties. *American College of Surgeons*. doi: 10.1016/j.jamcollsurg.2009.01.047

Jargowksy, P. A. (1997). *Poverty and Place: Ghettos, Barrios, and the American City*. New York, NY: Russell Sage Foundation.

Kaiser Family Foundation. (2013). Population Distribution by Race/Ethnicity, from http://kff.org/other/state-indicator/distribution-by-raceethnicity/

Kwan, M.-P. (2013). Beyond Space (As We Knew It): Toward Temporally Integrated Geographies of Segregation, Health, and Accessibility. *Annals of the Association of American Geographers, 103*(5), 1078-1086.

Landrine, H., & Corral, I. (2009). Separate and Unequal: Residential Segregation and Black Health Disparities. *Ethnicity & Disease, 19*.

Langford, M., & Higgs, G. (2006). Measuring Potential Access to Primary Healthcare Services: The influence of Alternative Spatial Representation of Population. *The Professional Geographer*, *58*(3), 294-306.

Link, G. B., & Phelan, J. (1995). Social Conditions as Fundamental Causes of Disease. *American Sociological Association, 36*(Extra Issue: Forty Years of Medical Sociology: The State of the Art and Directions for the Future), 80-94.

Luo, W. (2003). Using a GIS-based floating catchement method to assess areas with shortage pf physicians. *Health & Place, 10*.

Luo, W., & Qi, Y. (2009). An enhanced two-step floating catchement area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place*, *15*, 1100-1107.

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environmental an Planning B: Planning and Design, 30*, 865-884.

Luo, W., & Whippo, T. (2012). Variable catchement sizes for the two-step floating catchement area (2SFCA) method. *Health & Place, 18*, 789-795.

Massey , D., & Denton , N. (1988). The Dimensions of Residential Segregation. *Social Forces 67*(2), 281-315.

Massey, S. D., & Fong, E. (1990). Segregation and neighborhood quality: Blacks, Hispanics, and Asians in the San Francisco metropolitan area. *Social Forces, 69*, 15-32.

Massey, S. D., & Shibuya, K. (1995). Unraveling the tangle of pathology: the effect of spatially concentrated joblessness on the well-being of African Americans. *Social Science Research, 24*(352-366).

McGrail, R. M. (2009). A new index of access to primary care services in rural areas. *Australian and New Zealand Journal of Public Health, 33*(5), 418-423.

McGrail, R. M., & Humphreys, J. S. (2014). Measuring spatial accessibility to primary health care services: Utilising synamic catchment sizes. *Applied Geography, 54*, 182-188.

Penchansky, R., & Thomas, J. (1981). The Concept of Access. *Medical Care*, *19*(2), 127-140.

Phibbs, S. C., & Luft, S. H. (1995). Correlation of Travel Time on Roads Versus Straight Line Distance. *Medical Care Research and Review, 52*(4), 532-542.

Polzin, P., Borges, J., & Coelho, A. (2014). An extended kernel density two-step floating catchement area method to analyze access to health care. *Environmental an Planning B: Planning and Design, 41*, 717-735.

Probst, C. J., Laditka, B. S., Wang, J.-Y., & Johnson, O. A. (2007). Effects of residence and race burden of travel for care: cross selectional analysis of the 2001 US National Household Travel Survey. *BMC Health Services Research, 7*(40).

Reardon, F. S., & Firebaugh, G. (2002). Measures of Multigroup Segregation. *Sociological Methodology, 32*, 33-67.

Schabenberger, O., & Gotway, C. (2005). *Statistical Methods for Spatial Data Analysis*. Boca Raton, FL.

Shi, L., & Stevens, D. G. (2010). *Vulnerable Populations in the United States* (2nd ed. ed.). San Francisco, CA: Jossey-Bass.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*(2), 234-240.

USDA. (2011). from http://webarchives.cdlib.org/sw1wp9v27r/http://www.ers.usda.gov/Briefing/Rurality/RuralUrbCon/

Westra, E. (2013). *Python Geospatial Development* Birmingham, UK: Packt Publishing Ltd.

White, K., & Borrell, L. (2011). Racial/ethnic residential segregation: Framing the context of health risk and health disparities. *Health & Place, 17*, 438-448.

Williams, D. R., & Jackson, P. B. (2005). Social sources of racial disparities in health. *Health Affairs, 24*(2), 325-334.

Williams , R. D., & Collins , C. (2001). Racial Residential Segregation: A Fundamental Cause of Racial Disparities in Health. *Public Health Reports, 116*(September-October).