Death by Segregation: Does the dimension of racial segregation matter?[1]

Tse-Chuan Yang, Ph.D.
Department of Sociology
Center for Social and Demographic Analysis
University at Albany, State University of New York

Stephen A. Matthews, Ph.D.
Department of Sociology
Population Research Institute
Social Science Research Institute
Pennsylvania State University

Hui-Shien Tsao, Ph.D.
Center for Social and Demographic Analysis
University at Albany, State University of New York

Manuscript for PAA 2014
[Please Do Not Cite without Authors' Consent]

Words: 7,572 (excluding Tables and Figures)
Figures: 3
Tables: 3
Appendix: 1

[1] We acknowledge the support from the Center for Social and Demographic Analysis at the University at Albany, SUNY, which receives core funding from NICHD [R24-HD04494309].

Abstract

The county-level geographic mortality differentials have persisted in the past four decades in the United States (US). Though several socioeconomic factors (e.g., inequality) partially explain this phenomenon, the role of race/ethnic segregation, in general, and the different dimensions of segregation, more specifically, has been underexplored. Focusing on all-cause age-sex standardized US county-level mortality (2004-2008), this study has three substantive goals: (1) to understand whether segregation is a determinant of mortality, (2) to investigate whether the relationship between segregation and mortality varies by race/ethnic groups, (3) to explore whether different dimensions of segregation (i.e., evenness, exposure, concentration, centralization, and clustering) are associated with mortality. A fourth goal is methodological: to assess whether spatial autocorrelation influences our understanding of the associations between the dimensions of segregation and mortality. Race/ethnic segregation was found to contribute to the geographic mortality disparities. Moreover, the relationship with mortality differed by both race/ethnic group and the dimension of segregation. Specifically, white/black segregation is positively related to mortality, whereas the segregation between whites and non-black minorities is negatively associated with mortality. Among the five dimensions of segregation, evenness and exposure are more strongly related to mortality. Spatial filtering approaches also help identify six unique spatial patterns that significantly affect the spatial distribution of mortality. These patterns provide possible insights to help identify omitted variables related to the patterning of mortality in the US.

Keywords: mortality, segregation, spatial filtering, geographic health disparities

Introduction

While the United States (US) as a whole has experienced a significant decrease in mortality since World War II from approximately 20 deaths per 1,000 population to 8 deaths per 1,000 population (Hoyert 2012), the persistence of mortality disparities across US counties that has been stable for the past four decades (Cossman et al. 2007) motivates this study. For example, the counties in the Black Belt and lower Mississippi Valley have relatively high mortality rates, whereas those in the Great Plain, Mid-West, and along the US/Mexico border are much lower (Cossman et al. 2007). Explanations for this spatial pattern have mainly focused on socioeconomic factors (e.g., poverty, unemployment, educational attainment, income inequality, and social capital), demographic structure (e.g., racial compositions), and context (e.g., rurality) (Blanchard et al. 2008; Cossman et al. 2008; Lochner et al. 2003; McLaughlin et al. 2007; Sparks and Sparks 2010; Yang, Jensen and Haran 2011). While these factors are important, the spatial variation in mortality has not been fully explained. However, a potential determinant, namely race/ethnic segregation and the different dimensions of segregation, has–perhaps surprisingly–rarely been incorporated into the analysis.

The concept of segregation is complex but earlier studies on health and segregation largely overlook the complexity of how to measure segregation. Massey and Denton (1988) have shown that segregation can be classified into five different dimensions–evenness, exposure, concentration, centralization, and clustering–and each of these dimensions can in turn be calculated in multiple ways. Little attention has been paid to the nuances of race/ethnic segregation in health research (Acevedo-Garcia et al. 2003; Kramer and Hogue 2009). A recent study used three different measures of segregation to explore the associations with mortality (Sparks, Sparks and Campbell 2013); however, this analysis focused solely on white/black

segregation, ignoring other race/ethnic groups. This paper makes a contribution by examining whether race/ethnic segregation and specific dimensions of segregation are associated with US county-level mortality.

Until recently, research on race/ethnic segregation has tended to focus on US metropolitan areas; that is, the areas where most minorities live (Fischer 2003; Logan, Stults and Farley 2004; Wilkes and Iceland 2004). This metropolitan focus persists even though race/ethnic segregation has been reported to be more intense in nonmetropolitan areas; areas where the racial composition is generally less diverse (Fuguitt 1994; Johnson and Lichter 2010; Murdock, Hwang and Hoque 1994; Parisi, Lichter and Taquino 2011). Following Lobao and colleagues (2007), we adopt a county-level perspective, incorporating metropolitan and nonmetropolitan areas, as we believe that this provides a more comprehensive assessment of the relationship between segregation and mortality. It also allows for a direct comparison with other county-level studies, e.g., Cossman et al (2007). The use of county-level data in an ecological analysis raises several methodological issues. Most specifically, using county-level data without controlling for spatial dependence may result in biased estimates and lead to invalid or misleading interpretation of findings (Cressie 1991). In this paper, we directly address this issue via the use of spatial filtering regression methods.

This study has four main goals: (1) to understand whether race/ethnic segregation is a determinant of mortality in US counties, (2) to investigate whether the relationship between segregation and mortality varies by race/ethnic groups (i.e., black, Hispanics, and Asians/Pacific Islander), (3) to explore whether different dimensions of segregation are associated with mortality, and finally (4), a methodologically informed goal, to assess whether spatial dependence influences our understanding of the associations between the individual dimension

of segregation and mortality. Our analysis focuses on US county-level Mortality data for 2004-2008.

Segregation and Health by Race/Ethnicity

The conventional belief that race/ethnic segregation is adversely related to health is partially rooted in the ethnic stratification perspective. Collins and Williams argued (1999) that race/ethnic segregation could be understood as a structural manifestation of racism against minorities, and in particular, non-Hispanic blacks (hereafter black). Discrimination against minority groups can take on many forms. Among the most dominant forms is when discrimination fuels the residential sorting process, which as Logan (1978) noted, is a powerful mechanism maintaining the advantages of the majority group and generates ethnic stratification. From this perspective, people living in racially segregated neighborhoods are exposed to multiple health risk factors, such as poverty, crimes, and poor public services (Collins and Williams 1999; Williams and Collins 2001), and these risk factors are associated with poor health and racial health disparities (Brulle and Pellow 2006). The legal challenges to discrimination and discriminatory practices enacted since the 1960s has in par lead to the decline in white/black segregation (Glaeser and Vigdor 2012).

Extending the ethnic stratification perspective, there are three key mechanisms underlying the common belief that race/ethnic segregation is detrimental to health: (1) areas with high levels of race/ethnic segregation have poor social indicators, which may contribute to poor health outcomes (Collins and Williams 1999; Link and Phelan 1995; Williams and Collins 2001); (2) race/ethnic segregation is associated with political alienation and powerlessness and these factors may lead to relatively few resources being channeled into a minority area; and (3) the environment of an area with high race/ethnic segregation is more likely to be neglected and

lacking infrastructure (Greenberg and Schneider 1994; LeClere, Rogers and Peters 1997). These pathways, individually and in combination, may expose local residents to multiple health risks; a negative association between race/ethnic segregation and health is expected (Brulle and Pellow 2006). The conventional framework is heavily driven by the ethnic stratification perspective, but this framework may not be applicable to non-black minority groups, such as Hispanics and Asians/Pacific Islanders.

The landscape of racial composition has rapidly changed since the 1980s mainly due to the influx of immigrants from both Latin America and Asia (Hobbs and Stoops 2002). The geographically mixing and thus the race/ethnic segregation of Hispanics and Asians/Pacific Islanders with white have been transformed (Iceland, Weinberg and Steinmetz 2002; Lee et al. 2008). Researchers examining the residential sorting processes for different minority groups find differences between the growing race/ethnic groups and the white population compared to the processes accounting for white/black segregation. When Hispanic and Asian immigrants enter the US, they tend to have limited resources and tend to live in an ethnically bound neighborhood or enclaves. This living arrangement can help them to improve their socioeconomic situation and the process of adaption to the new society. That is, race/ethnic segregation between Hispanics and Asians from whites may be strategic for these minority groups. Logan and colleagues (2002) identified two types of neighborhood that serves the goal to help immigrants to survive and thrive: ethnic enclave and ethnic community. The distinction between these two neighborhoods is grounded in the motives of minority residents. Specifically, the former plays a temporary or transitional role in the process of adaption, whereas the latter is established by minority members who voluntarily live nearby, usually in the later stage of the process of adaption (Logan et al. 2002). Despite the difference, the shared and imperative function of both neighborhoods is to

help minorities to thrive or accumulate social and financial capital. This function may encourage Hispanics and Asians to be self-segregated to take advantage of ethnically bound neighborhoods.

That is, racial segregation may be beneficial to Hispanics and Asians. First, living in an ethnic enclave/community can translate into increased social support, frequent social engagement with people of the same race/ethnicity, and fewer challenges emerging from linguistic isolation (LeClere et al. 1997). These factors foster strong social cohesion that may facilitate health and well-being (Song, Son and Lin 2010). Second, and related, an ethnic enclave/community may provide social, economic, and structural resources generated by the close-knit social connections or among residents of the same race/ethnicity (Eschbach et al. 2004; Lee and Ferraro 2007). That is, the access to educational, information, and occupational opportunities in an ethnic enclave/community may be better in an ethnic enclave or community than in other types of neighborhood. Third, being segregated from the dominant racial group indicates a low level of exposure to direct racial discrimination, and in such a neighborhood, the norm that racial discrimination is intolerable would prevail thanks to a strong ethnic identity (Bécares, Nazroo and Stafford 2009; Whitley et al. 2006). These factors lead us to argue that race/ethnic segregation may be beneficial to population health, particularly for non-black minorities.

Since segregation and mortality are essentially ecological measures, their distribution across counties may be spatially dependent as many other social data (Cressie 1993). To obtain unbiased estimates, researchers have become increasingly aware of the importance to take spatial dependence into account, particularly in ecological demographic studies (Matthews and Parker 2013). The subsequent section will discuss how this study addresses this methodological issue.

Using Spatial Structure to Address Spatial Dependence

In this study, mortality and segregation both are measured for ecologic units (i.e., counties) and the spatial relationships between these units need to be explicitly incorporated into an analytic strategy if we are to obtain unbiased estimates of the relationships that exist between them (Voss 2007; Voss et al. 2006). Several demographers have explicitly incorporated a spatial perspective into mortality research based on a spatial econometric framework (Sparks et al. 2013; Sparks and Sparks 2010; Yang et al. 2011). Arguably, spatial econometric methods are the most popular in spatial demography but they can only provide an *overall* assessment of *how much* spatial structure matters. A common criticism of the spatial econometrics approach is that researchers do not know *how* spatial structure matters. Griffith (2000, 2003) and Tiefelsdorf and Griffith (2007) have proposed spatial filtering methods that use eigenfunctions to create a series of spatial patterns that are mutually unrelated but associated with the spatial structure underlying the spatial units. A recent study by Thayn and Simanis (2013) suggested that the spatial filtering approach effectively reduces spatial misspecification errors, increases model fit, and eliminates spatial dependence. In addition, the unrelated spatial patterns can be visualized (i.e., mapped) to gain further insight into how spatial structure contributes to the analysis and potentially shed new light on omitted variable bias (Griffith 2003). Indeed, perhaps the most distinctive feature of the spatial filtering approach lies in the decomposition of errors and that allows for the visualization of unknown spatial processes that affect the spatial pattern of the dependent variable (i.e., mortality). To our knowledge, our study is the first to use a spatial filtering approach to examine the mortality pattern across US counties.

Drawing from the discussion above, we propose two substantive research hypotheses and one derived from our use of spatial filtering approaches. The main substantive hypotheses are:

(1) white/black segregation is positively related to mortality as the segregation process (ethnic stratification) is rooted in discrimination, and

(2) white/Hispanic and white/Asians and Pacific Islanders segregation are beneficial to mortality as the segregation processes lead to the formations of enclaves and/or communities.

We also expect that

(3) spatial dependence undermines the estimates of the relationships between segregation and mortality and as such the spatial filtering approach can help refine our model specification and identify the spatial patterns. We will examine these issues focusing on five dimensions of segregation.

Data and Measures

The county-level mortality rate is the dependent variable of this study. Based on the Compressed Mortality Files maintained by the National Center for Health Statistics (NCHS 2010), we created the 2004-2008 five-year average mortality rates that are standardized to the 2006 US age-sex population structure. Using the five-year average rates minimizes the fluctuations across years and this approach has been used in recent ecological mortality studies (Yang et al. 2012; Yang, Teng and Haran 2009). As race/ethnic segregation plays a crucial role in this study, rather than standardize mortality rates with respect to racial groups, instead we include the race/ethnic composition of a county in the analysis.

Our independent variables can be classified into seven groups: segregation, urbanicity, socioeconomic status, racial composition, income inequality, social capital, and population health. We discussed them in detail below:

Segregation: We measured all five dimensions of racial segregation. Specifically, the "evenness" and "exposure" dimensions are measured with *entropy* (Theil 1972) and the *isolation index*, respectively. The entropy index assesses the average deviation of the sub-unit (i.e., tract in this study) from the county's racial diversity, whereas the isolation index measures "the extent to which minority members are exposed only to one another (p.288)" (Massey and Denton 1988). Entropy and isolation index both range between 0 and 1 and higher values of entropy and isolation suggest higher segregation. The "centralization" dimension is captured with the *absolute centralization index* that is developed to understand if the minority group is distributed around the center of a county. Absolute centralization index varies between -1 and 1 where a positive value indicates that minorities tend to live nearer the center of a county, whereas the negatives values suggest that minority populations live in the outlying areas (Massey and Denton 1988). The "concentration" dimension is based on the *delta index* (Duncan, Cuzzort and Duncan 1961) and is calculated to assess the proportion of minority members who live in the areas where the minority density is higher than average, ranging between 0 (no concentration) and 1. The delta index represents the proportion of a minority group that have to move to reach a uniform density within an area (Massey and Denton 1988). The fifth and final dimension, "clustering," is measured by the spatial proximity index (White 1986), which is the average of proximities within the minority and majority group, respectively. A spatial proximity index greater than 1 suggests that minority members live close to one another and so do the majority. When a spatial proximity index is less than 1, it means that majority and minority members live closer to each other than the members of their own groups.

We calculated the five segregation measures by the following three race/ethnicity combinations: non-Hispanic white vs. non-Hispanic black (white/black), non-Hispanic white vs.

Hispanic (white/Hispanic), and non-Hispanic white vs. Asians/Pacific Islanders (white/API). We only focused on the three largest minority groups in order to avoid unreliable segregation measures due to small number issues in many counties. The statistical procedures developed by Iceland and colleagues (2002) were applied to the 2010 Census Summary File 1 race/ethnicity data to obtain the fifteen segregation measures for all US counties.

Urbanicity: The rural-urban mortality differential in US counties has been documented and metropolitan counties have been found to have higher mortality rates than their rural counterparts (McLaughlin et al. 2007). Taking other socioeconomic covariates, such as poverty and educational attainment, into account does not fully explain the geographic mortality differential and thus it is important to consider urbanicity in this study. We employed the *metropolitan status* developed by the US office of Management and Budget in 2010 to dichotomize US counties into metropolitan and nonmetropolitan counties. While the heterogeneity within each group could be great, a recent study (Yang et al. 2011) reported that the conclusions based on the metro-nonmetro dichotomy were similar to those drawn from a finer urbanicity measure (i.e., the rural-urban continuum codes). That is, the dichotomous metropolitan status provides modeling parsimony to this study.

Socioeconomic status: As discussed previously, socioeconomic status is an important factor for mortality. We will use 2005-2009 American Community Survey (ACS) to obtain a set of social indicators and apply principal component analysis to reduce the total number of variables. This variable reduction approach is comparable with that proposed by Sampson and colleagues (1997). Indeed, more specifically, the following four indicators loaded on the concept of *social affluence* (factor loading in parenthesis): the log of income per capita (0.72), the percentage of population aged 25 or over with at least a bachelor's degree (0.91), the percentage

of population working in professional, administrative, and managerial positions (0.87), and percentage of families with annual incomes higher than $75,000 (0.87). Similarly, the concept of *social disadvantage* was derived from three indicators: the poverty rate (0.72), the percentage of population receiving public assistance (0.71), and the percentage of female-head households with children below 18 (0.81). These two concepts–social affluence and social disadvantage–account for more than 70 percent of the variation among these seven indicators. We used the regression approach to generate the factor scores included in the analysis.

Racial composition: As the dependent variable, county-level mortality rate, was not standardized with the race/ethnicity structure, we included racial composition variables namely *the proportion of black, the proportion of Hispanics, and the proportion of other races*. It should be noted that racial composition is highly correlated with the five segregation measures. For example, the proportion of black is strongly associated with the white/black isolation index (Pearson's R=0.90). To avoid multicollinearity, we excluded the racial composition variable from models that include the same group-specific minority segregation measures (e.g., excluding the proportion of black in the model with all five white/black segregation indices). Furthermore, even when the proportion of Asians or Pacific Islanders is separated from the proportion of other races and included in the analysis, its association with mortality was not statistically significant (due to vary small proportions across counties) and our conclusions were not altered. For the purpose of modeling parsimony, we just presented the results using the proportion of other races.

Income inequality: The relationship between income inequality and health has drawn much attention in the literature (Lynch et al. 2004) and income inequality is closely related to racial segregation (Sethi and Somanathan 2004). To understand whether the association between segregation and mortality is independent of income inequality, we used the 2005-2009 ACS

household income data to calculate the Gini coefficient and included it in the analysis to control for the level of income inequality in a county. The Gini coefficient ranges between 0 and 1 and a larger Gini coefficient indicates a higher level of income inequality. As the top-coded category in the ACS income data is an open category ($200,000 or above), the income inequality measure may be underestimated (a common drawback when calculating the Gini coefficient with grouped, instead of individual, income data).

Social capital: As for social capital, we adopted *social capital index* developed by Rupasingha et al (2006), measuring county-level social capital based on Putnam (2001). Four indicators were used to assess the strength of social capital in a county: the number of associations (e.g., sports clubs) per 10,000 population, the number of non-profit organizations per 10,000 population, the mail response rate for the decadal census, and the presidential election voting rate. Using principal factor analysis, Rupasingha and Goetz (2008) calculated the 2005 social capital index (the latest available) for US counties and a larger social capital index suggests stronger social capital in a county. We used the 2005 social capital index in the analysis.

Population health: The county-level mortality rate is a consequence of overall population health in a county and including population health covariates helps us to better clarify the segregation-mortality association. We obtained two reliable population health measures from the University of Wisconsin Population Health Institute (2011): *average unhealthy days per month* of the population in a county and the *adult obesity rate*. The unhealthy days include both mental and physical unhealthy days based on residents' answers to the question of "how many days during the past 30 days was your physical and mental health not good." The adult obesity rate indicates the percentage of adults with a body mass index greater than 30. Both measures were originally developed by the Centers for Disease Control and Prevention (CDC) and have been

included in the Behavioral Risk Factor Surveillance Surveys conducted and maintained by CDC. The reliability and validity of these measures have been examined by CDC and the University of Wisconsin Population Health Institute. The methodologies used to obtain these county-level health measures could be found elsewhere (CDC 2000, 2012, 2013).

Methodology: Spatial Filtering

Mortality rates are not evenly distributed across the US (Cossman et al. 2007) and importantly these patterns indicate strong spatial dependence. Based on a spatial filtering approach, the spatial pattern of mortality can be decomposed into three parts: (1) a spatial trend that can be explained by a set of independent variables related to mortality, (2) a spatial process that could only be captured by the factors that are not included as an independent variable, and (3) the random disturbances. The eigenvector spatial filtering approach aims to extract distinctive spatial patterns that are not only associated with the spatial process in (2) but also account for the spatial autocorrelation in the dependent variable (i.e., mortality). This eigenvector spatial filtering approach can be adopted by most classical regression models, such as the ordinary least squares (OLS) and logistic; the estimates of the spatially filtered models would be unbiased and the interpretations of the estimates remain the same (Tiefelsdorf and Griffith 2007).

Since the dependent variable of this study is continuous, we discuss the eigenvector spatial filtering approach under the OLS framework. The basic OLS regression model for mortality can be expressed as $\boldsymbol{y} = \boldsymbol{\beta X} + \boldsymbol{\varepsilon}^*$, where $\boldsymbol{y}$ is a vector of mortality rates, $\boldsymbol{\beta}$ represents the parameters associated with a set of independent variables, $\boldsymbol{X}$, and $\boldsymbol{\varepsilon}^*$ are spatially autocorrelated errors. The eigenvector spatial filtering approach further decomposes $\boldsymbol{\varepsilon}^*$ into $\boldsymbol{E\gamma} + \boldsymbol{\varepsilon}$, where $\boldsymbol{E}$ represents a set of unspecified factors that are related to the spatial autocorrelation of mortality, $\boldsymbol{\gamma}$ is a set of estimates for $\boldsymbol{E}$ (i.e., the relationships with mortality),

and $\boldsymbol{\varepsilon}$ denotes the random disturbances. In empirical research, $\boldsymbol{E}$ is often, if not always, unknown. However, as the goal of spatial filtering is to account for spatial autocorrelation in the dependent variable, $\boldsymbol{E}$ can be created based on Moran's I (Griffith 2003; Tiefelsdorf and Griffith 2007), a commonly used measure of global spatial autocorrelation (Moran 1950).

To obtain $\boldsymbol{E}$, we create a set of dummy variables ($\boldsymbol{B}$) based on the equation below:

$$\boldsymbol{B} = (\boldsymbol{I} - \mathbf{1}\mathbf{1}^T/n), \qquad (1)$$

where $\boldsymbol{I}$ is an *n*-by-*n* identity matrix and $\mathbf{1}$ is a vector of length *n* containing ones. The superscripted *T* indicates a transposed matrix and *n* is the total number of observations. We can use $\boldsymbol{B}$ to transform the spatial weight matrix ($\boldsymbol{C}$) underlying the spatial data:

$$\boldsymbol{\Omega} = \boldsymbol{BCB} = (\boldsymbol{I} - \mathbf{1}\mathbf{1}^T/n)\boldsymbol{C}(\boldsymbol{I} - \mathbf{1}\mathbf{1}^T/n), \qquad (2)$$

where $\boldsymbol{\Omega}$ is the transformed matrix and $\boldsymbol{C}$ is the spatial weight matrix based on the spatial relationships (defined by adjacency or distance) among units. The fact that the eigenvectors of $\boldsymbol{\Omega}$ are orthogonal (i.e., uncorrelated) is the reason why $\boldsymbol{C}$ has to be transformed. The orthogonal eigenvectors indicates the unique spatial patterns filtered from the spatially autocorrelated errors.

The Moran's I value for each eigenvector given a specific spatial weight matrix ($\boldsymbol{C}$) can be expressed as a function of the eigenvalues of $\boldsymbol{\Omega}$ (Griffith 2000, 2003):

$$\text{Moran's I} = (n/\mathbf{1}^T\boldsymbol{C}\mathbf{1}) * eigenvalue(\boldsymbol{\Omega}), \qquad (3)$$

Equation (3) suggests that the Moran's I values can be computed for any numerical values of a dependent variable (e.g., mortality) in a data set with *n* observations. It should also be noted that the first eigenvector (denoted as $\boldsymbol{E_1}$) will have the largest Moran's I value given the spatial structure $\boldsymbol{C}$ and the second eigenvector will be a set of numbers that will make Moran's I statistic largest, yet smaller, than the Moran's I of the first eigenvector. Similarly, the third eigenvector includes the real numbers that generate the largest Moran's I value that is smaller than the

Moran's I of the second eigenvector. In this fashion, spatial dependence, as measured by the Moran's I, decreases as the order of eigenvector increases (Griffith 2003).

In order to tie the spatial filtering procedures above to the OLS regression, one needs to incorporate a set of independent variables ($\boldsymbol{X}$) into Equation (1) as follows:

$$\boldsymbol{B_{OLS}} = \boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T, \qquad (4)$$

where $\boldsymbol{X}$ is a matrix containing the independent variables, which is the same $\boldsymbol{X}$ included in the basic OLS regression. Multiplying $\boldsymbol{M}$ by $\boldsymbol{y}$ results in the matrix of $\boldsymbol{\varepsilon}^*$ and the eigenvectors of the transformed matrix ($\boldsymbol{MCM}$) are hence derived from and are orthogonal to the independent variables ($\boldsymbol{X}$). This $\boldsymbol{MCM}$ matrix still could be applied to Equation (3) and the spatial filtering procedures could be implemented. Importantly, the selected eigenvectors based on Moran's I can be added to the OLS model as supplementary covariates that mainly account for spatial autocorrelation in the dependent variable (Getis and Griffith 2002; Griffith 2000; Tiefelsdorf and Griffith 2007).

Since the eigenvectors of $\boldsymbol{MCM}$ can be as many as *n*, the next stage of spatial filtering is to select a parsimonious subset of eigenvectors. The conventional approaches to eigenvector selection are either to set a Moran's I threshold for inclusion (Griffith 2000) or to apply stepwise regression to positive Moran's I values (Griffith and Peres-Neto 2006). However, both of these approaches used an iterative process, and this can be computationally demanding. To address the computational demands, Tiefelsdorf and Griffith (2007) proposed to minimize a Z-score objective function of Moran's I of residuals and they found that this new approach greatly facilitates the process of eigenvector selection and removes spatial autocorrelation. Furthermore, this approach guarantees that each eigenvector included in the regression is associated with the dependent variable. In this study, we use the Tiefelsdorf and Griffith (2007) approach; details of

how to implement this approach can be found in Chun and Griffith (2013). The spatial weight matrix was constructed based on the first order Queen specification (i.e., neighbors are defined as those counties that share the same boundary or a vertex).

The analytic strategy has several steps. First we examined descriptive statistics and Moran's I tests for measuring spatial autocorrelation. Second, we estimated both OLS and spatial filtering regression models. The second step will help us to better understand whether individual dimensions of segregation are associated with mortality. In the third step, we focused on each dimension of segregation by including the same segregation dimension measures for three race/ethnic groups in one regression model, e.g., simultaneously considering white/black, white/Hispanic, and white/API entropy indices. The regression model was estimated with both OLS and spatial filtering approaches. Following Tiefelsdorf and Griffith (2007), we only consider the eigenvectors that are statistically significant and help to account for the spatial autocorrelation in mortality. We used R for all statistical analyses in this study (The R Core Development Team 2013).

Results

The descriptive statistics and Moran's I values are shown in Table 1. The average age-sex standardized mortality rate in US counties was 8.9 deaths per 1,000 population between 2004 and 2008, which is comparable with the number in a recent report (NCHS 2012). The Moran's I of mortality was 0.55, suggesting that counties with similar mortality rates tend to cluster together and the strength of spatial autocorrelation is moderate. As for the segregation measures, two findings are notable. First, regardless of dimensions, white/black segregation measures are, on average, higher than white/Hispanic and white/API segregation indices. Second, all five dimensions of segregation were found to be spatially autocorrelated. Based on Moran's I, the

exposure and centralization dimensions have the strongest and weakest levels of global autocorrelation, respectively. This pattern is consistent across the three race/ethnic groups. We also computed the Pearson's correlation coefficients between mortality and all fifteen segregation measures in Appendix A.

[Table 1 Here]

With respect to other covariates, roughly 35 percent of counties were metropolitan counties and overall (and on average), a county had approximately 9 percent of non-Hispanic blacks, 8 percent of Hispanics, and 4 percent of other minorities. The average Gini coefficient in our data, 0.43, closely matched the national level of income inequality (DeNavas-Walt, Proctor and Smith 2009). Every month, the population in a county reported approximately 7 unhealthy days and the average adult obesity rate was almost 30 percent. All of the independent variables have moderate to strong spatial autocorrelations based on a Queen first order spatial weights matrix (see Moran's I in Table 1). The spatial structure evident in our dependent and independent variables implies the need for model specifications that are explicitly spatial.

Table 2 presents the OLS and spatial filtering regression results by race/ethnicity groups. There are several important findings. First, among the white/black segregation measures, only isolation index (the exposure dimension) was found to be positively related to mortality. While this association holds in both OLS and spatial filtering models, the estimated association between isolation and mortality decreased by more than 30 percent between the OLS and the spatial filtering model ((1.051-0.709)/1.051=0.33). Second, as we argued and expected, the white/Hispanic and white/API segregation measures are negatively associated with mortality; though the evidence for this is based only on the isolation index. Two segregation measures, white/Hispanic spatial proximity index and white/API absolute centralization, follow our

expectations in the OLS models, but in the spatial filtering models there are no longer significant (see spatial filtering models in Table 2). That said, their associations with mortality in the OLS models are confounded with the eigenvectors (i.e., omitted variables) as taking the eigenvectors identified by the spatial filtering approach into account fully explained the association. Third, we visualized the spatial distributions of the three significant isolation indices and mortality rates by quintiles in Figure 1. The white/black isolation index and mortality rates share a similar pattern where the Black Belt, Mississippi Delta, and eastern Texas have both high mortality rates and white/black isolation. By contrast, white/Hispanic and white/API isolation indices are higher along with the US/Mexico border and Pacific coast, in which mortality rates are relatively low. Figure 1, to some extent, provides an explanation for why the associations of the exposure dimension of segregation with mortality differ by race/ethnic groups. Even considering other potential explanatory variables in the models, our results offered support to the bivariate visual comparison in Figure 1.

[Table 2 Here]

[Figure 1 Here]

The spatial filtering approach identified more than 50 eigenvectors and improved the adjusted R-square by approximately 20 percent from the OLS models for each race/ethnicity group. After examining these eigenvectors (results not shown but available upon request), a fourth observation from our analysis is that the first four most important eigenvectors were the same across race/ethnic groups and they are eigenvectors 15, 19, 1, and 6. Following Griffith (2003), we visualized these shared eigenvectors (based on quintiles) to gain a better understanding of what their spatial patterns are. As shown in Figure 2, the four eigenvectors have distinctive patterns. For example, eigenvector 15 suggests that the highest component values of

this particular spatial pattern largely correspond to an area west of the Mississippi from Texas in the south through to the northern Plains and Mountain West (with the exception of parts of Nebraska). The Mid-Atlantic areas also have high values. The lowest components values are found in the west coast and an area covering parts of the Rust Belt and Appalachia. Eigenvector 1shows a pattern where the high component values surround all US borders. As discussed in the method section, the four eigenvectors are independent of one another and each represents a variable at the county level not considered in the analysis. Note that in the spatial filtering models, these eigenvectors were included in the regression model specification and their estimated relationships with mortality were in Table 2. Except eigenvector 1, all other eigenvectors are positively associated with mortality and the parameter estimates are stable across models. The estimates for these eigenvectors are relatively large due to the fact the eigenvalues in each eigenvectors are relatively small (i.e., decomposition of errors).

[Figure 2 Here]

Fifth, the relationships of urbanicity, socioeconomic status and racial composition variables with mortality all reflect the findings from recent mortality literature (Sparks and Sparks 2010; Yang, Noah and Shoff Forthcoming). For instance, both OLS and spatial filtering results indicated that metropolitan counties have higher mortality rates than nonmetropolitan counterparts, the so-called rural paradox (Yang et al. 2011). Socioeconomic status variables suggested that better socioeconomic environment is associated with lower mortality, as Link and Phelan (1995) argued. Regarding racial composition, the proportion of Hispanic population was negatively related to mortality, whereas the presence of other minority groups increased mortality. It should be noted that these associations between race/ethnicity groups and mortality were found even after taking racial segregation into account.

Finally, the statistically significant associations for both income inequality and social capital index with mortality observed in the OLS models were eliminated in the spatial filtering models. That is, once spatial structure is taking into account, the relevance of social capital and income inequality to mortality is reduced; as such, our findings contribute to the ongoing debates on these topics (Kawachi et al. 1997; Lynch et al. 2004).

Following our analytic strategy, we also implemented analyses by segregation dimensions and the results were summarized in Table 3. Since the main interest of this study is segregation (and the findings related to other independent variables, such as urbanicity and income inequality, were similar to those in Table 2), we focus our discussions on segregation and spatial filtering results. Again, there are several notable findings. First and foremost, the estimated relationships between the white/black segregation measures and mortality were all positive and they were statistically significant in four of the five segregation dimensions (except clustering). By contrast, the associations of white/API segregation measures with mortality were negative, with a statistically significant association for evenness, exposure, and clustering dimensions. While the white/Hispanic entropy index was marginally significant and negatively related to mortality, overall, white/Hispanic segregation does not affect all-cause mortality (in models including other covariates). Second, the OLS models seemed to overestimate the importance of segregation, such as the findings in the evenness and clustering dimensions. The spatial structure underlying the data contributes to this overestimation as the spatial filtering models generated weaker relationships between segregation measures and mortality, and improved the adjusted R-square.

[Table 3 Here]

Third, the total number of eigenvectors found in each model is comparable across segregation dimensions and, among them, six were commonly shared by the five segregation dimensions, i.e., eigenvectors 15, 19, 1, 6, 21, and 17. Comparing with the findings in Table 2, two additional eigenvectors, 21 and 17, were identified and they were shown in Figure 3. Again, both eigenvectors have spatial patterns that were different from those in Figure 2 and their associations with mortality were positive (see Table 3). As seen in Figures 2 and 3, the Moran's I values of the six most important eigenvectors were all positive, indicating that the high (low) component values of these eigenvectors are geographically close to one another. We would like to reiterate that the six eigenvectors capture the spatial processes that are not associated with the independent variables in the models but they contribute to the observed spatial pattern of mortality in the US.

[Figure 3 Here]

Last, as the spatial filtering approach aims to remove spatial autocorrelation in the dependent variable, we conducted Moran's I tests to assess if the residuals of these models are still spatially autocorrelated. The last row of Tables 2 and 3 indicated that spatial filtering effectively reduces spatial autocorrelation by approximately 85 percent from OLS models. The residuals' Moran's I values in spatial filtering models are all very close to zero (i.e., no spatial autocorrelation), while they remained statistically significant. The explanation for the statistical significance is that our eigenvector sets are optimal (based on statistical significance), rather than exhaustive (Chun and Griffith 2013). When we included those eigenvectors with a p-value between 0.05 and 0.1, the residuals' Moran's I became non-significant (results not show but available).

Discussion and Conclusions

We used the findings above to examine the research hypotheses. Following the ethnic stratification perspective, we first hypothesized that white/black segregation is positively related to mortality. When taking all five segregation dimensions into account (Table 2), only the exposure dimension was found significantly and positively related to mortality. Nonetheless, the dimension-specific analyses (Table 3) offered stronger evidence to support the first hypothesis as four out of the five white/black segregation measures followed our expectation. Our second hypothesis suggested that white/Hispanic and white/API segregation are beneficial to mortality as these minority groups tend to self-segregate from non-Hispanic whites. The results in Table 2 suggested that higher levels of isolation between non-Hispanic whites and these two minority groups are associated with lower levels of mortality in US counties. Though white/Hispanic spatial proximity index (clustering dimension) and white/API absolute centralization index (centralization dimension) supported the second hypothesis in the OLS models, these findings did not hold when the spatial structure underlying the data was considered. When considering one dimension of segregation at a time, we obtained stronger evidence for white/API than white/Hispanic segregation measures. Specifically, three of the five white/API dimensions of segregation were negatively related to mortality, whereas only the white/Hispanic entropy index was marginally significant.

The third hypothesis indicated that the spatial autocorrelation affects the estimates of the relationships between segregation and mortality and spatial filtering approach would identify the spatial patterns that are not only related to county-level mortality but also shared by various segregation dimensions. This hypothesis was confirmed as the OLS models tend to overestimate the importance of segregation (see both Tables 2 and 3). Furthermore, the race/ethnicity-specific models shared four eigenvectors and the dimension-specific analyses identified two additional

eigenvectors. The six eigenvectors have distinctive spatial patterns and each of them represents a dimension not included in the model. While it is not clear what these covariates may be, they provide scientific insights into future mortality studies as researchers could explore what factors correspond to these spatial patterns (Griffith 2003). For example, Eigenvector 1 in Figure 2 seems to suggest the edge effect that surrounds the US boundaries. Researchers may need to identify another variable that captures this unique spatial pattern or simply create a dummy variable to separate the counties at the edge from others.

Overall, we believe that our hypotheses received sufficient support, especially from the spatial filtering models. Several important discrepancies between our hypotheses and findings were in evidence. First, the race/ethnicity-specific analyses suggested that the relationship of isolation index with mortality is the most consistent among other segregation measures, which suggests that exposure to non-Hispanic whites may be the most important dimension of segregation. As Massey and Denton (1988) defined, the exposure dimension refers to the extent to which minority and majority group members interact within a given area and the isolation index captures the level of segregation experienced by minorities. This definition fits the ethnic stratification and ethnic community/enclave perspectives and the opposite associations between non-Hispanic blacks and other minority groups were expected. Furthermore, the evenness and exposure dimensions of segregation were more closely related to mortality than the other three dimensions (Table 3). Our findings seemed to justify a focus on evenness and exposure dimensions of segregation in the literature on segregation and health (Kramer and Hogue 2009).

Second, the relationships between white/Hispanic segregation and mortality were weakly supported by our analytic results. Though we suspected that white/Hispanic segregation measures may be highly correlated with racial composition or other independent variables, such

as social capital and income inequality, our sensitivity analyses (not shown) where these variables were excluded did not support this explanation. Thus, the possible explanation would be that white/Hispanic segregation measures are associated with the spatial processes that are not included or captured by our models.

Some of our findings speak to but do not resolve some debates in mortality inequality research. We found that income inequality and social capital are not important determinants of mortality when using a spatial filtering approaches, which contradicts several recent studies using other analysis techniques, such as spatial econometrics and weighted least squares modeling (McLaughlin and Stokes 2002; McLaughlin et al. 2007; Yang et al. 2011; Yang et al. 2009). As spatial filtering has not been commonly used in county-level mortality research, it is possible that the differences that emerge are the result of the use of different methods. Spatial filtering mainly focuses on the decomposition of error terms and the eigenvectors can be as many as the total number of observations in a study (Griffith 2003). By contrast, other conventional spatial models, such as spatial econometrics modeling, aim to provide an overall assessment of the impact of spatial autocorrelation on the dependent variable (LeSage and Pace 2009). This methodological issue is beyond the scope of this study but we recognize exploring this would be a fruitful future direction.

This study contributes to the mortality literature in the following ways. Only a few mortality studies have employed segregation as an explanation for mortality differentials across the US counties or cities (Collins and Williams 1999; Sparks et al. 2013) and even less has focused on the segregation between whites and non-black minority groups. Using the ethnic stratification and ethnic community/enclave perspective, we argued that white/black segregation is detrimental to overall mortality but white/Hispanic and white/API segregation are beneficial.

Our arguments were generally supported by the analytic results, which sheds new light on geographic mortality disparities. In addition, the spatial filtering approach identified several common eigenvectors that demonstrate unique spatial patterns in US counties. These eigenvectors represent missing variables that have implications for mortality but cannot be captured with any of the independent variables in our models. That said, mortality researchers should think outside the box to find determinants of geographic mortality differentials and these unique patterns offer some clues.

There are several limitations specific to this study. First, the ecological relationships between segregation measures and mortality cannot be generalized to individuals, and we would urge health researchers to investigate how the health of non-black minorities is affected by segregation from non-Hispanic whites. Second, this study combined multiple data sources from different time points to explore the associations between the independent and dependent variables. No causality can be derived from our analyses and the temporal misalignment should be noted. Third, like other ecological studies, our analysis is subject to the modifiable area unit problem (Openshaw 1984) as changing the unit of analysis (e.g., using health service areas) may alter the findings and conclusions. Fourth, this study used all-cause age-sex standardized mortality as the dependent variable as it is an overall evaluation of population health in a county. While race-specific mortality rates can be calculated, we encountered the small area/population estimation issues (Ghosh and Rao 1994) as numerous counties have zero death for non-black minorities. Finally, while our findings suggested that evenness and exposure dimensions are more important than others, future research may still need to investigate whether hypersegregation (Massey and Denton 1989; Osypuk and Acevedo-Garcia 2008) is a more useful measure in health research. As one of our goals is to understand whether the relationship

between segregation and mortality varies by segregation dimensions, the analysis using hypersegregation is beyond the scope of this study.

Some future research directions can be drawn from this study. First, our discussion on ethnic enclave and community is relevant to the literature on immigrant health (Jasso et al. 2004; Kandula, Kersey and Lurie 2004). The Compressed Mortality Files used in this study do not include the information on nativity of the deceased, which prevents us from directly addressing this issue. Future studies should use other data sources to investigate the relationship among immigration, segregation, and health. Second, most mortality research used the latest available data to explain the mortality differentials across social groups and long-term latency between disease onset and death has been overlooked (Matthews 1990). That is, the mortality differentials observed today may be the result of the socioeconomic or environmental factors in existence decades prior rather than those measured concurrently with mortality. Addressing the latency issue may better clarify the causality between the persistent mortality pattern and its determinants. Finally, the segregation measures used in this study only concern two race/ethnic groups. The measures of multi-group segregation (Reardon and Firebaugh 2002) should be employed to understand whether the measures of segregation alter the results/conclusions.

In sum, racial segregation is argued to be the major cause of health disparities (Williams and Collins 2001) and a determinant of health outcomes (Kramer and Hogue 2009). Previous evidence has been drawn heavily from the relationships between non-Hispanic blacks and white/black segregation. A growing body of literature has found that segregation may be beneficial to health outcomes or behaviors for non-black minorities, particularly Hispanics and Asians/Pacific Islanders (Osypuk et al. 2009; Walton 2009; Yang et al. 2014). This study echoes

the recent development in the literature and offers county-level evidence for the potential benefit of segregation for Hispanics and Asians/Pacific Islanders.

References

Acevedo-Garcia, D., K.A. Lochner, T.L. Osypuk, and S.V. Subramanian. 2003. "Future directions in residential segregation and health research: a multilevel approach." *American journal of public health* 93(2):215-221.

Bécares, L., J. Nazroo, and M. Stafford. 2009. "The buffering effects of ethnic density on experienced racism and health." *Health & Place* 15(3):700-708.

Blanchard, T.C., J.P. Bartkowski, T.L. Matthews, and K.R. Kerley. 2008. "Faith, morality and mortality: The ecological impact of religion on population health." *Social Forces* 86(4):1591-1620.

Brulle, R.J.and D.N. Pellow. 2006. "Environmental justice: human health and environmental inequalities." *Annu. Rev. Public Health* 27:103-124.

CDC. 2000. "Measuring Healthy Days: Population Assessment of health-related Quality of Life." edited by D.o.H.a.H. Services. Atlanta, GA.

—. 2012. "Overweight and Obesity: Causes and Consequences."

—. 2013. "Behavioral Risk Factor Surveillance System."

Chun, Y.and D.A. Griffith. 2013. *Spatial Statistics and Geostatistics: Theory and Applications for Geographic Information Science and Technology*: SAGE.

Collins, C.A.and D.R. Williams. 1999. "Segregation and mortality: The deadly effects of racism?" Presented at Sociological Forum.

Cossman, J.S., R.E. Cossman, W.L. James, C.R. Campbell, T.C. Blanchard, and A.G. Cosby. 2007. "Persistent clusters of mortality in the United States." *American journal of public health* 97(12):2148.

Cossman, R.E., J.S. Cossman, A.G. Cosby, and R.M. Reavis. 2008. "Reconsidering the rural–urban continuum in rural health research: a test of stable relationships using mortality as a health measure." *Population Research and Policy Review* 27(4):459-476.

Cressie, N.A.C. 1991. *Statistics for Spatial Data*. London, UK: John Wiley & Sons, Inc.

—. 1993. "Statisticsfor spatial data." *John Willey&Sons*.

DeNavas-Walt, C., B. Proctor, and J. Smith. 2009. "Income, poverty, and health insurance coverage in the United States: 2008." edited by U.S.C. Bureau. Washington, DC: U.S. Government Printing Office.

Duncan, O.D., R.P. Cuzzort, and B. Duncan. 1961. "Statistical geography: Problems in analyzing areal data." *Statistical geography: Problems in analyzing areal data.*

Eschbach, K., G.V. Ostir, K.V. Patel, K.S. Markides, and J.S. Goodwin. 2004. "Neighborhood context and mortality among older Mexican Americans: is there a barrio advantage?" *Journal Information* 94(10).

Fischer, M.J. 2003. "The relative importance of income and race in determining residential outcomes in US urban areas, 1970-2000." *Urban Affairs Review* 38(5):669-696.

Fuguitt, G.V. 1994. *Population change in nonmetropolitan America*: Center for Demography and Ecology, University of Wisconsin--Madison.

Getis, A.and D.A. Griffith. 2002. "Comparative spatial filtering in regression analysis." *Geographical Analysis* 34(2):130-140.

Ghosh, M.and J. Rao. 1994. "Small area estimation: an appraisal." *Statistical Science*:55-76.

Glaeser, E.and J. Vigdor. 2012. *The End of the Segregated Century: Racial Separation in America's Neighborhoods, 1890-2010*: Manhattan Institute for Policy Research.

Greenberg, M.and D. Schneider. 1994. "Violence in American cities: young black males in the answer, but what was the question?" *Social Science & Medicine* 39(2):179-187.

Griffith, D.A. 2000. "Eigenfunction properties and approximations of selected incidence matrices employed in spatial analyses." *Linear Algebra and its Applications* 321(1):95-112.
—. 2003. *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*: Springer.
Griffith, D.A.and P.R. Peres-Neto. 2006. "Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses." *Ecology* 87(10):2603-2613.
Hobbs, F.and N. Stoops. 2002. *Demographic trends in the 20th century*: US Census Bureau.
Hoyert, D.L. 2012. "75 Years of Mortality in the United States, 1935–2010." in *NCHS Data Brief*, edited by NCHS.
Iceland, J., D.H. Weinberg, and E. Steinmetz. 2002. *Racial and ethnic residential segregation in the United States 1980-2000*: Bureau of Census.
Jasso, G., D.S. Massey, M.R. Rosenzweig, and J.P. Smith. 2004. "Immigrant health: selectivity and acculturation." *Critical perspectives on racial and ethnic differences in health in late life*:227-266.
Johnson, K.M.and D.T. Lichter. 2010. "Growing diversity among America's children and youth: Spatial and temporal dimensions." *Population and Development Review* 36(1):151-176.
Kandula, N.R., M. Kersey, and N. Lurie. 2004. "Assuring the health of immigrants: what the leading health indicators tell us." *Annu. Rev. Public Health* 25:357-376.
Kawachi, I., B.P. Kennedy, K. Lochner, and D. Prothrow-Stith. 1997. "Social capital, income inequality, and mortality." *American journal of public health* 87(9):1491-1498.
Kramer, M.R.and C.R. Hogue. 2009. "Is segregation bad for your health?" *Epidemiologic reviews* 31(1):178-194.
LeClere, F.B., R.G. Rogers, and K.D. Peters. 1997. "Ethnicity and mortality in the United States: individual and community correlates." *Social Forces* 76(1):169-198.
Lee, B.A., S.F. Reardon, G. Firebaugh, C.R. Farrell, S.A. Matthews, and D. O'Sullivan. 2008. "Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales." *American Sociological Review* 73(5):766-791.
Lee, M.-A.and K.F. Ferraro. 2007. "Neighborhood residential segregation and physical health among Hispanic Americans: Good, bad, or benign?" *Journal of health and social behavior* 48(2):131-148.
LeSage, J.P.and R.K. Pace. 2009. *Introduction to Spatial Econometrics*: Chapman & Hall/CRC.
Link, B.G.and J. Phelan. 1995. "Social conditions as fundamental causes of disease." *Journal of health and social behavior* 35:80-94.
Lobao, L.M., G. Hooks, and A.R. Tickamyer. 2007. *The Sociology of Spatial Inequality*: State University of New York Press.
Lochner, K.A., I. Kawachi, R.T. Brennan, and S.L. Buka. 2003. "Social capital and neighborhood mortality rates in Chicago." *Social Science & Medicine* 56(8):1797-1805.
Logan, J.R. 1978. "Growth, politics, and the stratification of places." *American Journal of Sociology*:404-416.
Logan, J.R., B.J. Stults, and R. Farley. 2004. "Segregation of minorities in the metropolis: Two decades of change." *Demography* 41(1):1-22.
Logan, J.R., W. Zhang, and R.D. Alba. 2002. "Immigrant enclaves and ethnic communities in New York and Los Angeles." *American Sociological Review*:299-322.
Lynch, J., G.D. Smith, S. Harper, M. Hillemeier, N. Ross, G.A. Kaplan, and M. Wolfson. 2004. "Is income inequality a determinant of population health? Part 1. A systematic review." *Milbank Quarterly* 82(1):5-99.

Massey, D.S.and N.A. Denton. 1988. "The dimensions of residential segregation." *Social Forces* 67(2):281-315.
—. 1989. "Hypersegregation in US metropolitan areas: Black and Hispanic segregation along five dimensions." *Demography* 26(3):373-391.
Matthews, S.A. 1990. "Epidemiology using a GIS: the need for caution." *Computers, Environment and Urban Systems* 14(3):213-221.
Matthews, S.A.and D.M. Parker. 2013. "Progress in Spatial Demography." *Demographic Research* 28.
McLaughlin, D.K.and C.S. Stokes. 2002. "Income inequality and mortality in US counties: does minority racial concentration matter?" *American journal of public health* 92(1):99-104.
McLaughlin, D.K., C.S. Stokes, P.J. Smith, and A. Nonoyama. 2007. "Differential mortality across the U.S.: The influence of place-based inequality." Pp. 141-162 in *The sociology of spatial inequality*, edited by Linda M. Lobao, Gregory Hooks, and A.R. Tickamyer. Albany, NY: SUNY Press.
Moran, P.A.P. 1950. "Notes on continuous stochastic phenomena." *Biometrika* 37(1/2):17-23.
Murdock, S.H., S.S. Hwang, and M. Hoque. 1994. "Nonmetropolitan residential segregation revisited." *Rural Sociology* 59(2):236-254.
NCHS. 2010. "Compressed Mortality File, 1999-2007 (machine readable data file and documentation, CD-ROM series 20, No.2M)." Hyattsville, Maryland: National Center for Health Statistics.
—. 2012. "Health, United States, 2011: With Special Feature on Socioeocnomic Status and Health." Hyattsville, MD: Library of Congress Catalog Number 76-641496.
Openshaw, S. 1984. "Ecological fallacies and the analysis of areal census data." *Environment and planning A* 16(1):17-31.
Osypuk, T.L.and D. Acevedo-Garcia. 2008. "Are racial disparities in preterm birth larger in hypersegregated areas?" *American Journal of Epidemiology* 167(11):1295-1304.
Osypuk, T.L., A.V. Diez Roux, C. Hadley, and N.R. Kandula. 2009. "Are immigrant enclaves healthy places to live? The Multi-ethnic Study of Atherosclerosis." *Social Science & Medicine* 69(1):110-120.
Parisi, D., D.T. Lichter, and M.C. Taquino. 2011. "Multi-scale residential segregation: Black exceptionalism and America's changing color line." *Social Forces* 89(3):829-852.
Putnam, R.D. 2001. *Bowling alone: The collapse and revival of American community*: Simon and Schuster.
Reardon, S.F.and G. Firebaugh. 2002. "Measures of multigroup segregation." *Sociological methodology* 32(1):33-67.
Rupasingha, A.and S.J. Goetz. 2008. "US County-Level Social Capital Data, 1990-2005.". University Park: The Northeast Regional Center for Rural Development, Penn State University.
Rupasingha, A., S.J. Goetz, and D. Freshwater. 2006. "The production of social capital in US counties." *Journal of Socio-Economics* 35(1):83-101.
Sampson, R.J., S.W. Raudenbush, and F. Earls. 1997. "Neighborhoods and violent crime: A multilevel study of collective efficacy." *Science* 277(5328):918-924.
Sethi, R.and R. Somanathan. 2004. "Inequality and segregation." *Journal of Political Economy* 112(6):1296-1321.
Song, L., J. Son, and N. Lin. 2010. "Social Capital and Health." Pp. 184-210 in *The New Blackwell Companion to Medical Sociology*, edited by W.C. Cockerham. Malden, MA: Blackwell.

Sparks, J.P., C.S. Sparks, and J.J.A. Campbell. 2013. "An application of Bayesian spatial statistical methods to the study of racial and poverty segregation and infant mortality rates in the US." *Geojournal* 78:389-405.
Sparks, P.J.and C.S. Sparks. 2010. "An application of spatially autoregressive models to the study of US county mortality rates." *Population, Space and Place* 16(6):465-481.
Thayn, J.B.and J.M. Simanis. 2013. "Accounting for spatial autocorrelation in linear regression models using spatial filtering with eigenvectors." *Annals of the Association of American Geographers* 103(1):47-66.
The R Core Development Team. 2013. "R: A language and environment for statistical computing." ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. url: http://www.R-project.org.
Theil, H. 1972. *Statistical decomposition analysis: With applications in the social and administrative sciences*: North-Holland Publishing Company Amsterdam.
Tiefelsdorf, M.and D.A. Griffith. 2007. "Semiparametric filtering of spatial autocorrelation: the eigenvector approach." *Environment and planning A* 39(5):1193.
University of Wisconsin Population Health Institute. 2011. "County Health Rankings."
Voss, P.R. 2007. "Demography as a spatial social science." *Population Research and Policy Review* 26(5):457.
Voss, P.R., D.D. Long, R.B. Hammer, and S. Friedman. 2006. "County child poverty rates in the US: a spatial regression approach." *Population Research and Policy Review* 25(4):369-391.
Walton, E. 2009. "Residential segregation and birth weight among racial and ethnic minorities in the United States." *Journal of health and social behavior* 50(4):427-442.
White, M.J. 1986. "Segregation and diversity measures in population distribution." *Population index* 52(2):198-221.
Whitley, R., M. Prince, K. McKenzie, and R. Stewart. 2006. "Exploring the ethnic density effect: a qualitative study of a London electoral ward." *International journal of social psychiatry* 52(4):376-391.
Wilkes, R.and J. Iceland. 2004. "Hypersegregation in the twenty-first century." *Demography* 41(1):23-36.
Williams, D.R.and C. Collins. 2001. "Racial residential segregation: a fundamental cause of racial disparities in health." *Public health reports* 116(5):404.
Yang, T.C., Y.J. Chen, C. Shoff, and S.A. Matthews. 2012. "Using quantile regression to examine the effects of inequality across the mortality distribution in the US." *Social Science & Medicine* 74(12):1900-1910.
Yang, T.C., L. Jensen, and M. Haran. 2011. "Social capital and human mortality: Explaining the rural paradox with county-level mortality data." *Rural Sociology* 76(3):347-374.
Yang, T.C., A.J. Noah, and C. Shoff. Forthcoming. "Exploring Geographic Variation in US Mortality Rates Using a Spatial Durbin Approach." *Population, Space and Place*.
Yang, T.C., C. Shoff, A.J. Noah, N. Black, and C.S. Sparks. 2014. "Racial segregation and maternal smoking during pregnancy: a multilevel analysis using the racial segregation interaction index." *Social Science & Medicine* 107:26-36.
Yang, T.C., H.W. Teng, and M. Haran. 2009. "The Impacts of Social Capital on Infant Mortality in the US: A Spatial Investigation." *Applied spatial analysis and policy* 2(3):211.

Table 1. Descriptive Statistics of the Variables of This Study

	Min.	Max.	Mean	S.D.	Moran's I†
Dependent Variable					
Mortality (per 1,000 population)	2.904	18.889	8.913	1.453	0.554
Independent Variables					
Segregation					
Non-Hispanic Black					
Entropy (evenness)	0.000	0.691	0.096	0.105	0.320
Isolation Index (exposure)	0.000	0.890	0.152	0.200	0.665
Absolute Centralization (centralization)	-0.763	0.955	0.271	0.299	0.099
Delta (concentration)	0.000	0.990	0.518	0.257	0.202
Spatial Proximity (clustering)	1.000	1.749	1.049	0.081	0.313
Hispanic					
Entropy (evenness)	0.000	0.442	0.054	0.068	0.306
Isolation Index (exposure)	0.001	0.971	0.128	0.169	0.662
Absolute Centralization (centralization)	-0.585	0.958	0.249	0.271	0.102
Delta (concentration)	0.000	0.945	0.460	0.235	0.159
Spatial Proximity (clustering)	1.000	1.911	1.030	0.058	0.324
Asians/Pacific Islanders					
Entropy (evenness)	0.000	0.426	0.043	0.043	0.230
Isolation Index (exposure)	0.000	0.632	0.028	0.058	0.446
Absolute Centralization (centralization)	-0.561	0.970	0.266	0.285	0.088
Delta (concentration)	0.000	0.986	0.496	0.242	0.132
Spatial Proximity (clustering)	1.000	1.368	1.008	0.022	0.350
Urbanicity					
Metropolitan Status (1=yes, 0=no)	0.000	1.000	0.349	0.477	N.A.
Socioeconomic Status					
Social Affluence	-7.676	5.382	0.000	1.000	0.479
Social Disadvantage	-2.842	10.166	0.000	1.000	0.362
Racial Composition					
% Non-Hispanic Black	0.000	0.868	0.089	0.144	0.802
% Hispanic	0.000	0.986	0.076	0.128	0.769
% other races	0.000	0.910	0.040	0.069	0.414
Income Inequality					
Gini	0.272	0.621	0.431	0.037	0.309
Social Capital					
Social Capital Index	-3.804	15.222	-0.002	1.642	0.580
Population Health					
Unhealthy Days per Month	0.000	16.500	6.822	2.573	0.391
Adult Obesity Rate	11.700	43.700	28.930	3.700	0.653

†All Moran's I values are statistically significant at 0.01 level.

Table 2. OLS and Spatial Filtering Results by Race/ethnicity

	Non-Hispanic Blacks						Hispanics						Asians/Pacific Islanders					
	OLS			Spatial Filtering			OLS			Spatial Filtering			OLS			Spatial Filtering		
	β	S.E.		β	S.E.		β	S.E.		β	S.E.		β	S.E.		β	S.E.	
Intercept	5.827	0.571	***	7.506	0.524	***	6.575	0.669	***	6.975	0.614	***	4.320	1.713	*	5.976	1.531	***
Segregation																		
Entropy (evenness)	-0.462	0.357		-0.158	0.327		0.293	0.509		0.249	0.493		-0.027	0.760		-0.278	0.699	
Isolation Index (exposure)	1.051	0.173	***	0.709	0.171	***	-1.193	0.166	***	-1.406	0.182	***	-2.532	0.717	***	-1.369	0.653	*
Absolute Centralization (centralization)	-0.118	0.079		-0.034	0.070		-0.107	0.091		-0.071	0.081		-0.170	0.085	*	-0.091	0.076	
Delta (concentration)	-0.066	0.103		0.161	0.093		-0.088	0.110		0.163	0.100		-0.063	0.111		0.143	0.100	
Spatial Proximity (clustering)	-0.267	0.466		-0.400	0.413		-1.135	0.578	*	-0.044	0.525		1.265	1.687		1.023	1.506	
Urbanicity																		
Metropolitan Status	0.111	0.048	*	0.095	0.043	*	0.153	0.047	**	0.113	0.043	**	0.148	0.047	**	0.112	0.042	**
Socioeconomic Status																		
Social Affluence	-0.485	0.027	***	-0.501	0.025	***	-0.443	0.027	***	-0.460	0.025	***	-0.406	0.028	***	-0.440	0.026	***
Social Disadvantage	0.330	0.028	***	0.362	0.025	***	0.291	0.028	***	0.326	0.026	***	0.351	0.023	***	0.435	0.022	***
Racial Composition																		
% Non-Hispanic Black							1.652	0.175	***	1.377	0.197	***	1.306	0.161	***	0.524	0.167	**
% Hispanic	-2.097	0.164	***	-2.684	0.177	***							-1.704	0.173	***	-2.580	0.189	***
% Other Races	0.401	0.324		1.667	0.330	***	1.104	0.328	***	2.352	0.339	***						
Income Inequality																		
Gini	2.672	0.562	***	-0.081	0.531		2.637	0.558	***	-0.062	0.535		2.768	0.558	***	0.167	0.534	
Social Capital																		
Social Capital Index	-0.158	0.014	***	-0.016	0.014		-0.151	0.014	***	-0.005	0.014		-0.166	0.014	***	-0.033	0.014	*
Population Health																		
Unhealthy Days/Month	0.067	0.008	***	0.051	0.008	***	0.079	0.008	***	0.055	0.008	***	0.072	0.008	***	0.051	0.008	***
Adult Obesity Rates	0.063	0.007	***	0.050	0.007	***	0.062	0.007	***	0.052	0.007	***	0.060	0.007	***	0.055	0.007	***
Eigenvector 15		N.A.		7.010	0.873	***		N.A.		6.995	0.876	***		N.A.		6.857	0.870	***
Eigenvector 19		N.A.		10.505	0.890	***		N.A.		10.688	0.893	***		N.A.		10.647	0.892	***
Eigenvector 1		N.A.		-11.752	0.920	***		N.A.		-10.808	0.932	***		N.A.		-11.419	0.927	***
Eigenvector 6		N.A.		11.124	0.981	***		N.A.		11.023	0.992	***		N.A.		10.478	0.984	***
Total Eigenvectors‡		N.A.			55			N.A.			58			N.A.			52	
Adjusted R-square		0.551			0.657			0.552			0.656			0.557			0.657	
Residuals' Moran's I		0.243	***		0.038	***		0.238	***		0.037	***		0.237	***		0.036	***

***p<0.05; **p<0.01; ***p<0.001; ‡All eigenvectors are statistically significant (p<0.05). N.A.: Not Applicable**

Table 3. OLS and Spatial Filtering Results by Segregation Dimensions

	Evenness (entropy)						Exposure (isolation index)						Centralization (absolute centralization index)					
	OLS			Spatial Filtering			OLS			Spatial Filtering			OLS			Spatial Filtering		
	β	S.E.		β	S.E.		β	S.E.		β	S.E.		β	S.E.		β	S.E.	
Intercept	5.454	0.327	***	7.076	0.321	***	5.552	0.326	***	7.112	0.325	***	5.533	0.326	***	7.157	0.321	***
Segregation																		
Non-Hispanic Black/White	0.841	0.226	***	0.624	0.208	**	0.592	0.257	*	0.486	0.238	*	0.132	0.134		0.238	0.119	*
Hispanic/White	-1.759	0.359	***	-0.577	0.331	+	-0.144	0.352		0.286	0.321		-0.126	0.167		-0.127	0.148	
Asians/White	-1.826	0.520	***	-1.098	0.468	*	-2.687	0.433	***	-2.275	0.406	***	-0.220	0.144		-0.115	0.127	
Urbanicity																		
Metropolitan Status	0.127	0.047	**	0.110	0.043	**	0.128	0.048	**	0.111	0.043	**	0.132	0.047	**	0.114	0.042	**
Socioeconomic Status																		
Social Affluence	-0.453	0.027	***	-0.474	0.025	***	-0.441	0.027	***	-0.451	0.025	***	-0.474	0.027	***	-0.490	0.025	***
Social Disadvantage	0.302	0.028	***	0.354	0.026	***	0.296	0.028	***	0.333	0.026	***	0.317	0.028	***	0.359	0.026	***
Racial Composition																		
% Non-Hispanic Black	1.477	0.174	***	0.790	0.179	***	0.894	0.331	**	0.626	0.337	+	1.299	0.173	***	0.771	0.177	***
% Hispanics	-1.648	0.173	***	-2.397	0.190	***	-1.522	0.437	***	-2.298	0.426	***	-1.941	0.167	***	-2.576	0.179	***
% Other Races	0.939	0.329	**	1.880	0.336	***	1.122	0.343	**	2.225	0.353	***	0.581	0.329	+	1.751	0.336	***
Income Inequality																		
Gini	2.746	0.559	***	0.026	0.533		2.697	0.558	***	0.082	0.534		2.694	0.556	***	-0.031	0.529	
Social Capital																		
Social Capital Index	-0.156	0.014	***	-0.020	0.014		-0.158	0.014	***	-0.017	0.014		-0.158	0.014	***	-0.019	0.014	
Population Health																		
Unhealthy Days/Month	0.071	0.008	***	0.055	0.007	***	0.069	0.008	***	0.054	0.008	***	0.069	0.008	***	0.055	0.007	***
Adult Obesity Rates	0.062	0.007	***	0.051	0.007	***	0.058	0.007	***	0.046	0.007	***	0.061	0.007	***	0.049	0.007	***
Eigenvector 15		N.A.		7.003	0.870	***		N.A.		7.307	0.873	***		N.A.		6.795	0.869	***
Eigenvector 19		N.A.		10.409	0.886	***		N.A.		10.662	0.887	***		N.A.		10.490	0.887	***
Eigenvector 1		N.A.		-11.039	0.923	***		N.A.		-10.658	0.918	***		N.A.		-11.457	0.913	***
Eigenvector 6		N.A.		11.130	0.977	***		N.A.		10.541	0.985	***		N.A.		10.959	0.979	***
Eigenvector 21		N.A.		6.883	0.887	***		N.A.		6.942	0.882	***		N.A.		6.901	0.876	***
Eigenvector 17		N.A.		6.801	0.869	***		N.A.		6.615	0.999	***		N.A.		6.936	0.868	***
Total Eigenvectors‡		N.A.			53			N.A.			54			N.A.			54	
Adjusted R-square		0.558			0.659			0.558			0.662			0.553			0.658	
Residuals' Moran's I		0.237	***		0.036	***		0.244	***		0.033	***		0.243	***		0.038	***

+p<0.1; *p<0.05; **p<0.01; ***p<0.001; ‡All eigenvectors are statistically significant (p<0.05). N.A.: Not Applicable

Table 3. OLS and Spatial Filtering Results by Segregation Dimensions (cont.)

	Concentration (Delta index)						Clustering (spatial proximity index)					
	OLS			Spatial Filtering			OLS			Spatial Filtering		
	β	S.E.		β	S.E.		β	S.E.		β	S.E.	
Intercept	5.621	0.327	***	7.123	0.320	***	10.282	0.971	***	9.653	0.910	***
Segregation												
Non-Hispanic Black/White	0.200	0.134		0.285	0.120	*	0.589	0.278	*	0.344	0.256	
Hispanic/White	-0.445	0.162	**	-0.103	0.146		-1.716	0.378	***	-0.429	0.346	
Asians/White	-0.003	0.149		-0.066	0.133		-3.632	0.959	***	-2.576	0.878	**
Urbanicity												
Metropolitan Status	0.125	0.047	**	0.108	0.042	*	0.138	0.047	**	0.118	0.042	**
Socioeconomic Status												
Social Affluence	-0.476	0.027	***	-0.495	0.025	***	-0.458	0.027	***	-0.469	0.025	***
Social Disadvantage	0.312	0.028	***	0.353	0.026	***	0.296	0.028	***	0.351	0.026	***
Racial Composition												
% Non-Hispanic Black	1.336	0.178	***	0.858	0.180	***	1.388	0.182	***	0.885	0.200	***
% Hispanics	-1.882	0.169	***	-2.570	0.181	***	-1.520	0.182	***	-2.272	0.203	***
% Other Races	0.663	0.330	*	1.826	0.335	***	0.987	0.330	**	1.938	0.339	***
Income Inequality												
Gini	2.574	0.557	***	-0.081	0.528		2.703	0.556	***	0.126	0.530	
Social Capital												
Social Capital Index	-0.153	0.014	***	-0.015	0.014		-0.157	0.014	***	-0.019	0.014	
Population Health												
Unhealthy Days	0.069	0.008	***	0.053	0.008	***	0.071	0.008	***	0.057	0.007	***
Adult Obesity Rates	0.061	0.007	***	0.049	0.007	***	0.059	0.007	***	0.051	0.007	***
Eigenvector 15		N.A.		6.861	0.867	***		N.A.		6.804	0.871	***
Eigenvector 19		N.A.		10.564	0.887	***		N.A.		10.541	0.887	***
Eigenvector 1		N.A.		-11.639	0.922	***		N.A.		-10.968	0.921	***
Eigenvector 6		N.A.		11.093	0.979	***		N.A.		10.772	0.985	***
Eigenvector 21		N.A.		7.167	0.883	***		N.A.		6.693	0.880	***
Eigenvector 17		N.A.		7.105	0.869	***		N.A.		6.754	0.869	***
Total Eigenvectors[†]		N.A.			55			N.A.			55	
Adjusted R-square		0.553			0.659			0.558			0.660	
Residuals' Moran's I		0.241	***		0.037	***		0.239	***		0.033	***

+p<0.1; *p<0.05; **p<0.01; ***p<0.001; ‡All eigenvectors are statistically significant (p<0.05). N.A.: Not Applicable

Figure 1. Spatial Distributions of mortality and three isolation indices

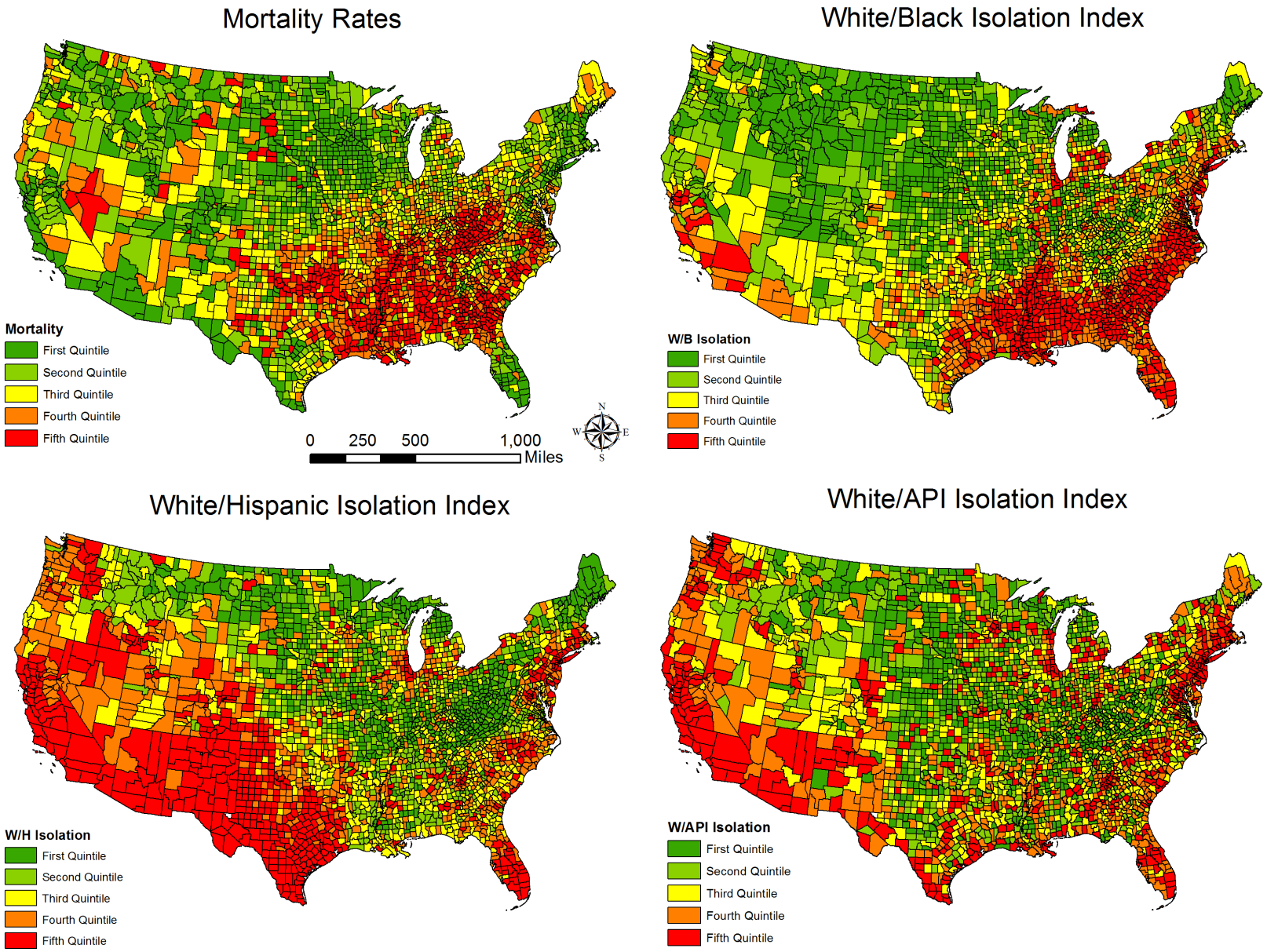


Figure 2. Spatial Patterns of the First Four Common Eigenvectors and their Moran's I Values

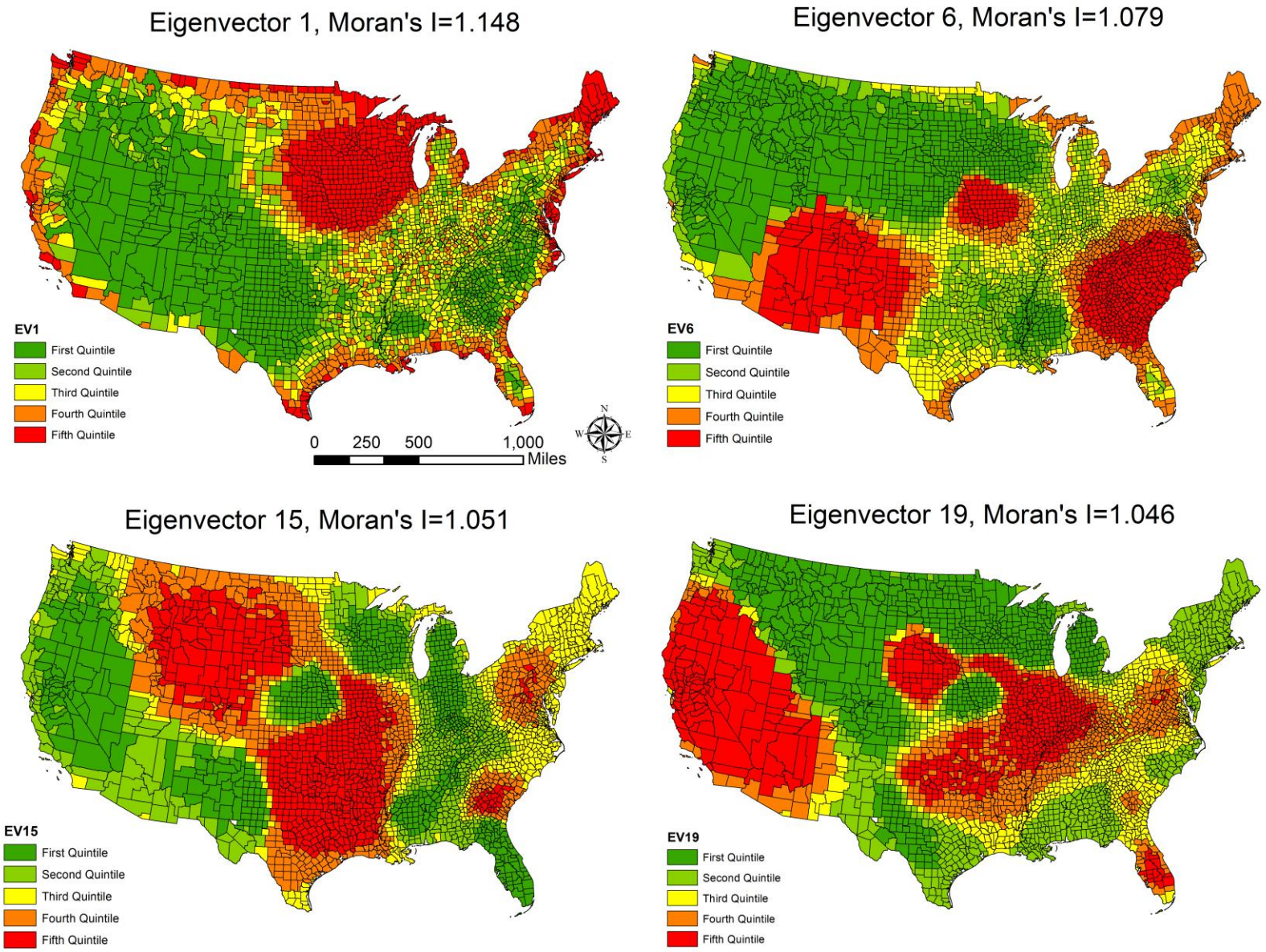
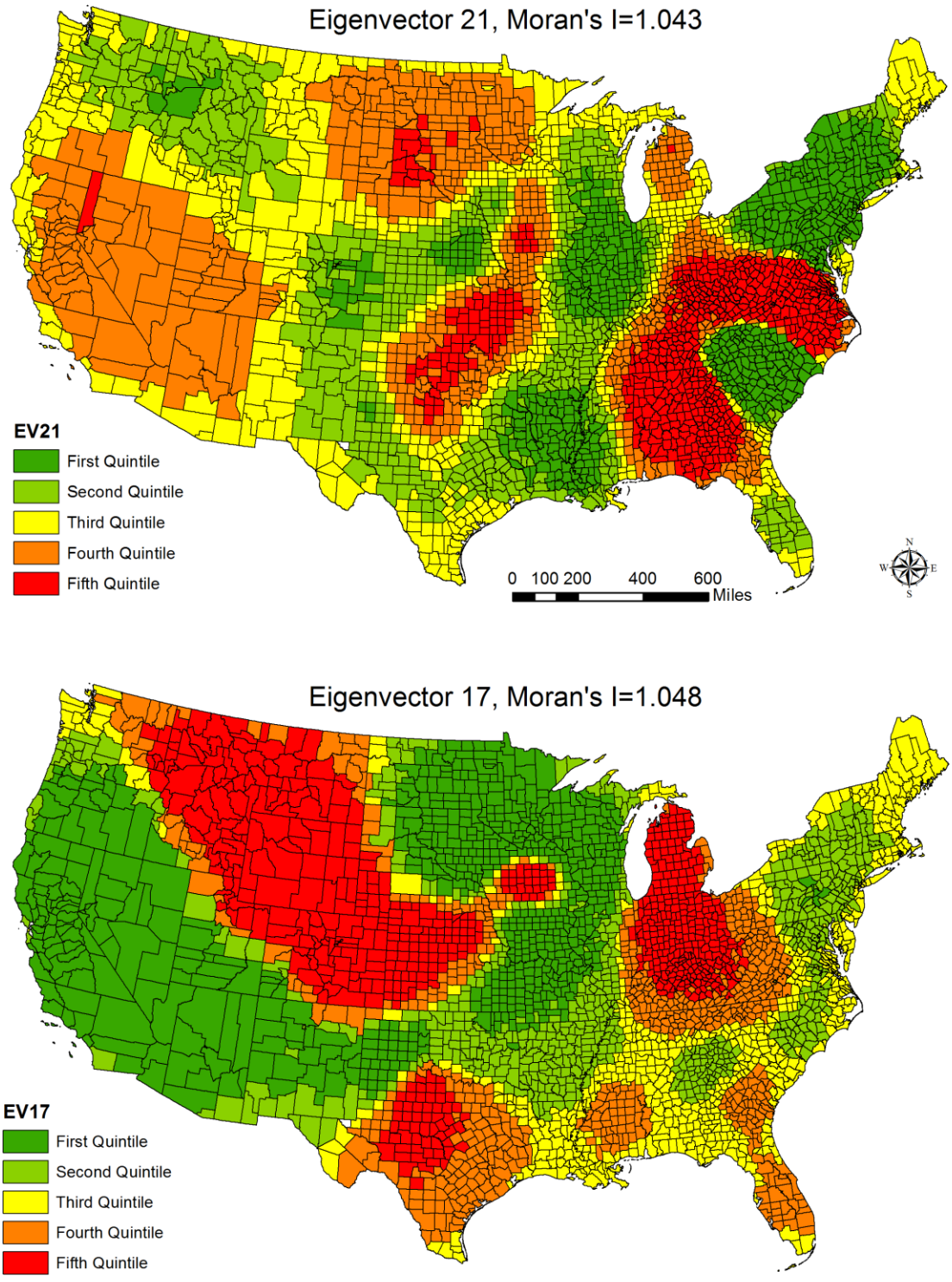


Figure 3. Spatial Patterns of the Two Additional Eigenvectors and Moran's I Values



Appendix A. Pearson's Correlation Coefficients between Mortality and Segregation Indices

	Mortality	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mortality	1															
Non-Hispanic Black																
1. Entropy (evenness)	.084**	1														
2. Isolation Index (exposure)	.354**	.628**	1													
3. Absolute Centralization (centralization)	-.063**	.338**	.150**	1												
4. Delta (concentration)	-.061**	.479**	.070**	.632**	1											
5. Spatial Proximity (clustering)	.154**	.828**	.754**	.234**	.252**	1										
Hispanic																
6. Entropy (evenness)	-.082**	.591**	.417**	.315**	.339**	.501**	1									
7. Isolation Index (exposure)	-.145**	.278**	.218**	.205**	.171**	.272**	.599**	1								
8. Absolute Centralization (centralization)	-.083**	.337**	.153**	.886**	.604**	.250**	.348**	.243**	1							
9. Delta (concentration)	-.112**	.352**	.084**	.620**	.830**	.211**	.488**	.283**	.669**	1						
10. Spatial Proximity (clustering)	-.128**	.394**	.273**	.267**	.269**	.364**	.815**	.681**	.292**	.383**	1					
Asians/Pacific Islanders																
11. Entropy (evenness)	-.008	.513**	.399**	.371**	.397**	.451**	.524**	.331**	.378**	.410**	.394**	1				
12. Isolation Index (exposure)	-.184**	.435**	.349**	.269**	.228**	.408**	.543**	.503**	.301**	.278**	.492**	.731**	1			
13. Absolute Centralization (centralization)	-.071**	.313**	.156**	.860**	.595**	.223**	.310**	.212**	.891**	.626**	.262**	.390**	.284**	1		
14. Delta (concentration)	-.030	.321**	.136**	.608**	.804**	.216**	.317**	.198**	.626**	.847**	.261**	.530**	.264**	.663**	1	
15. Spatial Proximity (clustering)	-.157**	.338**	.251**	.252**	.231**	.329**	.425**	.395**	.278**	.267**	.414**	.729**	.867**	.247**	.275**	1

**. Correlation is significant at the 0.01 level (2-tailed).