Computers, Environment and Urban Systems 35 (2011) 431–441

Contents lists available at SciVerse ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

Impact of spatial effects on income segregation indices ☆

Sergio J. Rey *, David C. Folch

GeoDa Center for Geospatial Analysis and Computation, School of Geographical Sciences and Urban Planning, Arizona State University, United States

ARTICLE INFO

Article history:
Received 15 February 2010
Received in revised form 13 July 2011
Accepted 27 July 2011
Available online 30 August 2011

Keywords:
Segregation
Spatial effects

ABSTRACT

Residential segregation is an inherently spatial phenomenon as it measures the separation of different types of people within a region. Whether measured with an explicitly spatial index, or a *classic* aspatial index, a region's underlying spatial properties could manifest themselves in the magnitude of measured segregation. In this paper we implement a Monte Carlo simulation approach to investigate the properties of four segregation indices in regions built with specific spatial properties. This approach allows us to control the experiment in ways that empirical data do not. In general we confirm the expected results for the indices under various spatial properties, but some unexpected results emerge. Both the Dissimilarity Index and Neighborhood Sorting Index are sensitive to region size, but their spatial counterparts, the Adjusted Dissimilarity Index and Generalized Neighborhood Sorting Index, are generally immune to this problem. The paper also lends weight to concerns about the downward pressure on measured segregation when multiple neighborhoods are grouped into a single census tract. Finally, we discuss concerns about the way space is incorporated into segregation indices since the expected value of the spatial indices tested is lower than their aspatial counterparts.

1. Introduction

Segregation indices have appeared in a vast literature focusing on the equity of urban areas.[1] A common theme in these studies is to examine the comparative levels of segregation either across cities at one point in time, or with respect to changes in segregation over time (e.g., Timberlake & Iceland, 2007; Stearns & Logan, 1986). An implicit assumption underlying these studies is that the values from the same index applied to different cities and/or points in time are in fact comparable. While comparisons comprise the core of many empirical segregation studies, the measures' distributional properties have received far less attention. As such, the extent to which the assumption of index comparability holds, especially in the face of cities with vastly different spatial structures, remains largely unexamined.

All residential segregation measures are in principle "spatial." Segregation indices use residents' location choices to proxy for unobserved attributes such as social networks—high segregation implies low interaction between people of different types within a region. However, the single index number is the result of many spatial decisions made by people far removed from the residents under study. In the United States, segregation studies typically use data collected by the US Census Bureau, which are disseminated via *aggregation units*, such as census tracts, to protect the privacy of individuals. The researcher must choose the *scale* of the aggregation units to investigate, e.g. blocks, block groups or census tracts, and this choice often involves an assumption that the socio-spatial construct "neighborhood" matches the official aggregation units chosen. In addition, the researcher must choose the *region size*, e.g. municipality, county or metropolitan statistical area, as the boundary to which a single index number should apply. Beyond the data themselves, segregation measures incorporate a variety of assumptions about spatial interactions. The explicitly spatial segregation measures that have emerged since the early 1980s can account for interactions between residents across Census-defined aggregation units, where earlier measures assumed interactions ended at the aggregation unit boundary. The formulation of these spatial measures assumes that regions with greater *clustering* of similar neighborhoods are more "segregated." The user of these spatial measures must choose among indices that incorporate space in different ways. All told, there are many explicit and implicit spatial decisions behind any single index number. Some of these decisions are expected to directly impact measured segregation, while others affect measured segregation more subtle ways.

The extensive literature on the measurement of segregation includes taxonomies of measures, rules defining what makes an appropriate measure and the proposal of new measures; few controlled simulation tests of the measures exist, however. Of

☆ Funding for this research was supported by National Science Foundation Award BCS-1102553.

* Corresponding author.
E-mail address: Sergio.Rey@asu.edu (S.J. Rey).

[1] For overviews see Clark (1986), Cutler, Glaeser, and Vigdor (1999), Jargowsky (1995), Jargowsky and Kim (2005), Massey and Denton (1988), Waldorf (1993), Wong (1993).

0198-9715/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2011.07.008

the limited examples, some authors explore the measures' properties using simulations based on a small number of extreme scenarios (e.g., Wong, 1997; Wong, 1993); others explore them using empirical data from a diverse set of regions (e.g., Wong, Lasus, & Falk, 1999; Krupka, 2007). While these approaches provide considerable direction, the first does not supply sufficient variation to explore the properties fully, and the second likely contains too much unexplained variation to isolate the factors affecting those properties.

This paper identifies areas of concern for researchers by exploring the impact of four spatial effects on the sampling distributions of four income segregation indices. By using a controlled simulation approach, we construct an extensive set of simulations involving 200 region designs, building a sampling distribution for each index for each design, 800 in total. We focus on two aspatial and two spatial income segregation indices: Dissimilarity Index (D), Neighborhood Sorting Index (NSI), Adjusted Dissimilarity Index (D (adj)) and Generalized Neighborhood Sorting Index (GNSI). The spatial effects considered are region size defined in terms of the number of neighborhoods, variation in region-wide interconnectivity through two spatial autocorrelation models, variation in spatial clustering of like neighborhoods via ten values of an autocorrelation parameter, and the effect of drawing aggregation units that do not match the underlying scale of the neighborhood pattern.

The remainder of the paper is organized as follows. We first provide the broader context for this work by revisiting some of the key methodological debates appearing in the segregation literature. This is followed by a description of our experimental design and the simulation procedures used to examine the impact of spatial effects on the sampling distributions of the segregation indices. The results of the experiments are examined in Section 4 and discussed in Section 5. The paper concludes with an overview of the key findings and the identification of directions for future research.

2. Motivation

The concept of residential segregation is relatively intuitive: it measures the spatial separation of groups within a region. But this is where the simplicity ends. Measuring segregation, as has been pointed out before (see Massey & Denton, 1988; James & Taeuber, 1985; Reardon, 2006), is composed of myriad theoretical and mathematical challenges. The theoretical challenges involve what it means to be "segregated." Is it a relative measure against some standard, or is it better described in absolute terms? Are highly segregated areas located near one another even more segregated (Massey & Denton, 1988)? From a mathematical perspective the challenges involve how *measured segregation* changes in the face of changes in the underlying data. Can different sized regions be compared? If one person is moved between areas, does the index reflect that shift appropriately? Should the index change if counts change but proportions remain the same (James & Taeuber, 1985)? The former challenges relate to choosing among the many candidate indices, while the latter look more into validating a particular index. This paper derives its motivation more from the latter line of literature than former as it explores how explicitly spatial changes in the underlying data affect measured segregation.

From an empirical perspective, the usefulness of segregation indices is often derived from cross-sectional or time-series comparisons. If we want to say that Region A is more segregated than Region B, or that Region C has become less segregated over the last 20 years, a critical criterion is that our measure and data do not prejudice the results; where no region has a lower (or higher) expected segregation value *a priori*. An example of this is the common finding that larger regions are more segregated than smaller ones (Krupka, 2007). If there are mathematical or data conditions predisposing larger regions to higher segregation, then this finding becomes ambiguous. It may be that larger regions are more segregated, but the methods and data available might lead to this finding regardless of some underlying truth. In their 1955 paper, Duncan and Duncan stated, "the problem of validating segregation indices is viewed as one of some importance, not only in its own right, but also as an illustration of the difficulties in finding an adequate rationale for much sociological research using index numbers" (p.210). This statement hints at the challenges faced when embarking on an empirical study involving segregation. Specifically, are the inferences made using segregation indices appropriate, valid and relevant?

Residential segregation operates at two distinct spatial levels: within neighborhoods and between neighborhoods. Traditional segregation indices ignore the latter level by treating neighborhoods as independent observations and are thus denoted "aspatial indices." Parallel to developments in exploratory spatial data analysis (ESDA), segregation researchers have recognized that the spatial configuration of the neighborhoods should be central to the empirical analysis. This reflects the recognition that traditional segregation indices are examples of "whole map" statistics in the sense that a permutation of the neighborhoods across the map would leave the aggregate segregation value unchanged.

Spatially explicit indices have been suggested throughout the social sciences for decades. However, their presence in the residential segregation literature was spotty until Massey and Denton (1988) took a broad view of the concept and brought the scattered ideas into a unified framework. In addition to identifying the existing segregation indices, they borrowed from other literatures to develop new measures.

Conceptually, spatial indices attempt to account for the empirical observation that a poor neighborhood surrounded by poor neighborhoods is more segregated than one surrounded by rich neighborhoods. Fig. 1 illustrates the concept by asking the question: is the region on the right more or less segregated than the one on the left? In both cases there 16 areas of one type and 84 of the other. Aspatial measures treat both the same; spatial ones should identify the region on the right as more segregated.

While neighborhood location is one spatial dimension impacting segregation measurement, another dimension relates to the spatial scale of the neighborhoods, and remains an important challenge facing empirical researchers since it has yet to find a satisfactory solution in either the literature on methods or in that on data. Johnston (1981) shows that holding the population size of each individual area fixed will result in higher segregation levels for more populous regions, given some mild assumptions. His argument revolves around the fixed tract size of say 4000 people, and that if all sub-groups of the population are not exactly divisible by 4000 then there is necessarily mixing of the subgroups. In smaller regions, if residents try to maximize segregation, the share of mixed to homogeneous neighborhoods would be higher than for large regions. The highly stylized explanation in Johnston (1981) is corroborated by Krupka (2007) which measures segregation for US metropolitan areas using different scales of US Census Bureau data. He shows decreasing *measured* segregation as the scale increases from census block groups to counties. Wong (1997) and Wong et al. (1999) approach the scale question from the perspective that variations in spatial scale will have less impact on measured segregation when the population is more clustered. A more clustered population is expected to have similar population patterns at various scales, thus having less impact on measured segregation no matter the scale of analysis chosen. Wong (2004) compares the scale effect for spatial and aspatial segregation measures for 30 US regions and finds that the spatial measures, like their aspatial counterparts, are sensitive to data at different scales.

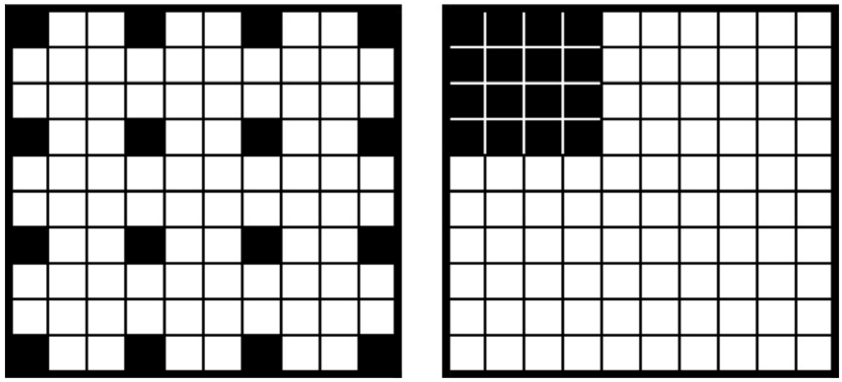

Fig. 1. Illustration of spatial segregation scenarios.

One challenge in testing these hypotheses with empirical data is that the underlying social process is unknown, as is the effect of the technical process used by data providers, such as the US Census Bureau, in the delineation of aggregation units used for data dissemination (Reynolds & Amrheln, 1998). The use of simulated regions with known properties is one avenue to overcome the ambiguity that often exists in empirical data; however, this approach is the exception rather than the rule in methodological research on segregation (Zhang, 2003). Some notable exceptions relate to comparisons of the Dissimilarity index (D) to a random distribution. Taeuber and Taeuber (1965) conducts a number of small simulations of Augusta, Georgia, in order to determine if Augusta's observed segregation level is different from a level that could have occurred from a random allocation of black and white households, and if the difference between Augusta's and Atlanta's segregation levels could have occurred from chance variation. In taking the random distribution idea even further, Cortese, Falk, and Cohen (1976) and Winship (1977) use Monte Carlo methods to compute expected values for D for various proportions of the minority population in a city (see Boisso, Hayes, Hirschberg, & Silber, 1994; Noreen, 1989; Carrington & Troske, 1997 for other imlimentations of permutation based approaches).

In more recent work, Zhang (2003) implements permutations of empirical data within various US cities to separate out patterns of racial and income segregation, using D as the test statistic. Wong et al. (1999) randomly aggregates census block data to test the effects of scale on measured segregation on 30 US regions. Although not focused on segregation, Amrhein and Reynolds (1996) uses a Monte Carlo approach to aggregate areas from a single region to different spatial scales and finds the G spatial autocorrelation statistic (Getis & Ord, 1992) is highly correlated to data aggregated to different scales. Possibly the largest use of simulation in the residential segregation literature involves agent based modeling of tipping points, exemplified in foundational works by Schelling (e.g., Schelling, 1971) and continuing today in contributions such as Fossett (2006). Although these generally look at causal factors of segregation as opposed to methodological issues of its measurement, Stearns and Logan (1986) uses this framework to identify how different causal theories of segregation influence measured segregation.

Beyond the challenges posed by space, residential *income* segregation, or the separation of rich and poor, must contend with a number of data issues. Typically, income segregation studies implement a discrete classification of people into categories such as rich-poor, low-middle-high income or an even greater number of income categories (e.g., Abramson, Tobin, & VanderGoot, 2002; Pendall & Carruthers, 2003). This is a direct translation of the *classic* race-based segregation indices to the income domain. There are drawbacks to this approach since income is a continuous variable, and even in the discrete case there is an explicit ordering of income levels unlike races (Reardon & Firebaugh, 2002). As a result, certain decisions must be made by the researcher as to the categorization of the data; these decisions are largely predetermined in the case of race-based segregation measurement. Jargowsky (1996) and Jargowsky and Kim (2005) directly tackle this issue in their segregation indices by allowing income to be treated more naturally as a continuous variable. As will be seen in the simulation section, we focus on two discrete indices and the two put forward by Jargowsky.

Whether measuring income, race or another form of residential segregation, researchers are often investigating a diverse set of urban areas. The United States urban system is one example of this diversity. Taking the June 2003 US Office of Management and Budget's Metropolitan Statistical Area (MSA) definitions, we examine 2000 census data for 359 US urban areas for internal measures of variation and found that in all cases the variation was heavily skewed. A country containing both New York City, NY and Carson City, NV is bound to generate variation in measured segregation. In addition to shear size[2] New York City was also the densest[3] and had the most income clustering.[4] As Table 1 and Fig. 2 show this is not the whole story. Another indicator, standard deviation of per capita income,[5] shows a similar trend of an extreme outlier, but in this case it is the Gulfport, MS region, which contains Biloxi and Pascagoula among other Mississippi cities (New York City ranks seventh behind the Bridgeport, CT and San Francisco, CA regions, among others).

3. Simulation procedure

This section details the simulation procedure we use to generate regions and areas and formalizes the hypotheses we test on the relationship between urban structure and segregation measurement. In the following discussion, the term "region" corresponds to "urban area" and "area" corresponds to "neighborhood." Studies

[2] Size is measured in terms of number of census tracts in 2000. This is a good indicator of size in the segregation context since most indices provide a region-wide value based on data at some subregional level.

[3] We use the median tract population density (persons per square mile) for each MSA. Since periphery tracts in many regions have large areas due to the presence of public or undeveloped lands, average MSA density is often a poor indicator of the spatial structure of a region.

[4] Income clustering is based on the standardized Moran's I value for each MSA. Moran's I is an indicator of spatial autocorrelation, or the tendency of like tracts to be located near one another.

[5] This indicator is the standard deviation of census tract level per capita income for each MSA.

Table 1
Summary statistics on indicators of US metropolitan statistical area variation, 2000.

	Number of tracts	Median pop density	Std dev per capita inc	Standardized Moran's I of per capita inc
Minimum	10	71.2	2000.9	0*
	Carson City, NV	Glen Falls, NY	Hinesville, GA	
Median	50	1220.6	6016.7	4.198
	Barnstable Town, MA	Bismarck, ND	Dalton, GA	Eugene, OR
	Champaign, IL			
	Lafayette, LA			
Maximum	4493	14093.9	60336.3	51.85
	New York, NY	New York, NY	Gulfport, MS	New York, NY

Note: City names represent the MSA central city.
* Zero represents insignificant values for Moran's I. 62 MSAs had an insignificant value indicating no spatial autocorrelation.

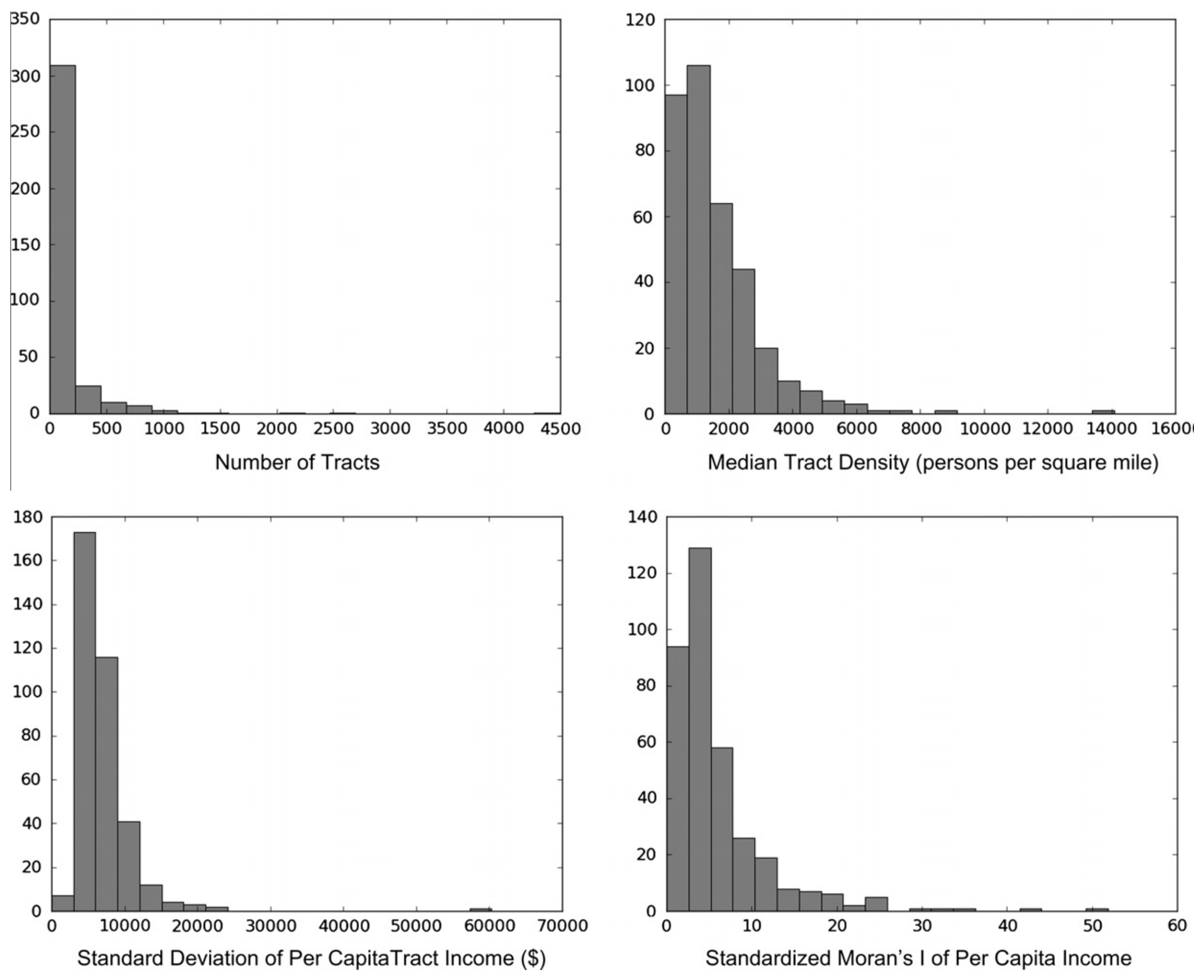

Fig. 2. US metropolitan statistical area variation, 2000.

conducted in the US typically define an urban area as a Metropolitan Statistical Area (MSA) and a neighborhood as a census tract.

3.1. Segregation indices

For this research, we calculate four segregation indices to compare measurement methods of income segregation: Dissimilarity Index (D), its spatial analogue D (adj), Neighborhood Sorting Index (NSI) and its spatial analogue Generalized Neighborhood Sorting Index (GNSI). First we define the indices mathematically and then discuss the intuition behind their selection. In all cases the indices range from 0 indicating no segregation to 1 indicating complete segregation.

We begin with D (Duncan & Duncan, 1955), which can be interpreted as the share of the poor population that would need to move to a new neighborhood to achieve an even distribution of population types (rich and poor) across all neighborhoods. Let $p_{i,g}$ denote the group g population (or count of households) in area i and $p_{i,\bar{g}} = p_i - p_{i,g}$ as the members of the population in area i that are not members of group g, where p_i is the total population of area i, and p_g is the total population of group g in the region. The dissimilarity index is given as:

$$D = \frac{1}{2}\sum_i \left| \frac{p_{i,g}}{p_g} - \frac{p_{i,\bar{g}}}{p_{\bar{g}}} \right| \quad (1)$$

$D(adj)$ (Morrill, 1991) *adjusts* D to explicitly incorporate space into the index. It is defined as:

$$D(adj) = D - \frac{\sum_i \sum_j |w_{ij}(s_{i,g} - s_{j,g})|}{\sum_i \sum_j w_{ij}}, \quad (2)$$

where $s_{i,g}$ represents the share of population g in area i (i.e. $s_{i,g} = p_{i,g}/p_i$) and w_{ij} signifies the value in the ith row and jth column of a spatial weights matrix where $w_{ij} = 1$ if i and j are neighbors and 0 otherwise, and $w_{ii} = 0$. The second term in $D(adj)$ considers how each area relates to its neighbors by reducing region-wide segregation if there is a difference in the group shares between the neighbors. The implication is that any difference between neighboring areas' shares increases overall interaction between the groups.

Individual personal income values drive NSI (Jargowsky, 1996). It can be viewed as a ratio of the standard deviation of neighborhood mean personal income to the standard deviation of individual personal incomes. NSI is defined as:

$$NSI = \sqrt{\frac{\sum_{i=1}^{N} h_i(\bar{y}_i - Y)^2}{\sum_{k=1}^{H} (y_k - Y)^2}}, \quad (3)$$

where N stands for the number of neighborhoods, i.e. areas, in the region, H denotes the number of households in the region, h_i is the number of households in area i, y_k is the income of household k, $\bar{y}_i$ indicates the mean household income for area i and Y is the mean household income for the region. If individual neighborhoods have mean values $(\bar{y}_i)$ far from the region mean, that implies a concentration of either high or low income households in those neighborhoods, i.e. segregation. This would raise the index value indicating higher segregation. Conversely, neighborhoods with residents above and below the region mean would have means near the region mean resulting in a low numerator, and hence low segregation.

GNSI (Jargowsky & Kim, 2005) *generalizes* NSI into a form that allows for the inclusion of information from the spatial neighbors of each area. To some extent, this is a smoothing technique that averages area i's average personal income $\bar{y}_i$ with that of its neighbors to arrive at a new income value for i, namely $\bar{y}_i^*$. GNSI is given as:

$$GNSI = \sqrt{\frac{\sum_{i=1}^{N} h_i(\bar{y}_{i-Y}^*)^2}{\sum_{k=1}^{H} (y_{k-Y})^2}}, \quad (4)$$

$$\text{where} \quad \bar{y}_i^* = \frac{h_i\bar{y}_i + \sum_{j=1}^{N} w_{ij}h_j\bar{y}_j}{h_i + \sum_{j=1}^{N} w_{ij}h_j}.$$

The only difference between Eqs. (3) and (4) is the replacement of $\bar{y}_i$ with $\bar{y}_i^*$ which represents the average of i with its neighbors.

The dissimilarity indices are structured entirely differently from the NSI indices, the primary difference being the root data on which they operate. D provides a measure of the binary "rich" and "poor," while NSI uses the continuous variable income as its base. NSI directly targets income segregation, while D, a more flexible measure, could equally be adapted to racial binaries such as "white" and "non-white."

We choose D because it is the most commonly encountered segregation index. D and NSI are also ideal candidates because they have direct spatial alternatives in $D(adj)$ and GNSI. Eqs. (1) and (3) do not use any spatial information and so the spatial patterning of the areas should have no bearing on measured segregation. In contrast, the other two indices explicitly include the spatial structure defined through the spatial weights matrix W. $D(adj)$ is the most straightforward among a family of similar spatial versions of D presented in Wong (1993). Wong (1993) extends Morrill (1991) by developing two variations on $D(adj)$ that consider different ways of defining neighborhood relationships. A key difference between $D(adj)$ and GNSI is the approach for incorporating spatial structure. $D(adj)$ joins the spatial and aspatial aspects of segregation in an additive form, while GNSI replaces data on an individual area with an average of its spatial neighbors.

Another reason to employ these measures is that they are all feasible using publicly available census data. This is obvious in the cases of D and $D(adj)$ which only use data at the area level. NSI and GNSI are designed for use with person level data to compute the denominator. For empirical implementation, however, Jargowsky (1996) utilized census data on the number of individuals within various income ranges to arrive at this denominator. In our study, individual people are simulated.

3.2. Region size and shape

In our simulations, each region is a square lattice made up of N areas, implying a shape of $\sqrt{N} \times \sqrt{N}$. Seven region sizes are simulated where N can take a value of 25, 81, 169, 289, 441, 576 or 625. To provide some perspective, the Phoenix-Mesa MSA contained 696 census tracts in the 2000 census. Each area within a region has the same number of residents, set at 5000. Cross-sectional studies comparing segregation between urban areas fill the empirical segregation literature. This spatial dimension is studied to help identify any systematic variation in the indices when regions of different sizes are compared.

3.3. Sampling from a spatially correlated distribution

Our simulation of spatially correlated variables follows a procedure of building the correlation structure first and then generating values that fit both this structure and a known distribution function. These two steps are as follows.

3.3.1. Correlation structure

The modeling of a spatial data generating process (DGP) requires the specification of a correlation structure. Our study examines two specifications: the spatial autoregressive (SAR) and spatial moving average (SMA), each of which considers different extents of a region's internal interconnectivity.

The SAR model can be characterized as generating a global autocorrelation structure for the region (Anselin, 2003). The estimation of any particular observation is dependent on its neighbors' values, which in turn are dependent on their neighbors' values. Therefore, this structure links the entire region together. Given the following SAR model,

$$y = \rho Wy + \varepsilon$$

the variance–covariance matrix of y is:

$$\mathrm{VC}[y] = \sigma^2[(I - \rho W)'(I - \rho W)]^{-1},$$

where y denotes an $N \times 1$ vector of observations on the variable of interest, ε is an independent and identically distributed vector of error terms, σ^2 represents the scalar variance, ρ is a spatial autocorrelation coefficient, W signifies an $N \times N$ matrix of spatial weights describing the neighbor structure of the region and I indicates the $N \times N$ identity matrix.

The second model under consideration, the spatial moving average (SMA) model, is the second most commonly seen spatial model in the empirical literature after the SAR model (Anselin, 2003). In this model, the spatial structure is a function of the error term alone, and as such the global interconnectivity seen in the SAR model is not present here. Anselin (2003) labels this a local model of autocorrelation since the correlation structure dies out after the second order neighbors. The specification of the SMA model is as follows,

$$y = \rho W\epsilon + \varepsilon$$

leading to the following spatial structure:

$$\mathrm{VC}[y] = \sigma^2[(I+\rho W)'(I+\rho W)].$$

In cross-sectional studies, researchers might mix regions that follow different spatial models. The two spatial models we explore in this study offer two representations of neighborhood interconnectivity, where one may expect that a neighborhood in a denser region is influenced by higher order neighbors.[6] In contrast, influence may dissipate quickly in a sprawling western US city where people move around by car and are able to easily leapfrog neighborhoods for shopping, services and employment. Anselin (2003) presents a thorough treatment of the two models described in this section and Rey and Dev (2006) provides an application in the context of economic convergence.

As the equations above illustrate, the term ρW is a key driver of the correlation structure, where the sign and magnitude of ρ dictate the strength and direction of the relationships. Positive values of ρ imply spatial clustering of like values, either high or low, while a negative ρ indicates spatial dispersion where high and low values are interlaced. Spatial segregation indices are designed to detect this type of spatial pattern. The idea is that a poor neighborhood surrounded by poor neighborhoods (i.e. $\rho > 0$) should be considered more segregated than one surrounded by rich neighborhoods (i.e. $\rho < 0$). In the latter scenario, the poor neighborhood is less isolated from the rich population and therefore has greater opportunity for interaction. The simulations are generated with the following sequence of values for ρ: 0, 0.1, ..., 0.8, 0.9. We do not include the negative case since it did not occur in the empirical cases tested.[7]

The spatial weights matrix defines which areas directly interact and the strength of that interaction (see Anselin, 2002 for a discussion on various types of weights matricies). All the region designs use a "queen" contiguity matrix that defines areas as neighbors if they share a border or a single point. The weights matrices in all cases are row standardized.

3.3.2. Sampling distribution

For the segregation indices chosen, it is necessary to simulate two values for each area: (1) the share of area population fitting a certain category, in this case "poor," and (2) the average income for an area. Both values require a distribution with a minimum value of zero, and in the former case, one with a distribution capped at 1.0. This precludes the normal distribution which produces values over the interval $(-\infty, \infty)$. The ideal distribution should also mimic empirical income distributions to the largest extent possible.

The lognormal distribution falls on the interval $[0,\infty)$[8] and often matches empirical income distributions for individuals due to its positive skew (Cowell, 1995).[9] To test if this holds for the distribution of neighborhoods, we plotted tract level per capita income for 359 MSAs using data from the 2000 US Census. In most cases, the lognormal distribution fits the data reasonably well. For the simulations, we use the average of the lognormal parameters from the 359 MSAs, $\mu_{ln} = 9.816$ and $\sigma_{ln} = 0.330$.

Incorporating the spatial aspects of this simulation requires the construction of a multivariate lognormal sampling distribution. The process for simulating correlated values from a lognormal distribution is as follows:[10]

1. Generate an $N \times N$ variance–covariance matrix, *VC*, via one of the spatial models defined above, with $\sigma^2 = 1$, ρ between 0 and 0.9 and a queen contiguity weights matrix.
2. Transform *VC* to a correlation matrix, Σ, containing values $\hat{\rho}_{ij}$.[11]
3. Since we essentially use a multivariate normal random number generator to create multivariate lognormal random values, the correlation matrix must be transformed in advance to account for the eventual exponentiation that will be performed. This is done by transforming each value in the correlation matrix Σ via $\hat{\rho}^{adj}_{ij} = \log[\hat{\rho}_{ij}(\exp\sigma^2_{ln} - 1) + 1]$ to create Σ^{adj}.
4. Compute the Cholesky Decomposition C of Σ^{adj}, which when multiplied by an uncorrelated vector (e.g. Z in the next step) results in a correlated vector matching the correlation structure of Σ^{adj}.
5. Generate a vector of uncorrelated random values Z, where $Z \sim N(0,1)$ and is of length N.
6. Transform Z into Z^{adj} via C so that $Z^{adj} \sim N(0, \Sigma^{adj})$.
7. Shift each simulated value to account for the overall mean by adding μ_{ln} to each value in Z^{adj} to get Z^{adj}_{μ}.
8. Exponentiate Z^{adj}_{μ} to arrive at correlated lognormal values.

Tests of the above procedure showed that on average the *a priori* spatial correlation and lognormal distribution were maintained in the simulated values. It should be noted that a small decline in σ_{ln} was observed as ρ increased resulting in a narrowing of the lognormal distribution.

As a result of sampling from $y \sim LN(9.816, 0.330)$, the expected value of y is $19,350. For the cases of NSI and GNSI, y is the per capita income of the area. The following section describes how the share of low income persons is determined for each area.

3.4. Sampling within Areas

The NSI and GNSI indices require a distribution of personal income within the individual areas, and the *D* and *D*(*adj*) indices require the share of the population defined as "poor." Following the literature cited above, we simulate personal income via a univariate lognormal distribution, avoiding the spatial correlation structure needed for the area level simulations. Given a known mean (E[y]) it is possible to build a specific lognormal distribution for each simulated area according to the following rule (modified from Greene, 2003):

$$\mu_{ln} = \ln(\mathrm{E}[y]) - \frac{\sigma^2_{ln}}{2}. \tag{5}$$

A four step process is used to implement within area sampling. The first step involves the generation of a spatially correlated simulation of per capita income for each area, as described in the previous section. Each area income, E[y], is then plugged into Eq. (5) along with $\sigma_{ln} = 0.330$ to define each area's reference distribution. Next, 5000 personal incomes are drawn from the distribution for each area. Finally, the area per capita income is replaced by the average of the 5000 simulated personal incomes to ensure that the entire system matches exactly. In total, these four steps generate an entire region

[6] Neighbor order refers to the neighbors of neighbors, where the higher the order the greater the number of areas included in the computation.

[7] Based on 2000 Census Bureau data, the Moran's I value for per capita income in 359 MSAs was always greater or equal to 0.

[8] It is acknowledged that in empirical data household incomes can go negative, but for this simulation exercise it was deemed reasonable to assume $y \geqslant 0, \forall y$.

[9] Cowell (1995) finds that incomes typically follow the lognormal or Pareto distributions, with the Pareto occurring more at the higher income levels.

[10] See Lienhard (2004) for further description of this method along with code for implementation in MatLab.

[11] $\hat{\rho}$ is used here simply to differentiate it from the autocorrelation coefficient ρ.

of personal incomes while maintaining the spatial correlation structure between areas (i.e. neighborhoods).

With this full set of neighborhoods and residents in hand, we can determine the ratio of "poor" residents. This is the ratio of residents with an income below $11,600 (i.e. approximately 60% of the expected personal income) to the total number of residents, which is always set to 5000.

3.5. Scaling

One motivation for varying spatial scale is to test the idea that the aggregation areas (e.g. census tracts) may not be representative of neighborhood boundaries within the region. Krupka (2007) provides empirical evidence of this by testing data for US regions at different US Census data scales (block groups, census tracts and counties). He hypothesizes that larger urban areas have larger neighborhoods and thus one complete neighborhood can fit into an aggregation boundary such as a tract, but with smaller neighborhoods in smaller urban areas, mixture occurs at the aggregation boundary level.

Cowgill and Cowgill (1951) raises the mismatch issue also, and advocates using smaller scale aggregation areas instead of tracts when measuring segregation. However, this approach has a known end game. If a region is continuously subdivided into smaller and smaller areas, in the limit, each area contains only one person. Since that person can only be rich or poor, the index must have a value reflecting complete segregation. Therefore, upscaling is expected to reduce segregation and downscaling to increase it, regardless of some true underlying level of segregation.

Simulations incorporating scale operate only on a 24×24 region. To test the *upscaling* hypothesis, we simulate regions with 576 areas and one of the spatial correlation structures described above. This initial run is intended to model the actual spatial patterning of neighborhoods. These areas then scale up by factors of 4, 9 and 16 producing regions of 144 (12×12), 64 (8×8) and 36 (6×6) areas. The per capita income of an upscaled area results from averaging the per capita incomes from its constituent lower level areas. These upscaled areas mimic the actual aggregation boundaries that might group mismatched neighborhoods together. All other simulations run for this study could be viewed as having a scale factor of 1, i.e. no scaling of the data.

3.6. Iterations

Combining all the pieces results in 200 region designs. There are six region sizes, two spatial correlation models and ten values for ρ (i.e. strength of correlation) resulting in 120 regions at a constant spatial scale. The scale cases results in an additional 80 regions for the four scaling factors. For each region design we run 500 simulations, compute the four segregation indices on each simulation, and find the mean and standard deviation of each index. The results of these 800 mean and standard deviation values are presented in the following section.

4. Simulation results

The simulation stage explores the response of four segregation indices to variation in four spatial dimensions. We present the simulation results in two forms. The first is a pair of figures, each encapsulating different dimensions and indices. Both figures present variation in autocorrelation (ρ) and the data generating model (SAR and SMA). We split the remaining two spatial dimensions across the two figures: Fig. 3 presents variation in region size, while Fig. 4 presents variation in scale. These figures provide a high level visual representation of the results.

The second presentation is an investigation of the sensitivity of measured segregation to the four spatial dimensions. This is done through a series of regressions with mean measured segregation as the dependent variable and the parameters used to build the regions as the independent variables (see Table 2). A significant coefficient would indicate that the associated spatial dimension affects measured segregation for the cases studied here. A total of eight regressions were run that split the results by index type and by tests on scale and size. All the regressions correspond directly to the two figures introduced above. Specifically, each pixel in the figure (mean values) is an observation in a regression; for example, there are 120 pixels in the left column of the upper-left quadrant of Fig. 3 corresponding to the 120 observations used for the regression in column 1 of Table 2. The remainder of this section summarizes key findings for each of the spatial dimensions investigated.

4.1. Variation in region size

Segregation indices, be they aspatial or spatial, will ideally be immune to variation in region size. Overall, we find this to not be the case as the coefficient on region size is positive and significant in all cases (Table 2). This implies that as region size increases measured segregation will tend to increase, a confirmation of Johnston (1981) using a more general approach than his single illustrative example. However, Fig. 3 does not mirror this bleak interpretation as two cases standout as likely drivers of the positive correlation. First, except for the smallest region size, 25 areas, measured segregation does not tend to vary much with region size. This is indicated by a generally homogeneous color in Fig. 3, e.g. the Dissimilarity Index for the SMA model or a lack of horizontal banding, e.g. GNSI for the SAR model. Second, the increase in segregation with increases in region size are only noticeable for very high values of ρ. To test these hypotheses, we reran the regressions for regions with more than 25 areas and with $\rho < 0.9$ and found the p-value on region size remained highly significant for D and NSI, but became insignificant for $D(adj)$ and GNSI (0.161 for $D(adj)$ and 0.905 for GNSI). Overall, caution should be employed when making any comparisons of regions of different sizes using D or NSI. However, $D(adj)$ and GNSI appear to correct for this problem in the typical scenarios considered in empirical work, i.e. when regions are not extremely small and simultaneously not exhibiting extremely high levels of spatial autocorrelation.

4.2. Variation in spatial scale

The feature we model in the tests on spatial scale is one of the most challenging for the empirical researcher. Recall that this investigates the mismatch between the actual neighborhood structure and the one revealed to the researcher via public data sources; where the implication is that actual neighborhoods are smaller than aggregation areas. Table 2 shows that scale has a highly significant and negative effect on measured segregation. Therefore, as spatial scale increases, i.e. as more "actual" neighborhoods merge into one aggregation unit, segregation decreases. All eight mean graphs in Fig. 4 show horizontal banding indicating systematic variation in segregation values as scale increases. Each column of a particular graph is based the same simulated data from a 576 area region. The difference from one row to the next is the level of aggregation of the simulated data, starting from a baseline of 576 areas (i.e. the 1:1 case) on the bottom row to the 16:1 case on top.

These results are what would be expected since blended aggregation areas will tend to smooth out segregated neighborhoods. Unlike the size case, the spatial measures, $D(adj)$ and GNSI, are not able to correct for the scale problem. While the "expected" results emerge, this does not imply that the concept of segregation has been adequately captured for the regions. The large decline

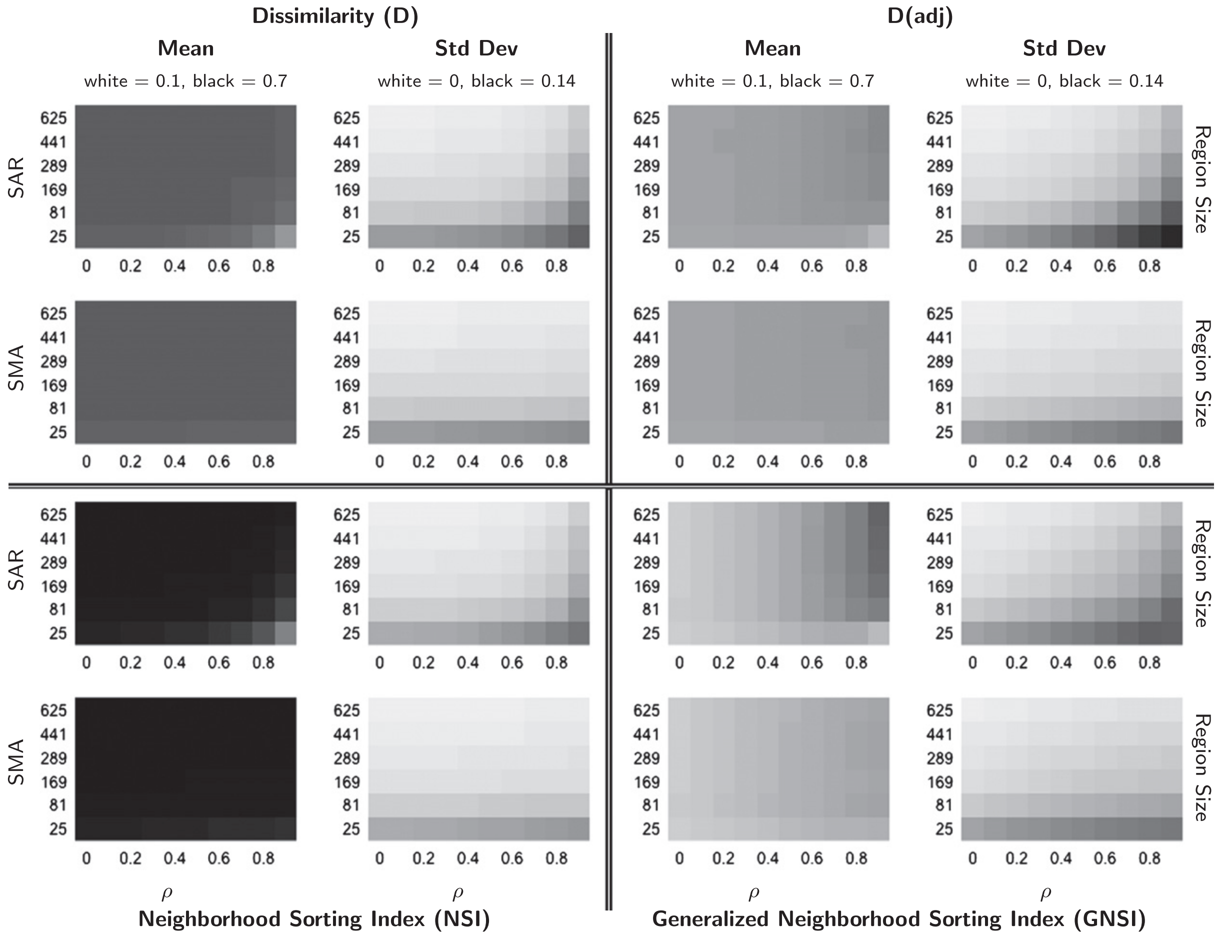

Fig. 3. Simulation results for the region size case. *Notes:* This figure is divided into quadrants, each containing four graphs. The left quadrants present the mean and standard deviation for the aspatial indices (D and NSI) with the right quadrants presenting the spatial indices (D (adj) and GNSI). The top pair of graphs in each quadrant present the results for the SAR model, with the lower pair presenting the SMA model. For each graph, rows represent the region size and columns the value of ρ, with color intensity being the magnitude of the mean or standard deviation value from 500 simulations.

from the 1:1 to 4:1 case (see Fig. 4) shows that segregation could be grossly under-measured even in the smallest aggregation scenarios. This issue is compounded in cross-region comparisons if one region's aggregation areas are at the 1:1 scale while another's are at 4:1.

4.3. Variation in the spatial DGP

The spatial DGP explores the difference between a broadly interconnected region, the SAR model, and one where the spatial relationships are more localized, the SMA model. A more interconnected region is expected to have smoother transitions from high to low income areas, where the alternative is a region with a more discontinuous surface of neighborhood income values. The regression results show that for six of the eight models (columns 2, 3, 4, 6, 7 and 8), a region built using the SAR model has a positive and significant impact on measured segregation over the alternative SMA model. Since neighboring areas are expected to be more similar in a SAR region than a SMA region, when actual neighborhoods are joined together into aggregation areas, as in the scaling case, or areas are smoothed together via a moving window as in the GNSI, it is more likely for similar populations to be joined for the SAR model.

The two regressions that do not involve aggregation or smoothing, columns 1 and 5, show a significant and negative impact of the SAR model over the SMA model. We attribute this result, which is low in magnitude, to the potential dampening of the variance caused by the data generating process, i.e. the conversion of random values to spatially correlated values. Lower overall variance will tend to reduce any measure of segregation since neighborhood values are more closely packed together. This implies then that the magnitudes on the coefficients for the SAR dummy variables in the other regressions are lower bounds of what might be expected in empirical data.

In all cases, the results from the regressions can be observed in the two results figures. While the differences are generally easy to see for small region sizes and high values of ρ, tests of this hypotheses did not result in changes to the significance or direction of the results presented in Table 2 meaning that the findings are not isolated to the extreme cases.

4.4. Variation in the magnitude of spatial autocorrelation

Regions constructed using higher levels of spatial autocorrelation have stronger clusters of high and low values, while lower spatial autocorrelation levels tend toward a random distribution of income values across the spatial landscape. This variation is implemented through values of ρ ranging from 0.0 to 0.9. In principle, the magnitude of ρ should have no impact on aspatial indices of segregation since the neighborhood relationships are not involved in the computation. In contrast, spatial autocorrelation is the spatial dimension that most spatial segregation measures, including

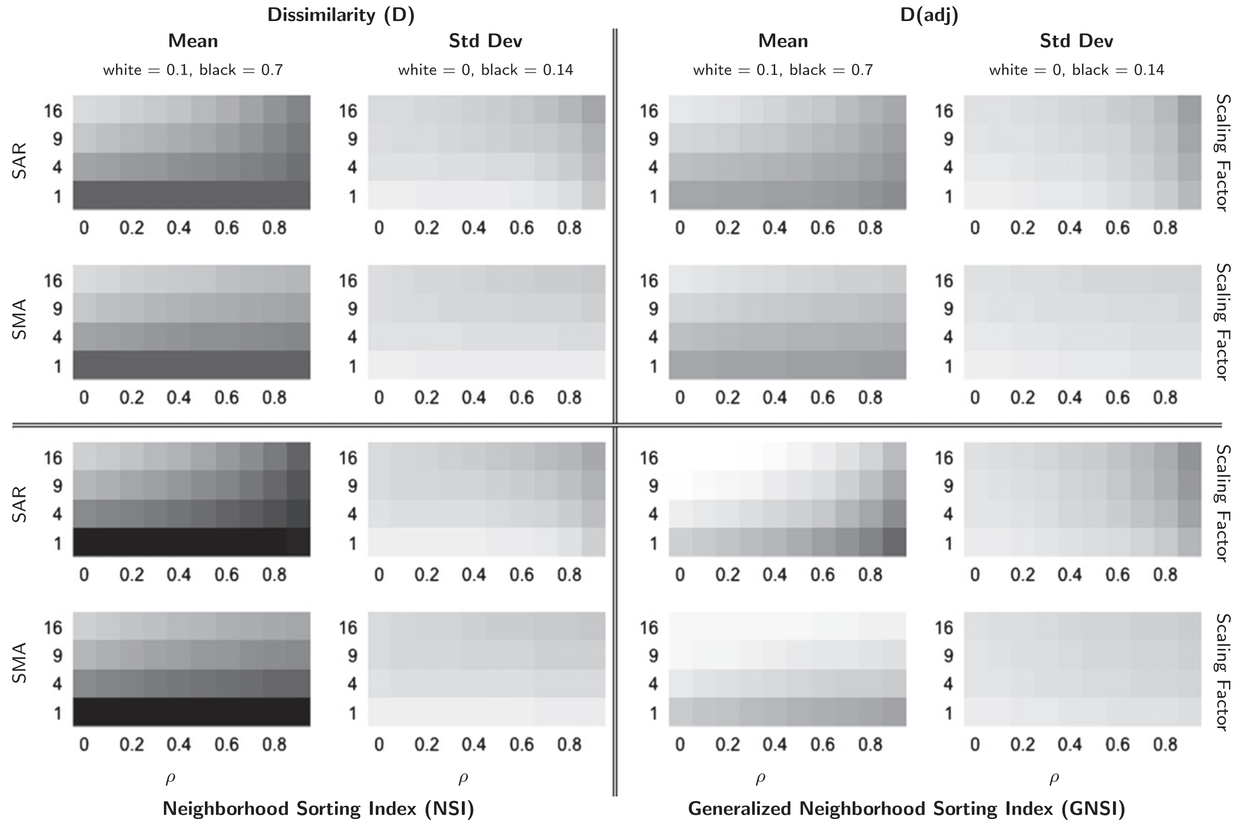

Fig. 4. Simulation results for the scale case. *Notes:* This figure is divided into quadrants, each containing four graphs. The left quadrants present the mean and standard deviation for the aspatial indices (*D* and NSI) with the right quadrants presenting the spatial indices (*D* (*adj*) and GNSI). The top pair of graphs in each quadrant present the results for the SAR model, with the lower pair presenting the SMA model. For each graph, rows represent the scaling factor and columns the value of ρ, with color intensity being the magnitude of the mean or standard deviation value from 500 simulations.

Table 2
Impact of spatial parameters on measured segregation.

	(1) *D* size	(2) *D* (*adj*) size	(3) *D* scale	(4) *D* (*adj*) scale	(5) NSI size	(6) GNSI size	(7) NSI scale	(8) GNSI scale
Constant	0.5463***	0.3378***	0.4540***	0.3108***	0.6772***	0.1982***	0.5611***	0.1970***
	(145.232)	(118.519)	(34.112)	(51.768)	(130.729)	(28.409)	(32.091)	(16.679)
ρ	−0.0206***	0.0413***	0.1211***	0.1019***	−0.0304***	0.1867***	0.1530***	0.1729***
	(−3.909)	(10.308)	(6.506)	(12.135)	(−4.178)	(19.062)	(6.257)	(10.468)
Size	0.0033***	0.0024***			0.0057***	0.0041***		
	(4.512)	(4.436)			(5.703)	(3.039)		
Scale			−0.0167***	−0.0095***			−0.0219***	−0.0117***
			(−17.714)	(−22.484)			(−17.728)	(−13.969)
SAR	−0.0075**	0.0069***	0.0255**	0.0224***	−0.0107**	0.0382***	0.0321**	0.0466***
	(−2.458)	(3.009)	(2.386)	(4.640)	(−2.556)	(6.782)	(2.282)	(4.912)
R squared	0.2644	0.5378	0.8264	0.8987	0.3276	0.7830	0.8251	0.8123
Observations	120	120	80	80	120	120	80	80

t-stats in parenthesis.
"sar" is a dummy variable, where 1 indicates the SAR model and 0 the SMA model.
"size" is region size (number of areas) divided by 100.
* Significant at 0.10 level.
** Significant at 0.05 level.
*** Significant at 0.01 level.

the two explored in this paper, specifically target higher clustering indicates higher segregation. To some extent these ideal results are borne out in our simulations. The four spatial segregation cases highlighted in Table 2, columns 2, 4, 6 and 8, all show a significant and positive impact on measured segregation for increases in ρ. Similarly, all but the 25 area regions on the right side of the two results figures show vertical banding in the mean segregation values indicating increasing values as ρ increases. The low coefficient on ρ for $D(adj)$ in the size case (column 2) is manifest in Fig. 3 where the horizontal banding is the least noticeable.

In contrast to the results for the spatial measures, we find that the aspatial measures have negative and significant coefficients for the models where regions vary in size, columns 1 and 5. As was stated above, we find a slight reduction in the variance of our simulated data after the conversion to spatially correlated values, and we suspect this is causing the small negative coefficients for the these two models. The remaining cases, columns 3 and 7, show significant and positive coefficients as ρ increases. The positive coefficients on ρ are caused by neighborhoods with similar income values being more clustered as spatial autocorrelation increases; as a result the smoothing and aggregating processes tend to join neighborhoods with similar values as ρ increases. The spatial measures' ability to pick up this pattern is encouraging; but the aspatial measures' sensitivity to ρ in the scale case is discouraging.

5. Discussion

As the previous section shows, we find that the four selected segregation measures are affected by variation in spatial configuration of the data meaning that care should be taken when segregation values are compared. We construct regions that are devoid of the myriad social dynamics that can vary from urban area to urban area. The regions are also isolated from the historic, technical and bureaucratic processes used to construct aggregation areas such as census tracts. Hence this provides a relatively clean testbed for exploring the variability of segregation measures when applied to all types of regions.

5.1. Summary of results

On the question of region size, we arrive at mixed results. The arguments in the literature that the positive correlation between region size and measured segregation is, at least partially, due to the construction of aggregation areas (Johnston, 1981; Krupka, 2007) are still not fully resolved here. Solid proof of this hypothesis would have come in the form of insignificant parameters on region size since our neighborhoods are random constructions. However, we found small but significant coefficients for the indices. On a positive note, we do find that the spatial measures we explore in this paper essentially remove this size effect for all but the smallest and most extremely clustered regions.

Variation in scale is shown to be a significant problem faced by empirical researchers. Wong (1997) provides a theoretical argument that greater autocorrelation will reduce the impact of scale variation. We were able to test that argument using our simulations and found it to hold in nearly all cases.[12] However, while the effect is reduced, even regions with extreme spatial autocorrelation, i.e. $\rho = 0.9$, can still be susceptible to the scale problem. A further confounding factor is the opposing effects from spatial scale and spatial autocorrelation. Fig. 4, which focuses on scale variation, clearly shows that for any particular value of ρ, measured segregation declines as scaling increases. However, it also shows that segregation increases as ρ increases. We have a situation where segregation is highest in the lower right corner of any particular mean graph in this figure, and decreasing both as scaling increases and as spatial autocorrelation decreases. These independent effects from scale and spatial autocorrelation would be difficult to disentangle in empirical data.

The finding that greater interconnectivity leads to significantly greater measured segregation for the spatial measures is understandable from a spatial analysis perspective, but may be counterintuitive from a social perspective. Smooth income transitions from neighborhood to neighborhood could be an indicator of greater interaction in the population of different types; while the sharper divisions resulting from the SMA model might imply a population with greater social divisions.

It is important to point out that the spatial measures do not uncover the differential impact space has on segregation relative to their aspatial counterparts. Both spatial measures capture the effect clustering of like neighborhoods has on measured segregation—both measures increase as ρ increases. However, as can be seen in the results figures and in the construction of the measures themselves, the expected value of the spatial measure is always less than its aspatial counterpart. Revisiting $D(adj)$ from Eq. (2) we can see that spatial segregation must always be lower than D except in the case of $D = 0$, i.e. $D = D(adj) = 0$.[13] The second term in Eq. (2) is only sensitive to absolute differences in neighboring pairs of areas, not whether the neighbor is more or less segregated. In principle, being surrounded by less segregated neighborhoods should contribute to reduced segregation, and the reverse if the neighborhoods are more segregated. As can be seen in Eq. (4), the key difference between NSI and GNSI is the introduction of spatially smoothed data. Since GNSI smooths neighboring values, the term $\bar{y}_i^*$ tends to be closer to the regional mean than $\bar{y}_i$ resulting GNSI being less than NSI on average. Unlike $D(adj)$, it is possible for GNSI to be greater than its aspatial counterpart, and it is in some simulations; however, it is not the expected result over many trials as shown in the simulation figures. Overall, comparison of the *magnitude* of measured segregation reported by an aspatial index and its spatial counterpart should be done with consideration of these general properties.

6. Conclusion

This paper presented the performance of four segregation indices, Dissimilarity Index (D), Adjusted Dissimilarity Index ($D(adj)$), Neighborhood Sorting Index (NSI) and Generalized Neighborhood Sorting Index (GNSI), tested against varying degrees of spatial effects. The aspatial indices showed mild increases as region size increased, but this problem was corrected in most cases by their spatial counterparts. They all failed to accommodate for the situation where neighborhoods are grouped into larger aggregation areas. In fact, the biggest decline in measured segregation occurred in the shift from a 1:1 scaling factor to 4:1, results that indicate significant challenges for many empirical studies. Regional integration, as measured by the spatial data generating process, had a positive impact on spatial segregation indicating that more interconnected regions may return an artificially high level of segregation.

Variation in spatial autocorrelation (ρ) was identified by the spatial segregation measures, but the expected values of these measures are lower than their aspatial counterparts. We would expect that bringing space into the formulation should allow segregation to either rise or decline, but this is not the case.

The current work can be expanded in a number of directions. First, additional measures of income segregation can be incorporated into the existing framework. Entropy measures (i.e. those based on the Theil Index), Gini index, coefficient of variation and many others offer possibilities. Second, multi-group measures of segregation would lend themselves well to the study of income segregation as they allow for multiple income ranges, and deserve further exploration. Third, this work does not address variation of the *overall* population. Specifically, all regions have the same expected value for average income or share of poor residents (even though the areas within could vary greatly). Since actual regions vary along this dimension an exploration of how these indices perform at different locations along the income or poverty distribution may provide interesting findings.

[12] We measured change in mean segregation between each one step scale difference, i.e. 1:1 vs. 1:4, 1:4 vs. 1:9 and 1:9 vs. 1:16. This resulted in 240 percent change measurements. We then identified any cases where percent change increased between values of ρ, again measured as one step intervals, i.e. 0 vs. 0.1, 0.1 vs. 0.2, … The only increases occurred for mean GNSI for the SMA model with high values of ρ.

[13] Even this does not hold in the case of neighborhoods with zero population. Zero population neighborhoods do not contribute to D, but they do contribute to $D(adj)$. While $D = 0$ is unlikely to occur in any empirical dataset, census tracts with zero population do occur. In addition, negative values of $D(adj)$ are possible (Wong, 2005) in some cases of negative spatial autocorrelation.

References

Abramson, A., Tobin, M., & VanderGoot, M. (2002). The changing geography of metropolitan opportunity: The segregation of the poor in US metropolitan areas, 1970–1990. *Housing Policy Debate, 6*(1), 45–72.

Amrhein, C. G., & Reynolds, H. (1996). Using spatial statistics to assess aggregation effects. *Geographical Systems, 3*, 143–158.

Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics, 27*, 247–267.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review, 26*(2), 153.

Boisso, D., Hayes, K., Hirschberg, J., & Silber, J. (1994). Occupational segregation in the multidimensional case: Decomposition and tests of significance. *Journal of Econometrics, 61*(1), 161–171.

Carrington, W. J., & Troske, K. R. (1997). On measuring segregation in samples with small units. *Journal of Business & Economic Statistics, 15*(4), 402–409.

Clark, W. A. V. (1986). Residential segregation in American cities: A review and interpretation. *Population Research and Policy Review, 5*(2), 95–127.

Cortese, C., Falk, R., & Cohen, J. (1976). Further considerations on the methodological analysis of segregation indices. *American Sociological Review, 41*(4), 630–637.

Cowell, F. A. (1995). *Measuring inequality* (2nd ed.). London: Prentice Hall/Harvester Wheatsheaf.

Cowgill, D., & Cowgill, M. (1951). An index of segregation based on block statistics. *American Sociological Review, 16*(6), 825–831.

Cutler, D., Glaeser, E., & Vigdor, J. (1999). The rise and decline of the american ghetto. *Journal of Political Economy, 107*(3), 455–506.

Duncan, O., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review, 20*(2), 210–217.

Fossett, M. (2006). Ethnic preferences, social distance dynamics, and residential segregation: Theoretical explorations using simulation analysis. *Journal of Mathematical Sociology, 30*, 185–274.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*, 189–206.

Greene, W. (2003). *Econometric analysis*. Prentice Hall Upper Saddle River.

James, D., & Taeuber, K. (1985). Measures of segregation. *Sociological Methodology, 15*(1), 1–32.

Jargowsky, P. (1995). Take the money and run: Economic segregation in US metropolitan areas. Institute for Research on Poverty, University of Wisconsin – Madison.

Jargowsky, P., & Kim, J. (2005). A measure of spatial segregation: The generalized neighborhood sorting index. Ann Arbor, MI: University of Michigan, National Poverty Center, Working Paper Series.

Jargowsky, P. A. (1996). Take the money and run: Economic segregation in US metropolitan areas. *American Sociological Review, 61*(6), 984–998.

Johnston, R. J. (1981). Some technical issues in the comparative study of residential segregation. *Transactions of the Institute of British Geographers*, 246–250.

Krupka, D. (2007). Are big cities more segregated? Neighbourhood scale and the measurement of segregation. *Urban Studies, 44*(1), 187.

Lienhard, S. (2004). Multivariate lognormal simulation with correlation. Last checked: September 14, 2010. <http://www.mathworks.co.uk/matlabcentral/fileexchange/6426>.

Massey, D., & Denton, N. (1988). The dimensions of residential segregation. *Social Forces, 67*, 281.

Morrill, R. (1991). On the measure of spatial segregation. *Geography Research Forum, 11*, 25–36.

Noreen, E. W. (1989). *Computer-intensive methods for testing hypotheses: An introduction*. New York: Wiley.

Pendall, R., & Carruthers, J. (2003). Does density exacerbate income segregation? Evidence from US metropolitan areas, 1980–2000. *Housing Policy Debate, 14*(4), 541–590.

Reardon, S. F. (2006). A conceptual framework for measuring segregation and its association with population outcomes. In J. M. Oakes & J. S. Kaufman (Eds.), *Methods in social epidemiology* (1st ed.. San Francisco, CA: Jossey-Bass.

Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. In R. M. Stolzenberg (Ed.), *Sociological Methodology* (pp. 33–67). Boston: Blackwell Publishing.

Rey, S. J., & Dev, B. (2006). σ-convergence in the presence of spatial effects. *Papers in Regional Science, 85*(2), 217–234.

Reynolds, H., & Amrheln, C. G. (1998). Some effects of spatial aggregation on multivariate regression parameters. In D. A. Griffith, C. G. Amrhein, J. M. Huriot, L. J. Gibson, A. Bailly, & N. L. Law (Eds.), *Econometric advances in spatial modelling and methodology: essays in honour of Jean Paelinck*. Kluwer Academic Publishers, pp. 85–106.

Schelling, T. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*, 143–186.

Stearns, L., & Logan, J. (1986). Measuring trends in segregation: Three dimensions, three measures. *Urban Affairs Quarterly, 22*(1), 124–150.

Taeuber, K. E., & Taeuber, A. F. (1965). *Negroes in cities: Residential segregation and neighborhood change*. *Population research and training center monographs*. Chicago: Aldine Pub. Co.

Timberlake, J., & Iceland, J. (2007). Change in racial and ethnic residential inequality in American cities, 1970–2000. *City & Community, 6*(4), 335–365.

Waldorf, B. (1993). Segregation in urban space: a new measurement approach. *Urban Studies, 30*, 1151–1164.

Winship, C. (1977). A revaluation of indexes of residential segregation. *Social Forces, 55*(4), 1058–1066.

Wong, D. S. (1993). Spatial indices of segregation. *Urban Studies, 30*, 55–572.

Wong, D. S. (1997). Spatial dependency of segregation indicies. *Canadian Geographer, 41*, 128–136.

Wong, D. S. (2004). Comparing traditional and spatial segregation measures: A spatial scale perspective. *Urban Geography, 25*(1), 66–82.

Wong, D. S. (2005). Formulating a general spatial segregation measure. *The Professional Geographer, 57*(2), 285–294.

Wong, D. S., Lasus, H., & Falk, R. F. (1999). Exploring the variability of segregation index D with scale and zonal systems: An analysis of thirty US cities. *Environment and Planning A, 31*, 507–522.

Zhang, J. (2003). Revisiting residential segregation by income: A Monte Carlo test. *International Journal of Business and Economics, 2*(1), 27–37.