doi:10.1111/pirs.12145

The centralization index: A measure of local spatial segregation*

David C. Folch[1], **Sergio J. Rey**[2]

[1] Department of Geography, Florida State University, PO Box 3062190, Tallahassee, FL 32306, USA (e-mail: dfolch@fsu.edu)
[2] School of Geographical Sciences and Urban Planning, Arizona State University, PO Box 875302, Tempe, AZ 85287, USA (e-mail: sergio.rey@asu.edu)

Received: 1 December 2013 / Accepted: 15 August 2014

Abstract. Segregated areas may occur around an attractive park or a waste incinerator, but the magnitude and group membership of the people in closest proximity will likely be different. We therefore introduce a local segregation measure that can be applied to any location within a metropolitan area, and that can identify the group that is relatively more concentrated around that reference location. We further introduce an inference approach to identify the statistical significance of a particular segregation value. In an exploratory setting the index can be used to generate a map of hot spots, and seed the question: "why is this group significantly concentrated around that location?"

JEL classification: C43, D63, J15, R23

Key words: Local segregation, spatial segregation, centralization

1 Introduction

US House of Representatives Speaker Tip O'Neill's aphorism, 'all politics is local', recognizes that voters care about what directly affects them and their local community; politicians that who ignore this concept do so at their own peril. A similar statement could be made that all residential segregation is local. Segregated outcomes are based on the neighbourhood level choices residents make. By implication, local analysis of segregation is expected to provide valuable information on the pattern of segregation, and hence open the door to policy solutions not visible at the global scale.

For decades, segregation literature focused on identifying the segregation level for an entire region. In the early 1980s, spatially explicit segregation measures began entering the discussion in earnest (e.g., White 1983), which were followed in the mid-1990s by the introduction of local measures (e.g., Wong 1996). In this paper we weave these two advances by repurposing the first spatial segregation measure, the centralization index (Duncan and Duncan 1955b), as a local measure.

* The authors would like to thank the editor and anonymous referees for their valuable feedback. Funding for this research was supported by National Science Foundation Grant BCS-1102553.

The centralization index was originally designed to quantify minority segregation relative to an urban area's centre. However, the index has fallen out of favour as urban areas become more polycentric. The implication is that minority concentration around a region's single urban core is an inherently bad outcome; this arguably made sense in the 1950s when the measure was developed. However, variation in downtowns from city to city, and the reemergence of downtowns as fashionable residential locations upends the idea that being near a region's centre is an inherent disamenity.

The original centralization index is computed using cumulative sums of the two population groups at ever increasing distances from the urban core. This is like shining a spotlight on a location, where the focus of the light is most intense at the centre, and then difiuses near the edge. While this is an effective approach to measuring segregation around a location, the resulting value is sensitive to the location chosen as the centre. The value is also sensitive to which population group is more concentrated around the focal location: positive values indicate one population group, while negative values point to the other group. We leverage these strengths and weakness of the global measure in the development of a local measure.

As an exploratory tool, the index can iteratively focus on every neighbourhood within the urban area to identify each neighbourhood's segregation level. Hot spots, or those areas with a value more extreme than expected from a random pattern of neighbourhoods, can be identified by measuring the statistical significance of each local centralization value. Alternatively, particular locations of interest can be identified as foci to determine if a statistically significant segregated pattern exists around the locations. This flexibility in usage makes the index a general tool useful for the detection of unusually high pockets of segregation on a map, which could in turn help generate hypotheses on the forces driving the segregation levels for those areas and the region as a whole.

This introduction is followed by the motivation for the local centralization index and how it fits within the broader segregation measurement landscape. We then introduce the local centralization index and highlight its properties relative to other measures. An empirical example of black-white segregation in Phoenix, Arizona is used to illustrate the measure, and the paper ends with a summary and directions for further research.

2 Motivation

2.1 Dimensions of segregation

Residential segregation measures attempt to capture the spatial separation of different types of people within an urban area. In 1988, Massey and Denton (1988) brought order to the multitude of available measures by categorizing twenty of them into five dimensions: evenness, exposure, concentration, centralization and clustering. While this pentad remains an important organizational model for theoretical and empirical research on segregation, alternative models have since been put forward. In most cases these new models have abandoned the centralization dimension that forms the core of the work at hand.

The five Massey and Denton dimensions can be split into ‘aspatial’ and ‘spatial’ categories where those in the former category are unaffected by a rearrangement of neighbourhoods within an urban area. The first of the three aspatial dimensions is evenness, which identifies if neighbourhoods tend to have a balanced ratio of different types of residents or if single population groups tend to dominate each neighbourhood. Exposure captures the likelihood of interaction of one population group with another, where urban areas with mixed neighbourhoods provide more exposure to the other group than urban areas where groups tend to be separated into isolated neighbourhoods. The third aspatial dimension, concentration, captures the density of a population group, identifying if one population group occupies a disproportionately small amount of land area within the urban area. Centralization, a spatial dimension, considers how population groups are distributed relative to the urban centre. Finally, clustering measures the tendency of neighbourhoods with similar demographic profiles to be located near one another.

More recent segregation models have tended toward a two dimensional approach. Johnston et al. (2007) identifies the dimensions of separateness and location based on a statistical approach designed to mimic Massey and Denton (1988). Separateness combines evenness, exposure and clustering, and measures the relative distribution of two population groups to one another. Location combines centralization and concentration, and captures the concept of population density. Reardon and O'Sullivan (2004) build a theoretical segregation model that eschews the constraint of neighbourhood boundaries and argues for conceptualizing segregation at the person scale. Spatial exposure, analogous to exposure, captures the likelihood of encountering a member of another population group; and spatial evenness, which combines evenness and clustering, describes similarity in the distribution of population groups within an urban area. Brown and Chung (2006) also argues for two dimensions. The name concentration-evenness represents two extremes of one dimension covering the uneven or even spatial distribution of the population. Similarly, clustering-exposure represents extremes of the other dimension capturing the amount of interaction between population groups. Wong (2008b) constructs two dimensions: clustering and exposure. The former represents spatial notions of residential segregation and subsumes evenness; the latter covers segregation in other 'spaces' such as work, religion, entertainment, etc. where proximity might not reflect the magnitude of segregation.

While all these segregation models approach the dimensions of segregation differently, they all share the abandonment (or minimization) of the centralization dimension from the original list of five. Brown and Chung (2006, p. 126) states, 'in today's increasingly polycentric, multinodal and sprawled city, centrality has little meaning' . Wong (2008b, p. 458) reinforces this position stating, 'it has now become obvious that centralization no longer plays a role in the measurement of segregation because the polycentric city is today's norm'. Johnston et al. (2007, p. 501) temper this position somewhat stating, 'in MAs [metropolitan areas] in which members of an ethnic group live at relatively high densities, they also tend to live close to the city centre – though not necessarily in a single cluster'.

Each of the Massey and Denton dimensions of segregation rely on some assumption(s) of what segregation is, and the subsequent translation of that assumption to a metric. In the case of centralization, it appears that the assumption is that disproportionate concentration around the urban core is bad. However, the urban core of each city is different, and so applying this type of general rule to all cities is inappropriate. In contrast, exposure to other population groups is a more universal rule that can be applied across different urban areas. For this, and other reasons to be discussed later, we generally agree with the abandonment of centralization as a measure of global segregation as it is currently defined. We propose to repurpose it as a tool for studying local spatial segregation.

2.2 Space and segregation

The adoption of spatially explicit approaches to segregation measurement has evolved incrementally over the decades. Through the late 1970s, segregation index research focused primarily on the technical details of variation between areas (e.g., census tracts), with little concern about the absolute or relative location of areas within their region. That being said, early researchers did identify a number of spatial concerns that went beyond the otherwise sole focus on the aspatial part of segregation measurement. Cowgill and Cowgill (1951) recognized that the scale of the areas affects measured segregation. They advocated using smaller scale administrative areas than the commonly used census tracts when measuring segregation. However, as one continues to subdivide the areas, the population within becomes more and more homogeneous, resulting in a corresponding increase in measured segregation. In the extreme there is only one person per area, or complete segregation as measured by the aspatial indices of the day (Wong 1997). Duncan and Duncan (1955a) raised the issue of concentrated minority population and speculated that a region where the highly segregated areas are clustered together is distinct from a region with distributed pockets of high segregation. Their initial observation was on the concentrated central location of minority populations in regions. Over time the concept was broadened and labeled the 'checkerboard problem' (White 1983) in recognition that any spatial rearrangement of areas in a region resulted in the same level of measured segregation. Jahn et al. (1947, p. 294) noted that the measures they presented did not " 'correct' for intra-tract segregation or overlapping of segregation areas in parts of different census tracts". The insights of these early pioneers along with advances in spatial analysis led to advances in segregation measurement starting in the early 1980s (e.g., White 1983). The explicitly spatial research efforts over the subsequent three decades have brought a steady stream of candidate indices to the fore, each addressing the spatial aspects of segregation differently.

These criticisms of early segregation measures are grounded in the idea that segregation is not distributed equally across the urban landscape, a problem that can also be addressed through local segregation indices (see Wong 2002, 2008a; Feitosa et al. 2007, for reviews and proposals). Local indices allow one to map the segregation pattern within a metropolitan area. Early local measures, such as the Shannon and Theil indices, are not 'spatial' as we have defined it here since a neighbourhood's value would be the same no matter its location on the map. In the 1990s new measures were introduced that borrowed from the literature on local spatial autocorrelation (e.g., Anselin 1995). These indices compute segregation using sums or averages of the population data in surrounding neighbourhoods to give spatial context. The centralization index introduces a unique approach for capturing the local context be by incorporating these surrounding data incrementally at ever increasing distances from the 'centre'.

2.3 Computational background

The global centralization index was first introduced by Duncan and Duncan (1955b) as a technique for understanding the spatial footprint of occupational segregation relative to the central business district. In 1955, US urban structure was emerging from a monoconcentric urban form once dominated by the central business district (Glaeser and Kahn 2004). The rapid decentralization of population resulted in an abandonment of the urban core in favour of suburban locations. However, the decentralizing population was largely white and affluent, leaving the older central city to lower income minority residents. The centralization method, as originally envisioned, looks at the differential spatial distribution of two groups, for example, white and black residents, relative to the urban core.

Centralization, as presented in Duncan and Duncan (1955b), is a measure based on the index Corrado Gini introduced in 1914 (Giorgi 1990). As such, it depends on an explicit ordering of the neighbourhoods being studied. For the classic Gini index of segregation, neighbourhoods are ordered from lowest per cent minority[1] to highest (Duncan and Duncan 1955a). For centralization the neighbourhoods are ordered by distance from the urban centre. Centralization was originally defined as:

$$C = \sum_{j=2}^{n} \tilde{X}_{j-1}\tilde{Y}_j - \sum_{j=2}^{n} \tilde{X}_j\tilde{Y}_{j-1}, \tag{1}$$

[1] Per cent minority in this case is the minority population's share of the total neighbourhood population.

where $\tilde{X}_j$ ($\tilde{Y}_j$) represents the cumulative percentage of the X (Y) group population through the jth neighbourhood; with the central business district represented by $j = 1$ and n being the number of neighbourhoods in the region. The measure ranges from −1 to 1, where positive values signify that the X group is located closer to the center, relative to the Y group, and negative values indicate the X group is located further from the center, relative to the Y group. A value of 0 indicates the two groups are distributed equally relative to the urban core.

Two methods have been proposed for ordering neighbourhoods relative to the centre. The first assigns each neighbourhood to one of a series of concentric rings rooted at the urban core, and then computes centralization based on aggregated population counts in each ring (e.g., Duncan and Duncan 1955b; Galster 1984). This approach finds grounding in the concentric ring model of urban growth (Burgess 1928). The researcher thus needs to define the width of each ring and a method for assigning the neighbourhood polygons to a ring. A second approach involves ranking the neighbourhoods based on their distance from the urban core (e.g., Massey and Denton 1988; Dawkins 2004). In this case the researcher needs to pick a mechanism for defining distance between polygons, for example, distance between polygon centroids.

A key weakness of centralization as a global measure of segregation is choosing the ‘centre’. For the analysis of a single region, local knowledge can be used to identify the centre, but a cross sectional study of hundreds of regions does not allow for this type of detailed research on each location. Most studies identify the central business district (CBD) as the region centre (e.g., Duncan and Duncan 1955b; Massey and Denton 1988), although few actually state the method used to determine the CBD. The US *Census Bureau last defined* CBDs for US cities in the *1982 Census of Retail Trade* (US Census Bureau Geography Division 2011). Iceland et al. (2002) depart from the CBD pattern by using the population centroid as the centre. Duncan and Lieberson (1959), in a segregation study of foreign born workers from various European countries, recommends caution when comparing centralization values. In 1950, Dutch immigrants had an abnormally low centralization value compared to other immigrant groups in Chicago. This low value was not the outcome of Dutch integration into the broader community, but the presence of a Dutch enclave 12 miles from the city centre. The inability of the centralization index to identify Dutch immigrants as segregated speaks more to the weakness of its use as a global measure, and less on its inherent mathematical structure.

2.4 *Global centralization*

We illustrate the global centralization index using a region with 100,000 residents split 45–55 between minority and majority groups. The centre of the region is designated by a cross, with the centroid of each neighbourhood represented by a dot (Figure 1a). Distances from the centre are defined based on centroid-to-centroid distance, and this particular example has a unique ordering of neighbourhoods relative to the centre. Figures 1b, 1c and 1d present three different spatial allocations of the same population. In all cases there are ten neighbourhoods with 3,000 minority residents and 1,000 majority residents, and 15 neighbourhoods with 1,000 minority residents and 3,000 majority residents. Figure 1c shows that the minority population is generally concentrated around the region’s centre, while Figure 1d is the case where the majority population concentrates near the centre. The centralization indices for these two cases are 0.484 and –0.484.[2] Figure 1b represents an intermixed case, relative to the centre, and the resulting centralization index is zero.

[2] If the neighbourhoods in Figures 1c and 1d were exclusively majority or minority then the centralization indices would be 1.0 and −1.0.

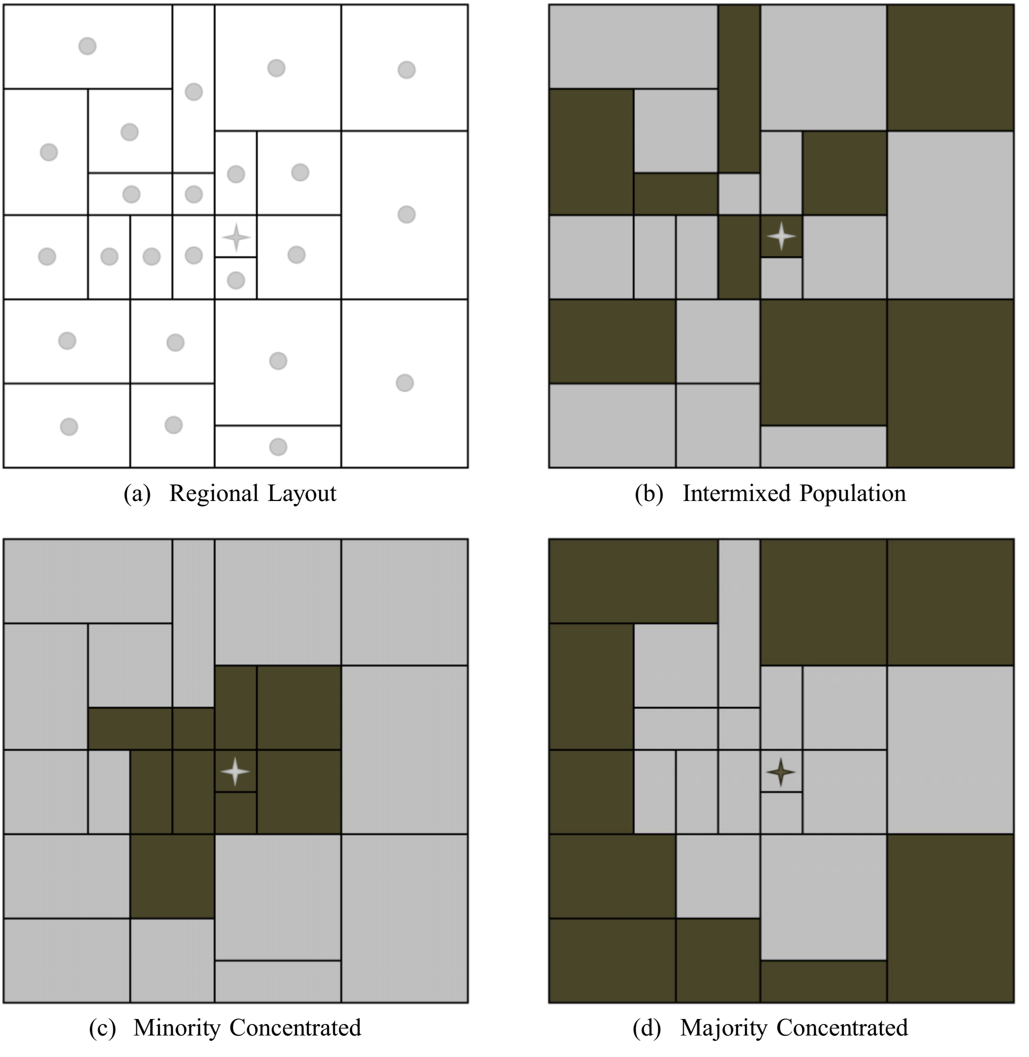


Fig. 1. Various distributions of minority and majority populations

A second way for the centralization index to achieve a value of zero is if every neighbourhood in the region has the same number of minority residents and the same number of majority residents; namely, each neighbourhood radiating out from the centre is identical to the neighbourhood one step closer to the centre. A strength of the centralization index is that a zero value is not constrained to this extreme case, but can emerge from any spatial pattern that results in a balanced distribution of the two population groups relative to the centre. This result can be visualized via the segregation curve (Duncan and Duncan 1955a) since the centralization index is a variation on the Gini index – the Gini index measures the area between the segregation curve and the 45 degree line of exact evenness. The segregation curve for the centralization index is derived by ordering the neighbourhoods by distance from the centre, as opposed to per cent minority for the classic segregation Gini index. It is therefore possible that the centralization index segregation curve can exceed the 45 degree line (Duncan 1957). Figure 2 shows that for the spatial arrangement from Figure 1c, with the minority population concentrated near the centre, the curve never exceeds the 45 degree line, note that for this example the classic Gini

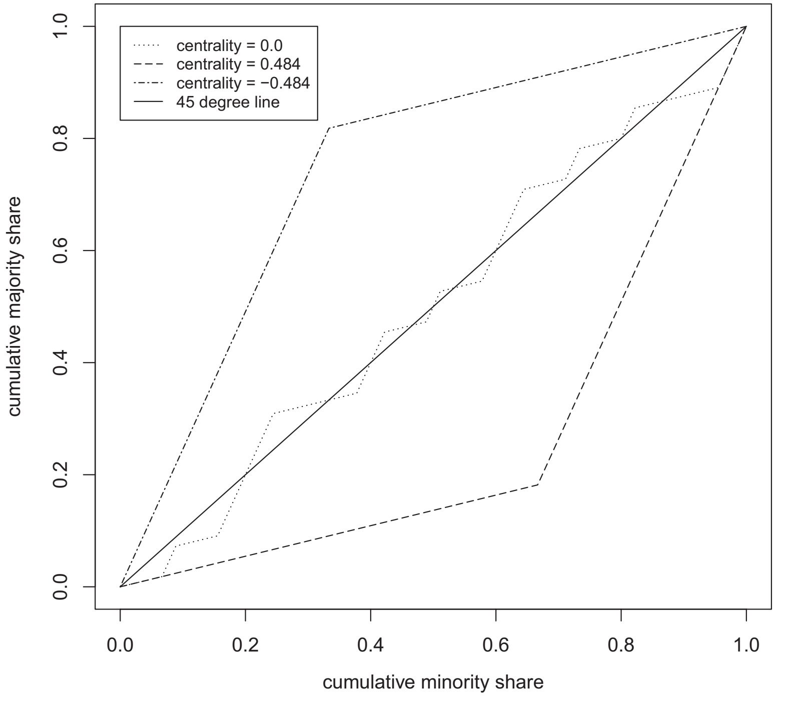


Fig. 2. Centralization index: segregation curves

index would also equal 0.484. When the majority population is concentrated around the centre, the centralization index is simply the negative value of the minority-centric case, and this is also revealed in the Figure 2, which shows the symmetry of the two curves. In the case where the neighbourhoods are interlaced throughout the region, the segregation curve zig-zags around the 45 degree line, a pattern that results in the area above the line of evenness being exactly equal to that below the line.

The segregation curve also shows that given an allocation of residents to neighbourhoods, the neighbourhood constrained bounds of the centralization index fall inside of the unconstrained bounds of −1 and 1. Specifically, the constrained bounds of the centralization index are the classic Gini index and the negative of the Gini index for that particular aspatial allocation. These maximum and minimum values are realized when the ordering of the neighbourhoods by distance from the centre is identical to the ordering based on population shares (Dawkins 2004).

2.5 Variations on the Gini index

While our focus is centralization, there are several other Gini-based indices that should be mentioned. Massey and Denton (1988) present absolute centralization, which focuses on the density of X group persons relative to the urban centre. Specifically, it replaces $\tilde{Y}_j$ and $\tilde{Y}_{j-1}$ in Equation 1 with $\tilde{A}_j$ and $\tilde{A}_{j-1}$, where the new terms represent the cumulative share of land area from the urban core to neighbourhood j. As before, the index is bounded by 1 and −1, with positive values indicating that the X group is disproportionately concentrated near the urban centre relative to the land area. Negative values represent the opposite pattern, where the X group tends to be less densely packed near the centre.

Wong (2008a) argues that at its core, segregation measures the difference in the spatial distribution of two or more groups of people, and therefore a segregation measure should include data on at least two groups of people. From this perspective absolute centralization would not be a measure of segregation. Another measure along these lines is the index of zonal redistribution, which captures one group's change in centralization over time (Redick 1956). In this case $\tilde{Y}_j$ and $\tilde{Y}_{j-1}$ are replaced by the $\tilde{X}_j$ and $\tilde{X}_{j-1}$ values from the following time period.

Dawkins (2004) presents a general class of spatial Gini segregation indices, within which the centralization index falls. He also presents a spatial Gini index in which neighbourhoods are first ordered as they would be for the classic Gini index, and then population counts for each neighborhood are swapped with the counts for the neighbourhood's nearest neighbour. Another spatial Gini index, which does not appear to have been presented previously in the literature, is one based on a spatial moving window around each observation. This generalized Gini index replaces the raw counts of group *X* and *Y* type persons in each neighborhood with a convolution of the counts from the neighborhood of interest and those that surround it. In its simplest form, the generalized Gini index could be the sum of the residents in the neighborhood of interest and those adjacent to it; more complex formulations could involve a kernel smoothing function that includes some or all of the other neighbourhoods in the region. The generalized Gini index is similar in form to the generalized dissimilarity index (Wong 2005; Feitosa et al. 2007). Both of these spatialized Gini indices attempt to capture the influence that proximate neighbourhoods have on a neighbourhood of interest. In the former case, if each neighbourhood's nearest neighbour was quite similar to itself, then the spatial Gini would be only slightly lower than the classic Gini. In contrast, if the nearest neighbour is quite different, as could be generated by a checkerboard pattern, then the spatial Gini would be near zero. The generalized Gini index has a similar premise, but in general would bring more 'neighbours' into the computation. Unlike these other Gini-based measures, the centralization index can in principle be applied to point level data.

3 Local centralization

The centralization index is a bivariate spatial measure of segregation. To this point the index has been used as a global measure of segregation, with the specific aim of determining if a minority population is disproportionally concentrated around a region's urban centre relative to the majority population. In this section we recast the index as measure of local segregation. Specifically, the index can be computed for any location by treating it as a center.

The local centralization index for location *i* and spatial extent *k* is:

$$C_{i,k} = \sum_{j=2}^{J} \hat{X}_{j-1}\hat{Y}_j - \sum_{j=2}^{J} \hat{X}_j \hat{Y}_{j-1}, \tag{2}$$

where vectors $\hat{X}$ and $\hat{Y}$ are defined as:

$$\begin{aligned} \hat{X} &= \frac{x_1}{X_k}, \frac{x_1+x_2}{X_k}, \frac{x_1+x_2+x_3}{X_k}, \ldots, \frac{x_1+\cdots+x_{J-1}}{X_k}, 1 \\ \hat{Y} &= \frac{y_1}{Y_k}, \frac{y_1+y_2}{Y_k}, \frac{y_1+y_2+y_3}{Y_k}, \ldots, \frac{y_1+\cdots+y_{J-1}}{Y_k}, 1. \end{aligned} \tag{3}$$

x_j represents the count of the *X* group in observation *j*, where the observations are ordered by distance from the reference location *i*. There are *J* observations between *i* and *k*, where $\sum_{j=1}^{J} x_j = X_k$. $\hat{X}_j$ is therefore the cumulative share of the X_k population through the *j*th observation. As a result, $\hat{X}_j = 1$ when $j = J$.[3] Similar definitions hold for the Y group. Equation 2 reduces to the classic centralization index, Equation 1, when $J = n$ and i is the central business district.

The local centralization index summarizes three spatial relationships: (*i*) the location of group X relative to a chosen reference point; (*ii*) the location of group Y relative to that point; and (*iii*) the location of group X relative to group Y. These spatial relationships are operationalized through the ordering of population data based on distance from the reference location. Some spatial information is lost in the process as spatial relationships are reduced to a simple ordering of observations. However, this simplification allows for the use of cumulative sums at each step away from the reference location, which means the index can capture the differential concentration of the two groups relative to a point. This approach can be contrasted to local measures of spatial autocorrelation (e.g., Anselin 1995; Ord and Getis 1995), which compare the central observation to the average (or sum) of the spatial neighbours. The cumulative summations behind the centralization index limit it to count type data, where spatial autocorrelation measures can be used on counts, rates and other spatially intensive data.

This mathematical structure ties centralization to concentration. Centralization can be characterized as measuring the relative level of concentration (Duncan and Lieberson 1959) – relative to the centre and relative to the other group. Therefore, a group is centralized if it tends to be closer to the reference location than the other group (Galster 1984). The literature on global centralization argues that a concentrated population is not necessarily centralized within the metro area, but that a centralized population is by definition concentrated (Gibbs 1961; Massey and Denton 1988). In the local context, the necessary relationship between concentration and centralization applies in both directions. Centralization captures population concentration relative to a point, and a central point could be found for any population concentration. The key difference is in the interpretation: concentration is measured relative to a two-dimensional reference area, while centralization is measured in reference to a point and area.

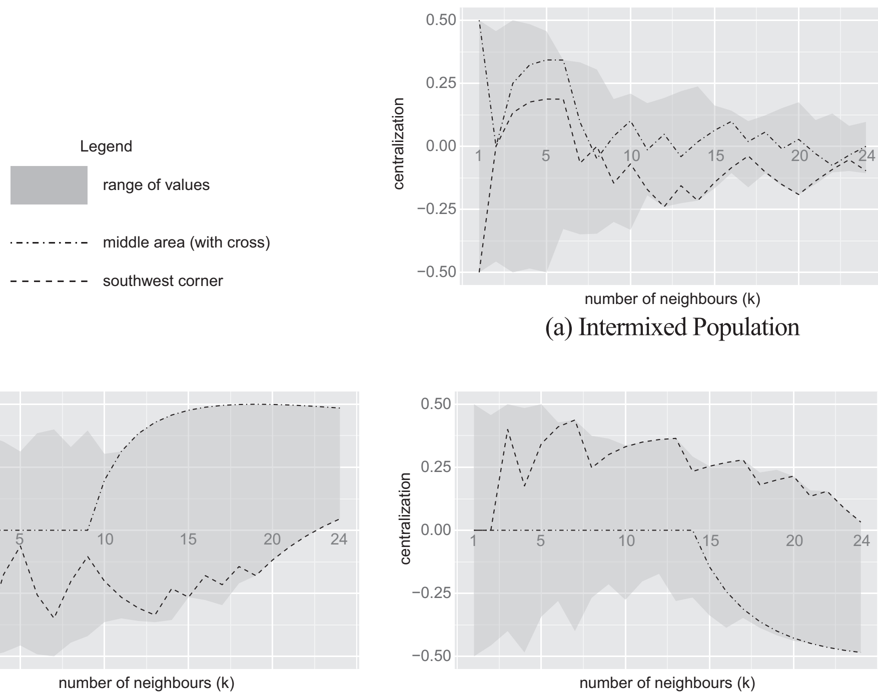
3.1 Local spatial relationships

As discussed in the previous section, there are a number of options for organizing the observations based on distance. There are two basic approaches: the first treats each observation as a discreet entity and the second sums observations based on some rule. Two examples of discrete approaches are (*i*) identifying the k nearest neighbours to the reference area and (*ii*) identifying all the observations that fall within a set distance band around the reference area. In either case, the observations need to be ordered by distance before entering the computation. An example of the summing approaches involves defining a series of distance based rings around the reference area, say at one mile intervals, then summing population data for the observations that fall within each ring. If ten rings are defined then $k = 10$.[4]

Like all local spatial measures, the definition of 'local' for the local centralization index is not clearcut – the choice of k depends on the context of the study. For example, the researcher may be interested in investigating local segregation within very close proximity of the reference location, within an intermediate area and then for a large area, in order to understand how segregation varies by distance. Each of these areas implies a different choice of k, which will likely result in different magnitudes of $C_{i,k}$. In Figure 3 we present the range of $C_{i,k}$ (shaded

[3] $\tilde{X}$, used in Equation 1, refers to an observation's population share relative to the entire region, while $\hat{X}$ is used when considering only the some subset J of the total observations. When $J = n$, $\tilde{X} = \hat{X}$.

[4] In Duncan and Duncan (1955b) the authors segment the metropolitan area using the ring method, and also a 'sector' method presumably based on the Hoyt (1943) sector model of urban structure. They compute a single centralization index for each sector based on the observations within that sector; all sectors share the same centre observation.



(a) Intermixed Population

(b) Minority Concentrated

(c) Majority Concentrated

Fig. 3. Centralization index ($C_{i,k}$) variation by number of neighbours (k)
Notes: The plots correspond to the maps in Figure 1. The lines represent variation in $C_{i,k}$ for a particular area as the number of neighbours increases.

regions) for the 25 areas in the maps presented in Figure 1 at values of k from 1 to 24 nearest neighbours; we also highlight two areas, the centre area (with the cross) and the area in the southwest corner. In all the plots there is more volatility in $C_{i,k}$ for lower values of k. This can be seen in the general smoothing as k increases and the large fluctuations at lower values. As k increases, the region's global pattern has increasing influence on the local values. Another important feature of $C_{i,k}$ is that context matters. In the two maps where the population groups are concentrated around the centre (Figures 1c and 1d), the $C_{i,k}$ values for the centre area (area with the cross) stay fixed at zero until k is large enough to encompass areas different from the reference area. Like most segregation indices, the local centralization index is sensitive to broader context, not absolute relationships between the two population groups at the location.

The summary measure $C_k = \Sigma_i |C_{i,k}|$ can be used to guide the selection of k. As the total centralization over all observations, C_k summarizes the combined local segregation pattern. In practice we are interested in the statistical significance of C_k as opposed to its magnitude. A significant C_k value can be interpreted as the combined local centralization from the actual map being statistically different from a spatially random map for that value of k. Significance is determined by randomly allocating the areas on the map, and computing C_k for the randomized map. This is done many times, with significance determined by counting the number of times the actual C_k is extreme relative to the random C_k values and comparing this to a chosen significance level using a two-tailed test.

Figure 4 applies this approach to the maps in Figure 1 and corresponds to the plots in Figure 3. We generated 999 random maps and computed C_k for $k = 1$ to 24 for each map. The plots all follow the same general pattern with a few significant values when k is small, then a zone of insignificant values, followed by another series of significant values. This transition zone

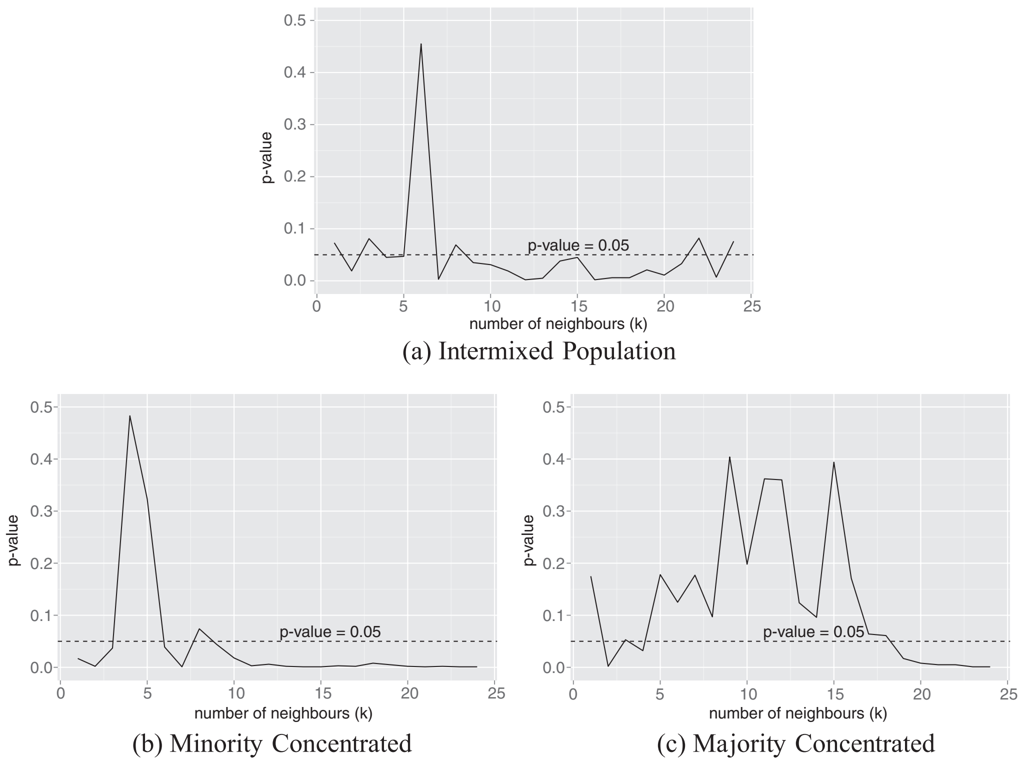


Fig. 4. Aggregate centralization index (C_k) significance by number of neighbours (k)
Notes: The plots correspond to the maps in Figure 1.

corresponds to the division between the high volatility in $C_{i,k}$ values when k is small and the more stable $C_{i,k}$ values as k increases. The recommended value for k is the smallest significant C_k after the transition zone – a k that is large enough to be beyond the high volatility region, but small enough to not be swamped by broad regional patterns. This would be $k = 9$, 11 and 19 for Figures 4a, 4b and 4c respectively.

3.2 Magnitude of the local centralization index

The local centralization index is bounded on the range [−1, 1], where negative values indicate centralization of the Y group and positive values indicate centralization of the X group around the reference location.[5] Unlike spatial autocorrelation measures, the centralization index is strictly bounded to this range and its extreme values are not directly affected by the spatial structure of the region (Tiefelsdorf and Boots 1995).

That being said, the empirical extrema for a reference location, its surrounding neighbours, and a given allocation of residents to neighbourhoods is the aspatial Gini index, and its negative complement.[6] The interpretation is that maximum (or minimum) $C_{i,k}$ occurs when the spatial ordering exactly matches an ordering by decreasing (increasing) minority racial shares. The local centralization index mirrors the Gini index by capturing the area between the segregation

[5] Typically the minority group is assigned to X and the majority group to the Y. Since the index is symmetric, the designation of X and Y can be reversed with a corresponding change to the interpretation.

[6] As discussed below, the local dissimilarity index (Wong 1996) also sits inside the [−1,1] range.

curve and line of complete evenness, although the segregation curve constructed via $C_{i,k}$ can be above or below the line of exact evenness (see Figure 2).

The magnitude of the local centralization index under various stylized conditions is presented in Figure 5 for a region with 16 neighbourhoods. Each neighbourhood can contain members of group A, group B or both. The top row of the figure presents the distribution of the group A population to the 16 neighbourhoods, and the leftmost column shows how group B is assigned to the region's neighbourhoods. The first three scenarios (1, 2 and 3) reflect clustering of similar neighbourhoods, and the latter three (4, 5 and 6) represent an interlaced, random and even spatial distribution of neighbourhood types respectively. Scenario 1 is the most extreme allocation of residents to neighbourhoods, with group population allocated as either high (4,000 people) or zero. The remaining five scenarios present the group population allocated to all 16 neighbourhoods; note that scenario 3 is simply one half of scenario 2. The interior of the figure presents the local centralization index for all scenario combinations, where k includes all observations; for example, case [A1, B5] represents the local centralization index when the group A population is concentrated in the northwest and the group B population has a random allocation.[7]

The extreme values of −1.0 and 1.0 can only be achieved when every observation in the region is entirely homogeneous, which corresponds to the only case in which the aspatial Gini index can equal 1.0. Case [A1, B1] in Figure 5 presents this extreme demographic pattern with six observations exclusively housing group A persons, and ten containing only group B persons. The northwestern most observation takes a $C_{i,k}$ value of 1.0 since the entire group A population is concentrated around this location. The southeastern corner ($C_{i,k} = -0.933$) falls just short of the other extreme since the particular orientation of neighbourhoods in this example does not entirely concentrate the group B population around this point. Each case in Figure 5 presents the Gini index for the particular population allocation to the 16 neighborhoods, and besides the [A1, B1] case all are less than 1.0. A standardized $C_{i,k}$ could be defined that divides the formulation in Equation 2 by the Gini index. This would essentially condition the spatial $C_{i,k}$ on the population's aspatial distribution.

Unlike the single extreme population pattern driving a $C_{i,k}$ value of 1.0 or −1.0, there are a number of ways the local centralization index can take a value of zero or near zero, indicating no segregation of the two groups. The lower right section of Figure 5 presents nine cases in which these low values can result for an entire map. The extreme case within this subset is case [A6, B6], where all the $C_{i,k}$ values are exactly 0 since every neighbourhood is exactly the same as all the others. This represents a case with no spatial or internal neighbourhood variation, hence no concentration of one population group relative to the other. Case [A4, B4] is the classic "checkerboard" pattern, again representing a case where there is essentially no concentration of either group around any point. The corner values being slightly different from 0 reflect a small boundary effect that can arise in a spatial measure. Case [A5, B5] is a random spatial pattern. The empirical extrema for these 16 observations is −0.447 and 0.447, but the actual maximum and minimum for this particular random spatial configuration is −0.056 and 0.056. The remaining six cases reflect combinations of the three scenarios, all of which result in near-zero $C_{i,k}$ values for all neighbourhoods.

The remaining maps present intermediate cases. In all cases the neighbourhoods along the diagonal stretching from southwest to northeast have $C_{i,k}$ values near zero, which reflects their spatial orientation on the border between the population group concentrations. Although in

[7] Since this is a regular lattice, some observations are equal distance from the central observation. These ties can be broken randomly or the observations summed and treated as a single observation. In this case we treat them as a single observation. Although summing can potentially generate abnormally large observations, the computation is driven by comparing the majority and minority counts in each observation meaning that a larger observation has the potential to have more majority and minority residents.

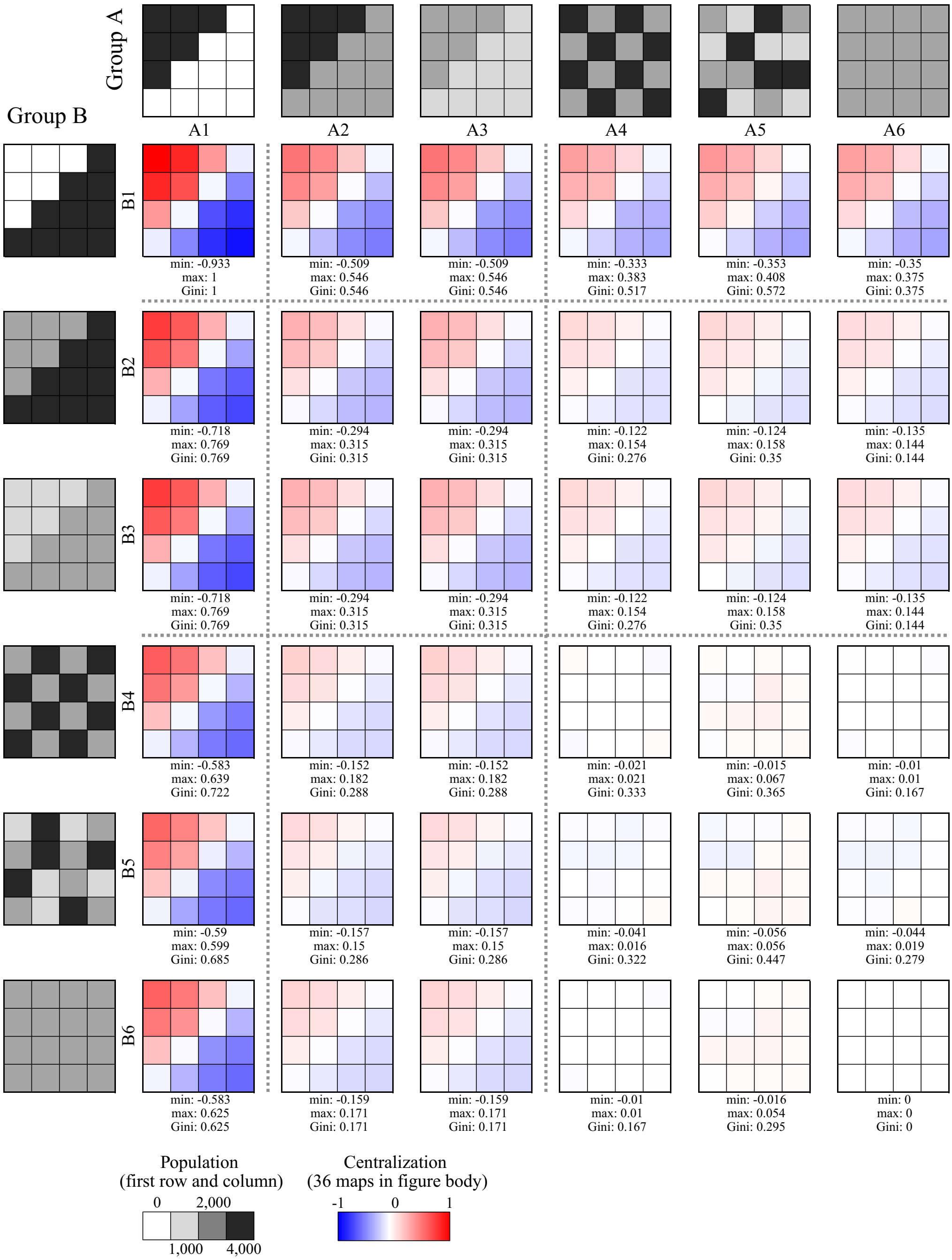


Fig. 5. Magnitude of the local centralization index

Notes: This figure presents 36 maps of the local centralization index, based on six population distribution scenarios. The population distributions driving the centralization values are in the first row and column. Below each map is the minimum (min) centralization index for that map, the maximum (max) and the aspatial Gini index. Segregation of group A is reflected by positive (red) values, and negative values (blue) represent group B segregation.

isolation these diagonal neighbourhoods may be homogeneous or dominated by one group, the local centralization index attempts to capture the broader spatial context in which the neighbourhood sits. The maximum and minimum values for all the cases in row B1 and column A1 are higher in absolute value than any maximum and minimum values for more interior maps in the figure. This indicates that a large number of clustered homogeneous neighbourhoods contribute greatly to the magnitude of $C_{i,k}$.

Another local centralization index characteristic is its invariance to a proportional population change in all neighbourhoods. In Figure 5, distribution A3 is one half of A2, and B3 is one half of B2. The effect of these proportional shifts in the population counts on $C_{i,k}$ is most clearly demonstrated by the identical results for cases [A2, B2], [A2, B3], [A3, B2] and [A3, B3]. Furthermore, each pair of $C_{i,k}$ maps in columns A2 and A3 are identical, and each pair in rows B2 and B3 are identical. The benefit of this property is that two regions, one where the minority population represents 40 per cent of the total population and one where the group represents 30 per cent, will have the same $C_{i,k}$ values if the relative spatial distribution is the same.

3.3 Comparison to related measures

To illustrate some of these properties in relation to other local measures, we again present a simple 4×4 region of 16 observations, with a minority population located exclusively in the northwest of the region, and the majority everywhere else (see Figures 6a and 6b). The local centralization values for these observations are presented in Figure 7a. For this simple example k includes all observations.

Comparing the centralization index to local measures of spatial autocorrelation requires some transformation of the data since these measures are univariate. The count data for each observation is transformed as $\overline{x}_i = x_i / (x_i + y_i)$, or the percentage of the observation's population in the minority group. In this simple example, this transformation results in three observations taking the value of 1.0, namely, 100 per cent minority population and the remainder taking values of 0. The local Moran's I for this map is presented in Figure 7b using a rook contiguity weights matrix.

The magnitude of the local Moran's I is a less than ideal measure for capturing residential segregation. Positive values of the measure indicate clustering of similar groups on the map, and negative values indicate an interlaced pattern of the groups. Therefore, while all values below zero can be ignored (there are three on the map presented), the raw index value does not provide information on whether the clustering is for the minority or majority population. Furthermore,

0	0	4,000	4,000
0	4,000	4,000	4,000
4,000	4,000	4,000	4,000
4,000	4,000	4,000	4,000

(a) Majority Population

4,000	4,000	0	0
4,000	0	0	0
0	0	0	0
0	0	0	0

(b) Minority Population

Fig. 6. Map of minority and majority resident counts by neighbourhood

1.0	0.8205	0.2564	–0.1538
0.8205	0.5641	–0.3077	–0.6154
0.2564	–0.3077	–0.8974	–0.8718
–0.1538	–0.6154	–0.8718	–1.0

(a) Local Centralization

4.0625	0.7292	–0.1683	0.2163
0.7292	–0.3606	0.2163	0.2163
–0.1683	0.2163	0.2163	0.2163
0.2163	0.2163	0.2163	0.2163

(b) Local Moran's *I*

2.0817	0.8006	0.1601	–0.4804
0.8006	0.5444	–0.4804	–0.4804
0.1601	–0.4804	–0.4804	–0.4804
–0.4804	–0.4804	–0.4804	–0.4804

(c) Local Getis–Ord $G*$

0.3333	0.3333	–0.0769	–0.0769
0.3333	–0.0769	–0.0769	–0.0769
–0.0769	–0.0769	–0.0769	–0.0769
–0.0769	–0.0769	–0.0769	–0.0769

(d) Local Dissimilarity

Fig. 7. Map of local measures

the magnitude of the local Moran's *I* is difficult to interpret since it is computed using *z*-scores, which are not bounded. In this simple example, the maximum local Moran's *I* value is 4.0625 for the observation in the northwest corner.

The Getis-Ord local G_i^* may be a more appealing candidate for this analysis. Using the $\bar{x}_i$ dataset again, positive values indicate clusters of the minority group and negative values indicate clusters of the majority group as seen in Figure 7c. While G_i^* can distinguish the type of cluster, interpreting the magnitude can be difficult since it is not bounded; the maximum value in this example is 2.0817 and the minimum is –0.4804.

Many local spatial segregation measures are derived from a global spatial measure, where the local measure is the individual contribution an observation would have made to the global measure. The local dissimilarity index (Wong 1993) is an example of this design. The spatial aspect comes through the replacement of each observation's individual data by summing its data it with that of its contiguous spatial neighbours. The index has the appealing property that positive values represent minority segregation in an observation, and negative values represent majority segregation. However, the magnitude of the index is not easy to interpret. The index is technically bounded on the range [–1, 1], however the end points become exceedingly difficult to reach as the number of observations within the region increases. As can be seen in Figure 7d, the maximum value is 0.3333, with a minimum value of –0.0769. It is also the case that the sum of local dissimilarity indices for all 16 observations is zero; this property holds for any population configuration. As the number of observations within a region increases, the local dissimilarity index for any particular observation tends toward zero. This does not mean that the index values become meaningless as the number of observations increases (see Wong 2008b, for an example using this index on Buffalo, New York). The magnitudes can be used

0	0	1,000	4,000
0	1,000	4,000	4,000
1,000	4,000	4,000	4,000
4,000	4,000	4,000	4,000

(a) Majority Population

4,000	4,000	0	0
4,000	0	0	0
0	0	0	0
0	0	0	0

(b) Minority Population

Fig. 8. Map of minority and majority resident counts by neighbourhood, variation in total population per neighbourhood

1.0	0.9225	0.3333	–0.1628
0.9225	0.6589	–0.2791	–0.6512
0.3333	–0.2791	–0.876	–0.8915
–0.1628	–0.6512	–0.8915	–1.0

(a) Local Centralization

0.3333	0.3333	–0.0233	–0.093
0.3333	–0.0233	–0.093	–0.093
–0.0233	–0.093	–0.093	–0.093
–0.093	–0.093	–0.093	–0.093

(b) Local Dissimilarity

Fig. 9. Map of local measures on region with variation in total population

to study the segregation distribution within a city, but they are difficult to interpret on their own and cannot be compared to values in cities with different numbers of neighbourhoods.

Some features of spatial autocorrelation measures make them less appealing as measures of local segregation. Since the autocorrelation measures require a conversion of the bivariate data to a single variable, some information is lost in the transition. Specifically, $\bar{x}_i$ can be insensitive to changes in population magnitude. Figure 8a shows that the majority population is reduced from 4,000 to 1,000 in three observations. The obvious result here is that the spatial autocorrelation measures are insensitive to this change since $\bar{x}_i$ does not change for any of the observations. In contrast, the centralization index, Figure 9a, reflects these changes in the population counts as increased minority population concentration in the northwest, and increased majority concentration in the southeast. The local dissimilarity index also captures change in population by showing lower majority group segregation in the observations with reduced population (Figure 9b).

The examples presented here involve areal observations since this type of data is most often available to social science researchers. However, the index could be applied, without modification, to point data on the locations of individuals. The cross-*k* function (Lotwick and Silverman 1982; Diggle 1983), which also operates on bivariate point data, bears some similarities to $C_{i,k}$. The function could be applied in a similar manner to $C_{i,k}$ by drawing a radius around a reference location to determine the spatial relationship between the two point patterns. However, it only measures the location of one group relative to the other; it does not take into consideration their positions relative to the reference location as $C_{i,k}$ does.

4 Empirical example

We explore the properties of the local centralization index using the black and white population in the Phoenix metropolitan area for 1990, 2000 and 2010. This section is not intended to be a comprehensive analysis of black and white segregation within the region, but a demonstration of the index in an empirical setting.

The Phoenix metropolitan statistical area (MSA) is one of the fastest growing regions in the US, having a population that nearly doubled between 1990 and 2010 (see Table 1). The majority white population has grown by approximately 30 per cent, while the black, Asian and Pacific islander and Hispanic populations have each grown by more than 60 per cent. The metropolitan area is becoming more diverse overall, with the white population dropping from 76.2 per cent in 1990 to 58.7 per cent in 2010.

The dissimilarity index indicates that segregation between the black and white populations has declined over the three time periods studied (Table 1). The largest decline came between 1990 and 2000, and the index had only a small drop over the subsequent decade. However, this global measure tells only part of the story – local segregation indicates that at the neighbourhood level segregation may be increasing.

We can hold the Phoenix metro area boundary constant over the years studied, but the number of census tracts within the region grew each decade as the population grew. If we set k to a constant number of nearest neighbours, the resulting $C_{i,k}$ values would tend to be computed over smaller and smaller land areas as the decades progress. We therefore define a fixed spatial extent of local segregation using a distance band around each tract; all tracts whose centroid falls within that band are ordered based on distance for the $C_{i,k}$ computation. To determine the distance band, we compute C_k at 0.2 mile increments from 1 to 20 miles, and check the statistical

Table 1. Population and segregation data, 1990, 2000 and 2010 for Phoenix MSA

	1990		2000		2010		Percentage Change
	Count	%	Count	%	Count	%	1990–2010
Total	2,238,480	100	3,251,876	100	4,192,887	100	87.3
White	1,705,976	76.2	2,140,171	65.8	2,460,541	58.7	44.2
Black	74,312	3.3	113,179	3.5	193,497	4.6	160.4
Am Indian	41,672	1.9	58,122	1.8	76,662	1.8	84.0
Asian and Pac Isld	34,435	1.5	69,399	2.1	142,627	3.4	314.2
Other	2,525	0.1	4,255	0.1	5,995	0.1	137.4
Two or More	na	na	49,738	1.5	77,847	1.9	na
Hispanic	379,560	17.0	817,012	25.1	1,235,718	29.5	225.6
Dissimilarity (Black–White)	0.5005		0.4508		0.4363		
Number of Tracts	490	100	695	100	991	100	102.2
Seg. White	24	4.9	43	6.2	67	6.8	179.2
Seg. Black	27	5.5	54	7.8	83	8.4	207.4
Max White Seg	–0.6707		–0.5671		–0.4765		
Max Black Seg	0.6498		0.6161		0.4349		

Notes: Local segregation uses a significance level of $\alpha = 0.05$.

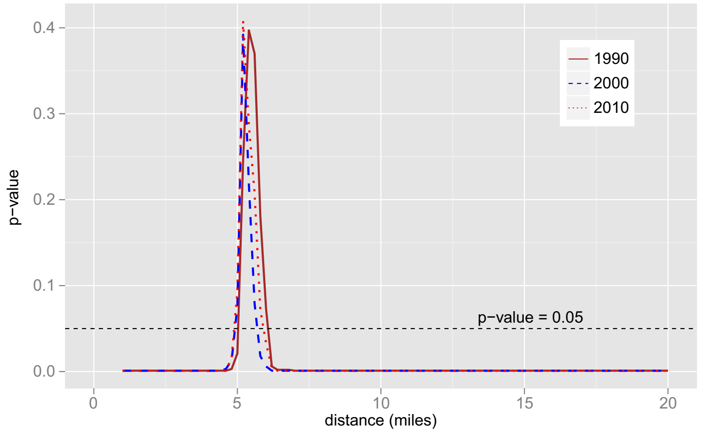


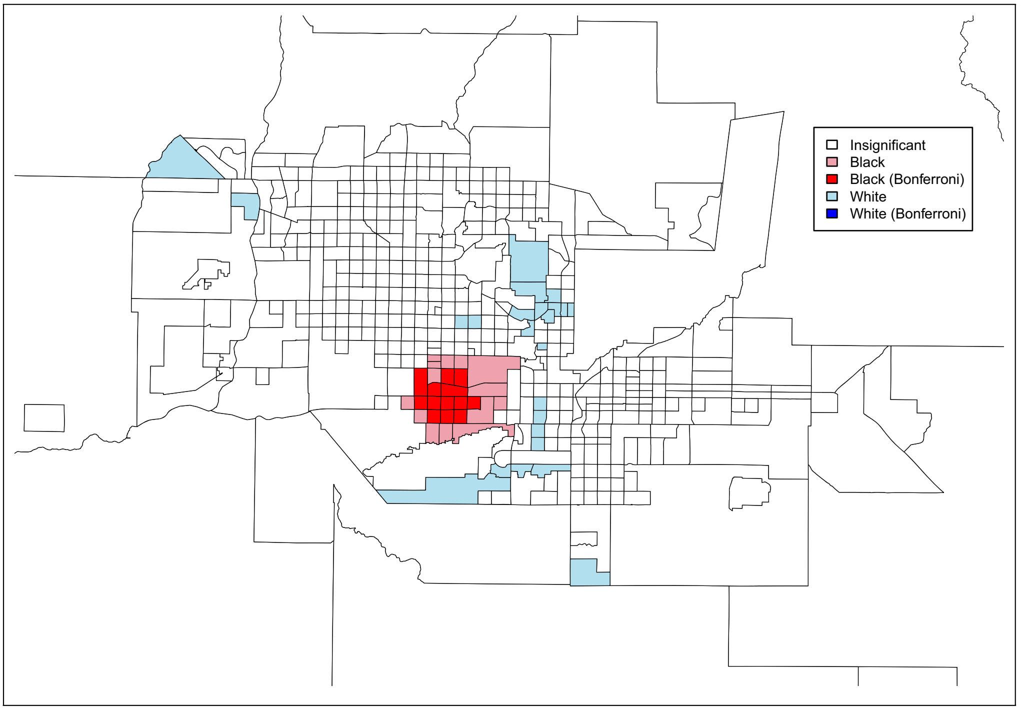
Fig. 10. Aggregate centralization index (C_k) significance by distance for Phoenix 1990, 2000 and 2010

significance using 999 spatially random permutations of the population data. The results of these tests for the three years (Figure 10) show the same basic pattern of insignificant values between approximately five and six miles. Based on these results we select a distance band of 6.2 miles, which is the smallest distance that is significant in all three years, but is beyond the spike of insignificant results. Considering the region's generally low population density, this corresponds to a distance that will extend beyond any localized population concentration, but not be too large to overly smooth the data. Segregation is not computed for large outlying census tracts, with no tracts within the defined radius.

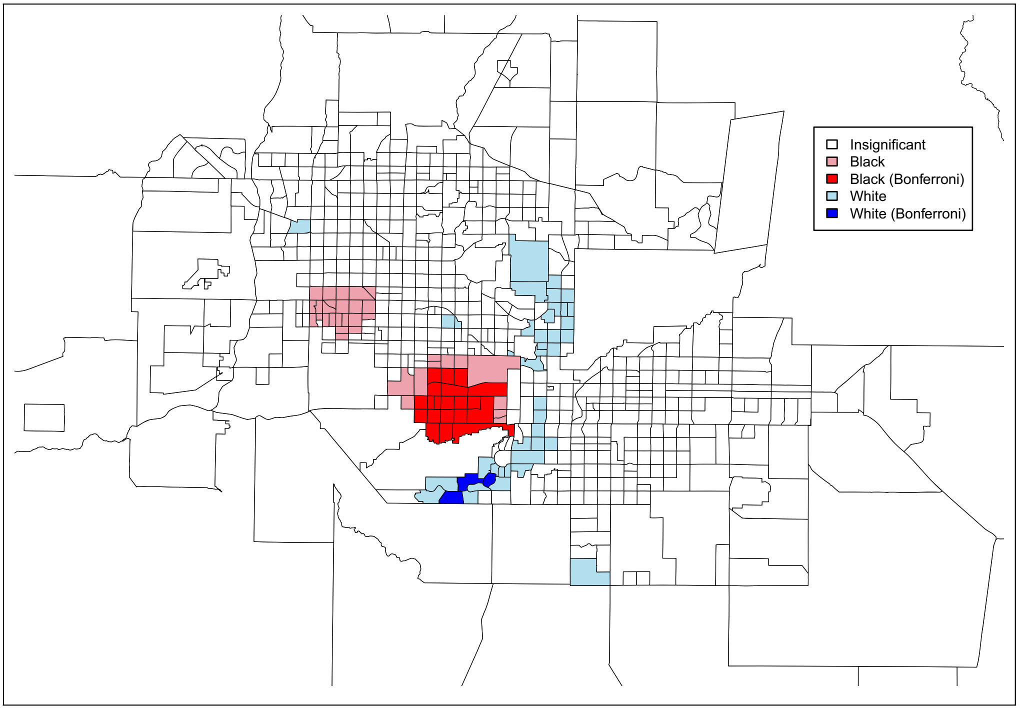
One challenge of local spatial measures is sorting through the mass of results to find meaningful information on which to concentrate. Here we borrow from the literature on local indicators of spatial autocorrelation (Anselin 1995) to measure the statistical significance of each tract's $C_{i,k}$ value. This allows for the identification of so-called hot spots where an observation or set of proximate observations is significant (Besag and Newell 1991; Ord and Getis 1995). The pseudo-significance of each measured segregation value is computed by permuting all the census tracts 9,999 times, and computing $C_{i,k}$ on each permutation. If the actual $C_{i,k}$ value is extreme relative to the distribution generated from the permuted maps, then that $C_{i,k}$ is considered significant. This is a two-tailed test, with a significance value of $\alpha = 0.05$.

Local measures such as $C_{i,k}$ are susceptible to the multiple testing problem. One typical solution involves using a significance level of α/n, the Bonferroni correction. Since the Phoenix MSA has a relatively high number of census tracts, the Bonferroni corrections for all three years reduces an α of 0.05 down to 0.0001, the minimum possible critical value for 9,999 permutations. A more liberal approach involves a correction of $\alpha/\bar{k}$, where $\bar{k}$ would be the average number of neighbours in the 6.2 mile radius. For the three points in time, $\bar{k}$ is 63, 77 and 81, which results in an adjusted significance value of approximately 0.0007. Using this approach the number of significant tracts declines from the numbers presented in Table 1 to 13, 23 and 23 for the respective years. We present the results for both significance levels in the maps below.

We define those tracts with statistically significant local centralization indices as being 'segregated'; and then split this set of tracts into 'black segregated' or 'white segregated' depending on the sign of $C_{i,k}$. Over the three time periods studied, the number of total census tracts increased by approximately 100 per cent, while the number of segregated tracts has increased by nearly 200 per cent (see Table 1). The rate of increase was similar for the two groups, but the number of black segregated tracts increased faster. While the number of black segregated tracts was always higher than those segregated white, the gap widens over the two

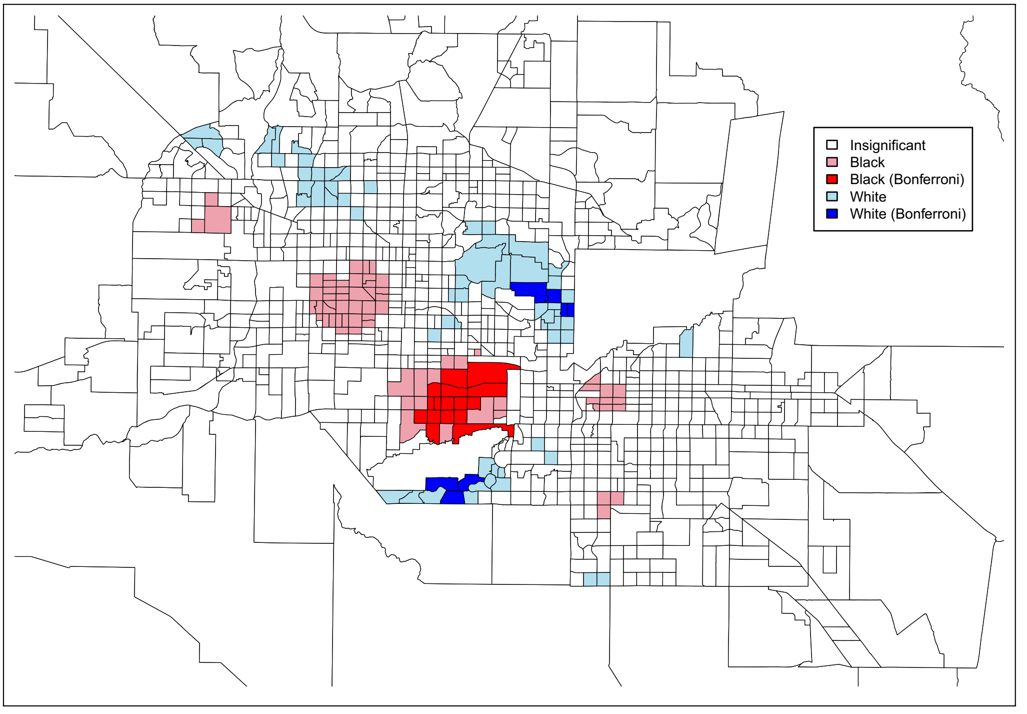


(a) 1990



(b) 2000

Fig. 11. Phoenix MSA black–white segregation, local centralization index



(c) 2010

Fig. 11. Continued

decades; ending with 6.8 per cent of tracts segregated white, and 8.4 per cent segregated black. Over the three time periods maximum and minimum segregation values have decreased in absolute value – the white level steadily decreasing and the black level decreasing mostly between 2000 and 2010. Overall we see a pattern of segregation that is becoming wider, in terms of its spatial footprint, but flatter in terms of its magnitude.

Figure 11 shows the changing spatial segregation pattern by identifying significantly segregated tracts, where darker colors indicate significance after a Bonferroni correction ($1/\overline{k}$). Over the three time periods three pockets of segregation remain intact: an area of black segregation (coloured red on the 1990 map) in the southern part of the city of Phoenix; a white segregated area in the northern part of Scottsdale (northeastern most blue area on the 1990 map), an affluent suburb of Phoenix; and the Ahwatukee area (south of the core black segregated area), a rapidly growing and geographically isolated community that is surrounded by a large mountain preserve to the north and west, an Indian reservation to the south and a freeway to the east. The south Phoenix pocket of segregation grew over the decades; but the number of tracts meeting the Bonferroni criterion peaked in 2000. This might be a reflection of the gentrification trend in many US cities as a more diverse population returns to older central cities near downtowns. The Ahwatukee area similarly maintained a core set of segregated tracts over the years, some eventually meeting the Bonferroni criterion, with the overall footprint expanding. North Scottsdale also followed this pattern, but the footprint noticeably changed at each decade.

Beyond these three pockets, other areas of interest emerged between 1990 and 2010. As a rapidly growing region, the metro area has engulfed many formerly isolated municipalities. These suburban cities generally have an older core area of homes surrounded by new subdivisions on former farm land. There appears to be a pattern of black segregation emerging around these older suburban areas. There is also indication of white segregation in large age restricted retirement communities in the northwest (1990 and 2010) and southeast (all years) parts of the region. It must be added that the vast majority of tracts do not show any significant segregation of either population.

Overall, the trends in black-white segregation in Phoenix show mixed results. The magnitude of segregation, appears to be declining but the share of tracts that are segregated is increasing, and there appear to be new pockets of segregation emerging. The results hint at potential explanations of these patterns such as gentrification and filtering as older housing is occupied by minority families.

5 Conclusion

The centralization index is by no means a new tool for the analysis of segregation. However, its use over the past 60 or so years was confined to a global measure of segregation concerned with the concentration of a minority population around the urban centre. Over the past 15 years most researchers have abandoned the index as a measure of segregation due to its weaknesses in the context of modern polycentric regions and reinvigorated downtowns. These weaknesses of the centralization index as a global segregation measure make it an ideal local index.

By sequentially designating each neighbourhood within a region as the centre, the index shines a spotlight on the social processes occurring within the region. The index does not just assess the population as a whole around each location, it systematically evaluates which population group is closer to the reference location. This results in a unique view into the pattern of segregation. The example of black-white segregation in the Phoenix metropolitan area showed the power of the index as a tool for exploratory spatial data analysis. By looking at significantly segregated neighbourhoods, the stable cores of segregation could be identified, and emerging patterns of segregation were visible at various locations across the region.

The local centralization index is presented here as an additional tool for the study of residential segregation. While a number of advantages have been shown, the measure has a number of limitations. Space is captured by a simple spatial ordering of observations. This has the advantage of ensuring that the measure always remains bounded to the $[-1, 1]$ range, but might be an oversimplification of complex spatial relationships in actual cities. The index also may be considered hyper-local, as region-wide information is only included when k encompasses all observations within the region. Like most segregation measures, the magnitude of $C_{i,k}$ is relative to the context in which it sits, and this context is entirely determined by k. Therefore the choice of k is an important decision for the user as some concentrations might be missed when k is too small, and an overly large k can overly smooth the data. We present C_k as a tool to help identify k in an empirical setting. Finally, local measures are susceptible to the multiple comparison problem. Due to the spatial correlation in the data, the problem is likely not on the order of $1/n$ as in the classic Bonferroni correction, but some adjustment is warranted to ensure that the concentrations identified truly reflect the stated significance level.

In this paper we used a permutation based approach to identify significance levels, future research could include identifying an analytic mean and standard deviation to provide further insight into the index. Variations on the index could also be explored, including dividing $C_{i,k}$ by its corresponding Gini index value to control for the aspatial population distribution; and applying the index to the location of individuals. Finally, all the examples presented here assume an isotropic spatial pattern around the focal location; a more nuanced bounding area could be defined to reflect non-symmetric patterns around the location of interest.

References

Anselin L (1995) Local indicators of spatial association: LISA. *Geographical Analysis* 27: 93–115
Besag J, Newell J (1991) The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 154: 143–155
Brown L, Chung S (2006) Spatial segregation, segregation indices and the geographical perspective. *International Journal of Population Geography* 12: 125–143
Burgess EW (1928) Residential segregation in American cities. *Annals of the American Academy of Political and Social Science* 140: 105–115
Cowgill D, Cowgill M (1951) An index of segregation based on block statistics. *American Sociological Review* 16: 825–831
Dawkins CJ (2004) Measuring the spatial pattern of residential segregation. *Urban Studies* 41: 833–851
Diggle P (1983) *Statistical analysis of spatial point patterns*. Academic Press, London
Duncan OD (1957) The measurement of population distribution. *Population Studies* 11: 27–45
Duncan O, Duncan B (1955a) A methodological analysis of segregation indexes. *American Sociological Review* 20: 210–217
Duncan OD, Duncan B (1955b) Residential distribution and occupational stratification. *American Journal of Sociology* 60: 493–503
Duncan OD, Lieberson S (1959) Ethnic segregation and assimilation. *American Journal of Sociology* 64: 364–374
Feitosa FF, Câmara G, Monteiro AMV, Koschitzki T, Silva MPS (2007) Global and local spatial indices of urban segregation. *International Journal of Geographical Information Science* 21: 299–323
Galster GC (1984) On the measurement of metropolitan decentralization of blacks and whites. *Urban Studies* 21: 465–470
Gibbs JP (1961) *Urban research methods*. Van Nostrand, Princeton, NJ
Giorgi GM (1990) Bibliographic portrait of the Gini concentration ratio. *International Journal of Statistics* 48: 183–221
Glaeser EL, Kahn ME (2004) Sprawl and urban growth. In: Henderson JV, Thisse JF (eds) *Handbook of regional and urban economics*. Elsevier, Amsterdam
Hoyt H (1943) The structure of American cities in the post-war era. *American Journal of Sociology* 48: 475–481
Iceland J, Weinberg DH, Steinmetz E (2002) Racial and ethnic residential segregation in the United States: 1980–2000. US Census Bureau, Series CENSR-3. US Government Printing Office, Washington, DC
Jahn J, Schmid CF, Schrag C (1947) The measurement of ecological segregation. *American Sociological Review* 12: 293–303
Johnston R, Poulsen M, Forrest J (2007) Ethnic and racial segregation in US metropolitan areas, 1980–2000: The dimensions of segregation revisited. *Urban Affairs Review* 42: 479–504
Lotwick HW, Silverman BW (1982) Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society. Series B (Methodological)* 44: 406–413
Massey D, Denton N (1988) The dimensions of residential segregation. *Social Forces* 67: 281–315
Ord JK, Getis A (1995) Local spatial autocorrelation statistics: Distributional issuesand an application. *Geographical Analysis* 27: 286–306
Reardon SF, O'Sullivan D (2004) Measures of spatial segregation. *Sociological Methodology* 34: 121–162
Redick RW (1956) Population growth and distribution in central cities, 1940–1950. *American Sociological Review* 21: 38–43
Tiefelsdorf M, Boots B (1995) The exact distribution of Moran's I. *Environment and Planning A* 27: 985–999
US Census Bureau Geography Division (2011) Central business districts 1982 census of retail trade. URL: http://www.census.gov/geo/www/cbd.html
White MJ (1983) The measurement of spatial segregation. *American Journal of Sociology* 88: 1008–1018
Wong DWS (1993) Spatial indices of segregation. *Urban Studies* 30: 559–572
Wong DWS (1996) Enhancing segregation studies using GIS. *Computers, Environment and Urban Systems* 20: 99–109
Wong DWS (1997) Spatial dependency of segregation indices. *The Canadian Geographer/Le Géographe canadien* 41: 128–136
Wong DWS (2002) Modeling local segregation: A spatial interaction approach. *Geographical and Environmental Modelling* 6: 81–97
Wong DWS (2005) Formulating a general spatial segregation measure. *The Professional Geographer* 57: 285–294
Wong DWS (2008a) Conceptual and operational issues in incorporating segregation measurements in hedonic price modeling. In: Baranzini A, Ramirez JA, Schaerer C, Thalmann P (eds) *Hedonic methods in housing markets*. Springer, New York
Wong DWS (2008b) A local multidimensional approach to evaluate changes in segregation. *Urban Geography* 29: 455–472

doi:10.1111/pirs.12145

Resumen. Alrededor de un parque atractivo o una planta incineradora de residuos se pueden observar áreas segregadas, pero es probable que la magnitud y la participación en el grupo de las personas más próximas sean diferentes. Por lo tanto, se introduce una medida de segregación local que se puede aplicar a cualquier localización dentro de un área metropolitana, con la que se puede identificar el grupo que está relativamente más concentrado en torno a la localización de referencia. Se introduce, además, un enfoque de inferencia para identificar la significación estadística de un valor de segregación en particular. En un escenario de exploración, el índice se puede utilizar para generar un mapa de puntos de interés, sobre los que preguntarnos: “¿Por qué este grupo se concentra significativamente en torno a ese lugar?”

要約: 魅力的な公園または廃棄物焼却施設の周りに隔離された区域ができることがあるが、その近接する地域に住む人口の規模やその集団の特性は異なる可能性が高い。そこで、大都市圏内のどの区域にも適用でき、かつ当該地域の周辺に、比較的密度が高い集団を識別できる地域分離指標を導入する。さらに、推計アプローチを用いて、特定の分離値の統計的有意差を識別する。予備的セッティングにおいて、当該指標は、ホットスポットの分布を示す地図を作成するのに使用することできるが、そこから、「なぜその集団が当該地域の周辺に特に集中するか？」という疑問が生じる。