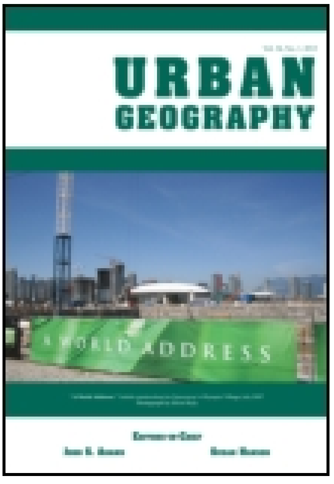


Urban Geography
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rurb20

Comparing Traditional and Spatial Segregation Measures: A Spatial Scale Perspective[1]

David W. S. Wong [a]
[a] George Mason University

Published online: 16 May 2013.

COMPARING TRADITIONAL AND SPATIAL SEGREGATION MEASURES: A SPATIAL SCALE PERSPECTIVE[1]

David W. S. Wong[2]
Earth Systems and GeoInformation Sciences Program
School of Computational Sciences
George Mason University

Abstract: Measuring the level of segregation often encounters two methodological issues: measures are sensitive to changes in the geographical scale of the data and the effectiveness of the measure in reflecting spatial segregation. Several spatial measures have been suggested to measure spatial segregation, but whether they are more or less sensitive to changes in spatial scale has not been investigated, while some spatial measures are relatively scale-insensitive. Using the 1990 Census data of 30 selected U.S. metropolitan areas, this paper demonstrates that these spatial measures, similar to the aspatial measure, report higher levels of segregation when smaller areal units are used in the analysis. Some spatial measures are even more sensitive to scale changes than aspatial measures. Certain patterns of the scale sensitivity were identified, but no general rules can be formulated. A preliminary explanation of the scale effect on spatial segregation measures is offered. [Key words: the MAUP, scale effect, dissimilarity index, spatial segregation measures, multigroup comparison.]

Since sociologists (Duncan and Duncan, 1955) attempted to quantify the level of segregation nearly half a century ago, contributions from geographers in the research of measuring segregation has been intermittent (e.g., Morgan, 1975, 1983; Jakubs, 1979, 1981). Recent research efforts related to Morrill's article (1991) have attempted to develop more effective tools to capture the spatial dimension of segregation. In short, the issue is that many traditional measures of segregation, including the most frequently used and overwhelmingly endorsed index of dissimilarity D, fail to differentiate different population distribution patterns. This problem has been illustrated with extreme cases such as the checkerboard pattern in which two population groups exclusively occupy alternate cells in a cellular landscape, and the two-region pattern in which the two groups exclusively occupy two different parts of the study region. The D values for the two distinctive patterns are both unity. The basic problem is that those traditional measures are aspatial in nature.

Several approaches have been suggested to introduce the spatial dimension into measuring segregation. One of the earlier approaches is to incorporate distance measures to capture the propinquity among population groups (Jakubs, 1981). Distance-based

[1] This project is partially supported by the National Institute of Health/Child Health and Human Development (NICHD) under National Institute of Health (NIH) grant number 1 R03 HG38292-01.
[2] Correspondence concerning this article should be addressed to David W. S. Wong, School of Computational Sciences, George Mason University, Fairfax, VA 22030; telephone: 703-993-1212; fax: 703-993-1980; e-mail dwong2@gmu.edu

Urban Geography, 2004, **25**, 1, pp. 66–82.

measures are effective as long as directional bias of population distribution is not a concern. Nevertheless, most distance-based measures of segregation require intensive optimization computing methods, which are not accessible to many researchers and practitioners. The recent approach to model spatial segregation involves capturing different aspects of the spatial dimension of population distribution that are potentially relevant to the spatial separation of population groups in the measures. These spatial aspects include the neighborhood relationship among population groups and the geometric characteristics of enumeration units from which population data were gathered.

From a geographical perspective, segregation index D and other traditional segregation measures are sensitive to the spatial scale of the population data. This problem is pervasive among almost all analytical measures and techniques adopted to analyze spatial data. The scale effect on segregation index D is one of the sub problems of the so-called modifiable areal unit problem (MAUP), and its implications in segregation studies have been briefly discussed (Wong, 1997). Some argued that because the MAUP is a spatial problem, the solution, if one exists, probably is a spatial one (Tobler, 1989). In addition, it is appropriate to use only inherently spatial techniques to handle spatial problems.

In the context of segregation studies, scale-independent segregation indices have not appeared yet. We do, however, have a set of spatially explicit segregation measures. Therefore, the intent of this paper is to determine whether this set of spatial segregation measures can handle the scale effect better or not. In other word, this study is to explore if these measures are as sensitive to scale changes as the traditional measures. If these spatial measures are not immune to the scale effect, then a related practical question is if they are more or less sensitive to scale changes than the traditional measures. Or will spatial measures offer more consistent results than that offered by traditional measures across scale levels?

This study selected a set of 30 U.S. metropolitan areas to explore the above issues. These 30 cities were selected because they were used in several previous studies (Massey and Denton, 1988; Wong et al., 1999). Census data at the block group and census tract levels for 1990 were used. Both traditional and spatial measures of segregation were computed for two-group and multigroup comparisons, while previous studies on the scale effect in segregation focused only on D for the two-group situation. The levels of segregation based upon these measures among the 30 areas were compared across the two scale levels. The next section provides a brief review of the MAUP and spatial segregation measures. Then the third section reports results from aspatial and spatial measures for two-group comparisons, and the fourth section reports results from multigroup comparisons. The final section presents the conclusion and a discussion.

REVIEW

The MAUP

In general, the MAUP consists of the subproblems of scale effect and zoning effect. When data at different scale levels or resolutions yield inconsistent results, this is referred to as the scale effect. Inconsistent results can also be obtained when the study region is partitioned by different spatial configuration schemes from which data are compiled. This is known as the zoning effect. Several papers have provided an overview of the MAUP.

Fotheringham and Wong (1991) offered a pessimistic view that the MAUP is omnipresence and may not have any resolution, while Reynolds and Amrhein (1998) demonstrated that the MAUP effects could be modeled.

In the past several decades or so, the MAUP has drawn the attention of geographers intermittently. Starting with the classical work by Openshaw and Taylor (1979), who demonstrated that correlation coefficients could change quite significantly when data at different scale levels were used, more recent publications have focused on deriving solutions (e.g., Holt et al., 1996). While King (1997) argued that the MAUP was “solved” using the technique of error bound in the context of ecological fallacy, geographers seem to be unable to come up with a unified or consistent verdict (Sui, 2000). In addition to the call for multiscale and multizonal sensitivity analysis to acknowledge the presence of the MAUP effects (Fotheringham, 1989), several application-specific solutions have been proposed so far. Tobler (1989) suggested a scale-independent framework to model inter-regional migration. Fotheringham et al. (1998) suggested that the spatial weighted regression technique might solve the MAUP in a regression framework. Wong (2001) demonstrated that a specific spatial correlation method is relatively stable to scale changes. All these potential solutions to the MAUP are spatial techniques, as Tobler (1989) suggested that we might need spatial tools to solve spatial problems.

While most of the discussions on the MAUP are on general spatial analysis or using traditional statistical techniques, very few discussions have focused on the impacts of the MAUP on measuring segregation, an issue of great concern in sociological and population studies. Wong (1997) discussed certain conceptual issues in measuring segregation as far as the scale effect is concerned. In general, the smaller the areal units are adopted for enumerating the population count data, the more homogeneous is the population within the areal units in general. As most segregation measures have a direct relationship to the internal homogeneity of population within areal units, smaller enumeration units will likely produce a higher level of segregation according to these measures. This has been clearly demonstrated empirically using the popular index of dissimilarity D. Nevertheless, it is not clear if this relationship between the size of areal unit and the changing direction of segregation level will hold when spatial measures of segregation are employed in the analysis. These spatial measures were proposed to overcome a major limitation shared among many traditional measures. If populations in different areal units of the study area are swapped (i.e., a change in the spatial relation of the population), most traditional measures of segregation cannot reflect this geographical change and return the same level of segregation. Beyond what has been examined in terms of the MAUP effects on D in two-group situations only (Wong et al., 1999), the purpose of this paper is to evaluate how sensitive spatial segregation measures are to the scale effect in both two-group and multigroup situations. In addition, the paper will explore if spatial and traditional measures will provide similar results under the influence of scale effect. These are the major objectives of this paper.

Spatial Segregation Measures

It has been well documented that many traditional measures of segregation, including the popular index of dissimilarity D, are incapable of handling the so-called checkerboard problem. A set of spatial measures has been introduced to address this limitation of D.

Based upon Newby's (1982) conceptualization of spatial segregation, Morrill (1991) suggested to modify D with a term to compare the ethnic mixes of neighboring units. This measure is denoted at D(adj). Wong (1993) modified Morrill's D(adj) by incorporating the length of shared boundaries to derive D(w). Subsequently, by incorporating the shape factor or the compactness measure, the D(s) index was introduced (Wong, 1993). For multigroup comparisons, White (1986) favored the entropy-based diversity index due to several statistical properties. Reardon and Firebaugh (2002) recently reaffirmed this claim. However, Morgan (1975) and later Sakoda (1981) derived the D(m) index, the multigroup version of D. A spatial version of D(m), denoted as SD(m), has also been suggested (Wong, 1998). Finally, a spatial measure S, which is based upon the concept of using deviational ellipses to capture the distribution characteristics of multiple groups, was introduced (Wong, 1999). Interested readers can refer to the original literature or to a more recent review of these spatial measures (Wong, 2002). This paper will not review these measures in detail.

The rest of this paper examines the question of whether or not the scale effect is more prominent in these spatial segregation measures than in the most commonly used D index and its multigroup counterpart. To accomplish this task, 1990 census data of 30 U.S. metropolitan areas tabulated at the census tract and block group levels are used. This set of metropolitan areas has been investigated in several studies (Massey and Denton, 1988; Wong et al., 1999). The comparisons between spatial and traditional measures are organized into two-group and multigroup comparisons.

EVALUATING SCALE EFFECT ON TWO-GROUP MEASURES

The Dissimilarity Index

Besides the popular index of dissimilarity D, spatial measures that were used for two-group comparisons include the Morrill's D(adj), the boundary-adjusted D(w), and the D(s), which has incorporated shape-compactness measures. Each of these indices was computed for the 30 selected metropolitan areas at the block group level and the census tract level. Whites and Blacks were the selected population groups for the two-group comparisons.

We focus on the scale effect on the D index first. Table 1 shows all the D values of the 30 metropolitan areas at the two census levels and their differences. The left column of metropolitan listing includes the states in which the metropolitan areas are found in order to specify which metropolitan areas were used because some metropolitan names are found in more than one state. In those cases, one of the states where the metropolitan area is primarily found is used.

The list of metropolitan areas in Table 1 is arranged in ascending order according to their differences in D values computed at the two census levels. Several observations from Table 1 deserve our attention. First, as indicated by existing literature that using smaller areal units most likely will yield a higher D value (Wong, 1997), all D values at the block group level among the 30 metropolitan areas in Table 1 were higher than their counterparts at the census tract level. Even their mean at the block group level is higher than that at the tract level. The higher segregation values at the block group level across all metropolitan areas reflect the nature of D that it is a measure of evenness such that

TABLE 1. D VALUES FOR THE 30 METROPOLITAN AREAS AT THE CENSUS TRACT AND BLOCK GROUP LEVELS, AND DIFFERENCES BETWEEN THE TWO LEVELS

Name	D(Tract)	D(BG)	DIFF(BG – Tract)
Detroit, MI	0.8749	0.8880	0.0130
Cleveland, OH	0.8494	0.8649	0.0155
New York, NY	0.7599	0.7770	0.0171
Gary, IN	0.8845	0.9020	0.0175
Chicago, IL	0.8458	0.8637	0.0179
Milwaukee, WI	0.8250	0.8431	0.0181
District of Columbia	0.7670	0.7886	0.0216
Los Angeles, CA	0.6922	0.7153	0.0231
Newark, NJ	0.7978	0.8238	0.0260
Buffalo, NY	0.8136	0.8460	0.0324
Atlanta, GA	0.6759	0.7115	0.0356
Miami, FL	0.7084	0.7442	0.0358
St. Louis, MO	0.7689	0.8059	0.0370
Kansas City, KS	0.7219	0.7589	0.0371
Pittsburgh, PA	0.7096	0.7479	0.0382
Baltimore, MD	0.7117	0.7503	0.0386
Philadelphia, PA	0.7671	0.8063	0.0392
Columbus, OH	0.6724	0.7118	0.0395
Cincinnati, OH	0.7569	0.7991	0.0422
Indianapolis, IN	0.7419	0.7856	0.0437
New Orleans, LA	0.6859	0.7297	0.0438
San Francisco, CA	0.6106	0.6568	0.0462
Boston, MA	0.6868	0.7345	0.0477
Houston, TX	0.6433	0.6938	0.0505
Greensboro, NC	0.6088	0.6642	0.0554
Dallas, TX	0.6138	0.6701	0.0563
Memphis, TN	0.6920	0.7506	0.0586
Tampa, FL	0.6847	0.7440	0.0593
Birmingham, AL	0.7166	0.7789	0.0622
Norfolk, VA	0.4923	0.5573	0.0650
Averages	0.7260	0.7638	0.0378

smaller areal units such as the block groups are likely to have a more uniform racial mix within each unit than that in the census tracts. Then each unit is more likely to be occupied exclusively by one population group, and thus the region will have a higher segregation value. However, the differences in D values between the two census levels are not uniform across all areas. As Table 1 shows that the differences range from the smallest of 0.013 in Detroit to the largest of 0.065 in Norfolk. This variation across different metropolitan areas is potentially attributable to the local population distribution patterns and how the census enumeration units are delineated locally.

D(Tract)		D(BG)
Norfolk		Norfolk
Greensboro		San Francisco
San Francisco		Greensboro
Dallas		Dallas
Houston		Houston
Columbus		Atlanta
Atlanta		Columbus
Tampa		**Los Angeles**
New Orleans		New Orleans
Boston		Boston
Memphis		Tampa
Los Angeles		Miami
Miami		Pittsburgh
Pittsburgh		Baltimore
Baltimore		Memphis
Birmingham		Kansas City
Kansas City		**New York**
Indianapolis		Birmingham
Cincinnati		Indianapolis
New York		District of Columbia
District of Columbia		**Cincinnati**
Philadelphia		St Louis
St Louis		Philadelphia
Newark		Newark
Buffalo		Milwaukee
Milwaukee		Buffalo
Chicago		Chicago
Cleveland		Cleveland
Detroit		Detroit
Gary		Gary
	Spearman's	
	0.9840	

Fig. 1. The 30 metropolitan areas according to their D values at the two census levels, in ascending order.

To examine if the two census levels will yield different results based upon D, one may compute the correlation coefficient of the two series of D values obtained from the block group and census tract data to examine how close they correspond to each other. However, the relative levels of segregation among metropolitan areas are probably of greater concern than the absolute magnitudes of variation among the measures. Therefore, the metropolitan areas were ranked according to D derived at the block group level and the census tract level. Then the Spearman's rank correlation coefficient was computed for the two D-series. The level of correlation, 0.98, was extremely high, implying that the rankings of the metropolitan areas according to the D values at the two census levels were basically not different.

Relying only on the Spearman's rank coefficient to draw a conclusion may be premature. Figure 1 displays the rankings of the metropolitan areas in ascending order according to the D values obtained at the two census levels. It is true that in general, ranks of many metropolitan areas did not shift between census levels, especially those at the lower range (e.g., Norfolk, Dallas, Houston) and the higher range (e.g., Gary, Detroit, Cleveland, Chicago) in D values. However, 20 out of the 30 metropolitan areas did not maintain at the same ranks or similar ranks when the analysis shifted from one census level to the other. In Figure 1, arrows are used to indicate those metropolitan areas that their ranks have shifted more than one rank from the census tract to block group level. Areas such as Los Angeles, Tampa, Memphis and New York experienced the largest shifts in their ranks, followed by Cincinnati and Birmingham. These results indicate that when levels of

TABLE 2. SUMMARY STATISTICS OF D AND ITS SPATIAL VERSIONS AT THE TRACT AND BLOCK GROUP LEVELS

	Tract averages	CV_tract	BG averages	CV_BG	Average differences
D	0.7260	0.1204	0.7638	0.0992	0.0378
D(adj)	0.6060	0.1806	0.6610	0.1490	0.0550
D(w)	0.6178	0.1762	0.6697	0.1452	0.0520
D(s)	0.7131	0.1224	0.7550	0.0992	0.0419

segregation among places are compared, one cannot assume that their relative positions will be maintained when different levels of census data are used. The conclusions arrived at one census level of geography may not be consistent with the conclusions derived from another level.

This finding has significant implications in public policy analysis. Quite often, public policy formulations, including those related to racial and ethnic issues, are based upon analyses using data at one geographical level. Results applicable to one geographical level likely do not hold in other geographical levels. This problem falls into the classical ecological fallacy problem. In addition, the patterns observed in the above analysis were based upon the comparison of the Whites and Blacks. If two other groups were chosen in the comparison, such as Blacks and Asian, some of the relationships discovered above, except that smaller areal units always yield a higher level of segregation level, may no longer hold. Given this unpredictable nature of the scale effect, Fotheringham (1989) called for analysis using multiple-level data to assess the range of possible results.

Two-Group Spatial Indices

Using an approach similar to the one above to compare the values of D obtained from different census levels, values of different two-group spatial segregation measures for the 30 metropolitan areas from the two census levels are compared. Table 2 reports the summary statistics of two-group spatial measures at the two census levels. Among all spatial two-group measures, values at the block group level are usually higher than that at the tract level. This is a pattern consistent with the aspatial measure D. However, the variation of any given segregation measure, spatial or aspatial alike, is higher at the tract level than that at the block group level based upon the coefficients of variation (CV_Tract and CV_BG in Table 2). The variations of D(adj) and D(w) among the 30 areas, given a census level, are higher than that found in the other two-group measures. As far as the absolute difference between census levels are concerned, the aspatial D has the smallest average difference, followed by D(s), the boundary-adjusted D(w), and D(adj).

These summary statistics comparing segregation measures at the two census levels indicate several interesting aspects. Similar to the traditional measure D, the spatial measures of segregation are not immune to scale effect of the MAUP. In fact, the absolute differences in the levels of segregation are larger for the spatial measures than for the aspatial measure D. Spatial measures among the 30 areas also have a larger degree of variability than that found in D. In other words, we should expect a larger degree of inconsistency in results between census levels when spatial measures are used.

Similar to the comparison of D at the two census levels, the relative levels of segregation reflected by spatial measures should be evaluated. Figure 2 shows the rankings of the 30 metropolitan areas according to the three two-group spatial measures in ascending order. Their Spearman's rank correlation coefficients are also reported. According to the Spearman coefficients, spatial segregation levels of the 30 areas at the two census levels have relatively high levels of rank correlation. The correlation ranges narrowly from 0.94 for D(s) to 0.98 for D(w). However, these correlation coefficients are not quite effective to show qualitatively the degrees of correspondence of the 30 areas between the two census levels according to the spatial measures. Using arrows again to highlight those metropolitan areas that experienced more than a minor shift in their rankings between the two census levels, areas sensitive to scale change according to the aspatial D index are also quite sensitive according to the spatial indices. In the case of D(adj), these sensitive areas include Memphis, Pittsburgh, Los Angeles, Cincinnati, and New York. Those with smaller shifts are Miami, Columbus, Chicago and Gary. When D(w) was used, Memphis, Miami and Gary dropped out from the list of most changes, but Milwaukee was added. When D(s) is used, those areas with lowest and highest segregation levels are less sensitive to scale changes. Nevertheless, the areas at the middle ranges of D(s) are more sensitive to scale changes, and they are shaded in Figure 2.

In short, it is well known that the traditional D index is sensitive to scale changes as demonstrated here again. The unexpected findings from the above exercise are that spatial segregation measures are not just subject to scale effect, but they are even more sensitive to scale changes than the aspatial measure. Therefore, researchers should expect that there should be more inconsistent results when using spatial segregation measures on multiple scale data than using the aspatial D index. In other words, the scale effect is more prominent when the two-group spatial segregation measures are used as compared to the traditional dissimilarity index.

EVALUATING MULTIGROUP MEASURES

Morgan (1975) and subsequently, Sakoda (1981) suggested the multigroup version of D, D(m). A spatial version of D(m), SD(m), was introduced later (Wong, 1998). Another spatial measure, S, based upon the concept of spatial congruence among different population groups was suggested quite recently. It is computed by overlaying ellipses representing the spatial distributions of population groups. Values of these three multigroup segregation measures were derived for the 30 metropolitan areas at both the census tract and block group levels. The 1990 Census population counts of White, Black and Asian (Asian and Pacific Islanders) were used in the computation. Summary statistics of the three multigroup measures at the tract and block group levels are reported in Table 3.

As expected, the average levels of segregation as indicated by the three measures increase slightly from the tract to block group level. In terms of variations of the three measures among the 30 metropolitan areas, the block group level data produce smaller variations. An observation that deserves attention is that the variations (the coefficients of variation) among these metropolitan areas are larger for the spatial measures than that for the aspatial measure, despite the level of census data employed. The ellipse-based

D(adj)(tract)		D(adj)(BG)	D(w)(Tract)		D(w)(BG)	D(s)(Tract)		D(s)(BG)
Norfolk		Norfolk	Norfolk		Norfolk	Norfolk		Norfolk
Greensboro		Greensboro	Greensboro		Greensboro	Greensboro		San Francisco
New Orleans		New Orleans	New Orleans		Dallas	San Francisco		Greensboro
Memphis		Dallas	Dallas		New Orleans	Dallas		Dallas
Dallas		Houston	Houston		Houston	Columbus		Houston
Houston		San Francisco	Memphis		San Francisco	Houston		Atlanta
San Francisco		Memphis	San Francisco		Memphis	Memphis		Columbus
Atlanta		Atlanta	Atlanta		Atlanta	Birmingham		**Los Angeles**
Miami		Birmingham	Birmingham		Birmingham	Atlanta		New Orleans
Birmingham		Columbus	Miami		Columbus	New Orleans		Boston
Baltimore		Miami	Baltimore		Miami	Tampa		Memphis
Columbus		Baltimore	Columbus		Baltimore	Boston		Kansas City
Tampa		**Los Angeles**	Tampa		**Los Angeles**	**Los Angeles**		Miami
Pittsburgh		Tampa	Pittsburgh		Tampa	District of Columbia		Tampa
Kansas City		**New York**	**Los Angeles**		**New York**	Miami		Pittsburgh
Los Angeles		Kansas City	Kansas City		Kansas City	Indianapolis		Baltimore
Cincinnati		Pittsburgh	Boston		Pittsburgh	Pittsburgh		District of Columbia
Boston		District of Columbia	**Cincinnati**		Boston	Baltimore		Indianapolis
District of Columbia		Boston	District of Columbia		District of Columbia	Cincinnati		Birmingham
Indianapolis		Indianapolis	Indianapolis		Indianapolis	Kansas City		Cincinnati
New York		**Cincinnati**	**New York**		Philadelphia	New York		New York
Philadelphia		Philadelphia	Philadelphia		**Cincinnati**	Philadelphia		Philadelphia
St Louis		St Louis	St Louis		St Louis	St Louis		St Louis
Newark		Newark	Newark		Newark	Buffalo		Buffalo
Buffalo		Buffalo	Buffalo		Buffalo	Newark		Newark
Chicago		Cleveland	Chicago		Cleveland	Milwaukee		Milwaukee
Gary		Milwaukee	Cleveland		Milwaukee	Chicago		Chicago
Cleveland		Chicago	Gary		Chicago	Cleveland		Cleveland
Milwaukee		Gary	Milwaukee		Gary	Detroit		Detroit
Detroit		Detroit	Detroit		Detroit	Gary		Gary
	Spearman's			Spearman's			Spearman's	
	0.9751			0.9804			0.9373	

Fig. 2. The 30 metropolitan areas in ascending order according to the two-group spatial measures at the tract and block group levels. Areas at the middle ranges of D(s) are more sensitive to scale changes and are shaded.

TABLE 3. SUMMARY STATISTICS OF MULTIGROUP SEGREGATION MEASURES AT THE TRACT AND BLOCK GROUP LEVELS

	Tract averages	CV_tract	Min	Max	BG averages	CV_BG	Min	Max
D(m)	0.7166	0.1242	0.4876	0.8774	0.7212	0.1242	0.5150	0.8660
SD(m)	0.6254	0.1590	0.3781	0.7994	0.6418	0.1552	0.4070	0.8211
S	0.7312	0.1861	0.4912	0.9592	0.7708	0.1582	0.5440	0.9661

measure especially has the greatest variation among the three. At the block group level, Greensboro has the lowest S (0.554) and Chicago has the highest (0.966). At the tract level, Houston has the lowest value (0.491) and Washington, DC has the highest value (0.959).

The 30 metropolitan areas were ranked according to the three measures (Fig. 3). A Spearman's rank correlation coefficient was calculated for each pair of the same measure for the set of metropolitan areas at the tract and block group levels. The rankings of the areas by the multigroup version of the dissimilarity index D(m) have the highest coefficient (0.931), while the rankings by the ellipse-based measure have the lowest correlation among the three, but still at a relative high level (0.883). Similar to the two-group index comparison, the Spearman's rank coefficient possibly conceals the major shifts in rankings for some observations. In Figure 3, arrows are used to indicate these major shifts in rankings.

For the aspatial D(m), the ranks of large metropolitan areas, such as Los Angeles, Boston and New York, were highly sensitive to scale changes. Smaller areas, such as Memphis, Kansas City, Birmingham, Cincinnati and St. Louis, were less sensitive, but still far from being immune to the scale effect. An interesting observation is that most of the areas with relatively high levels of segregation, such as Gary and Detroit, were rather insensitive to scale changes according to the aspatial index.

The doubt that Spearman's coefficient is not an effective indicator of scale effect becomes more apparent when the rankings of the metropolitan areas according to the multigroup spatial measures were compared. According to SD(m), 13 out of 30 areas have ranks shifts greater than one. Not only have large metropolitan areas experienced major changes in their ranks, even smaller metropolitan areas have the same experience. Also, it seems that the scale effect does not discriminate between areas with relatively low segregation levels (e.g., Greensboro and New Orleans) and those with relatively high levels (e.g., Gary and Chicago). Therefore, scale effect is more pervasive when using the spatial multi-group index SD(M) than using the aspatial multi group index D(m). Again, the multi group spatial measure is more sensitive to scale effect as in the case of two-group indices.

On the contrary, most areas with relatively high levels of segregation according to S remain at similar ranks when the analysis shifted from tract to block group level. Most of the shifts in ranking happened among those with lower segregation levels while some of the largest metropolitan areas (e.g., New York and Los Angeles) were not among them.

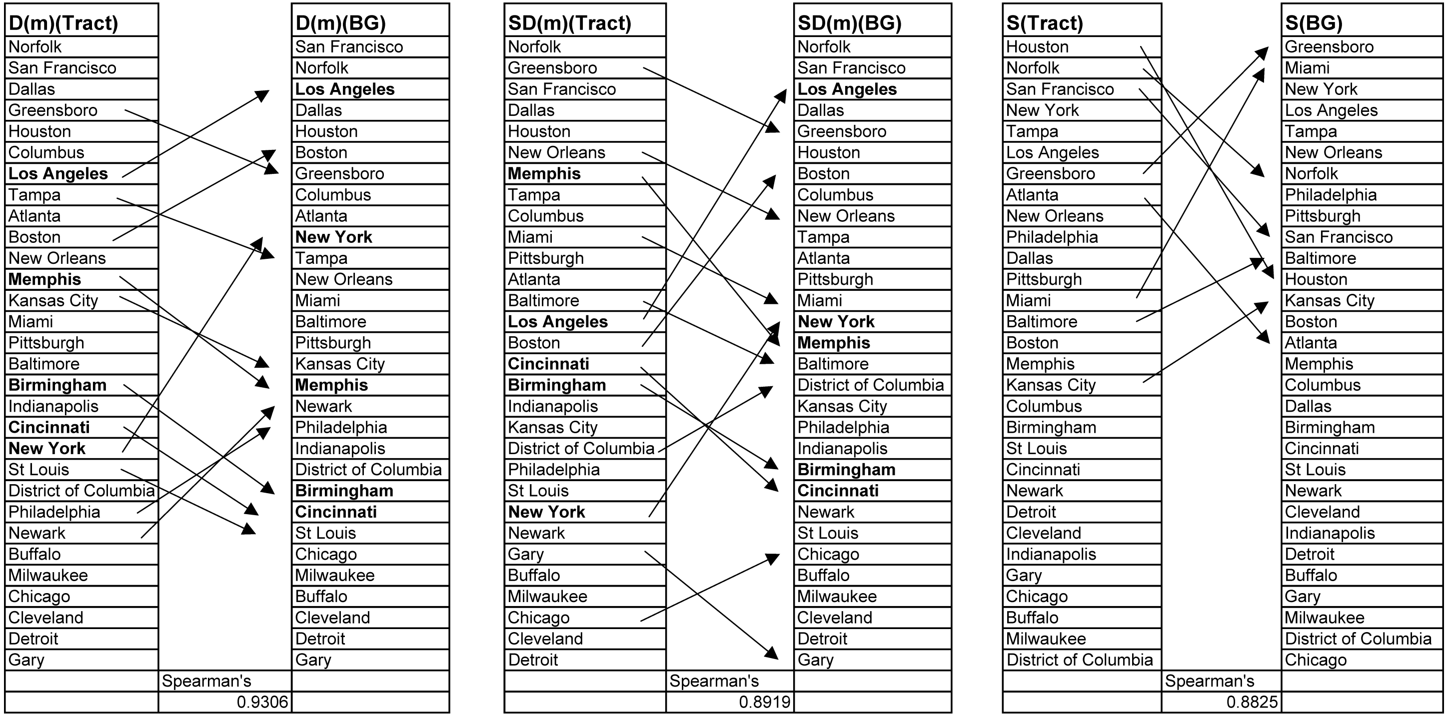


Fig. 3. The 30 metropolitan areas in ascending order according to the multigroup measures at the tract and block group level.

Apparently, S is quite different from SD(m) in its response to changing scale, at least in those extreme cases. This specific reaction of S to scale effect reflects the nature of the S index, which is a global summary measure that fails to account for detailed changes at the local scale. As the analysis shifted from block group level to census tract level, the overall regional pattern of segregation is maintained even though the local distribution patterns may have changed. The S index is effective in capturing the macro spatial pattern, but fails to detect local changes caused by changing scale.

CONCLUSION AND DISCUSSION

Using 30 metropolitan areas as the sample, this paper explores the spatial scale sensitivity of both aspatial and spatial measures of segregation in both two-group and multigroup comparisons. Based upon previous studies, aspatial measures yielded higher levels of segregation when data gathered for smaller areal units were used. This paper shows that this scale effect is also pervasive among spatial measures for both two-group and multigroup comparisons. In addition, variations in the level of segregation among the 30 metropolitan areas are larger for the spatial measures than that for the aspatial measures. When the rankings of the 30 metropolitan areas at the tract and block group levels were compared, their rank correlation coefficients were usually very high for all segregation measures. But when individual metropolitan areas were examined, major shifts in their rankings were revealed. When the two-group measures were used to compare Whites and Blacks, large metropolitan areas seemed to be more likely to shift in rank. But when multigroup indices were used to compare Whites, Blacks and Asian, both large and small metropolitan areas were equally likely subject to shift in ranks based upon the multigroup version of D. The ellipse-based multigroup spatial index has its unique responses to scale effect.

It was suggested that the spatial distribution patterns of different population groups might affect the scale effect of segregation measures, especially if population groups exhibit clustering or dispersion patterns (Wong, 1997). This paper, though not reported, did explore the relationship between scale effect and the spatial autocorrelation of population groups based upon Moran's I. Unfortunately, no correlation or pattern was detected when the level of spatial autocorrelation and scale effect were considered. For instance, Houston experienced a great shift in its rank according to the ellipse-based measure when the analysis shifted from the tract level to the block group level. But the level of spatial autocorrelation of the Black and White ratios at the block group level was moderate. It is, however, possible that spatial autocorrelation measures, such as Moran's, are not effective in capturing population distribution with a local clustering pattern. It is because spatial autocorrelation measures such as Moran's I are global measures that summarize the overall magnitude of spatial autocorrelation within the study area. Local clustering, if not widespread, cannot be revealed unless local measures are used (Anselin, 1995; Getis and Ord, 1992). Nevertheless, it will be difficult to relate the overall level of local clustering of population to the global measures of segregation focused on this paper.

With the exception of the multigroup spatial segregation index SD(m), other measures show that areas with higher levels of segregation generally seemed to be more likely to remain in similar ranks when the level of analysis shifts from one census level to the other. To explore this aspect further, we focused on the two extreme cases based upon the

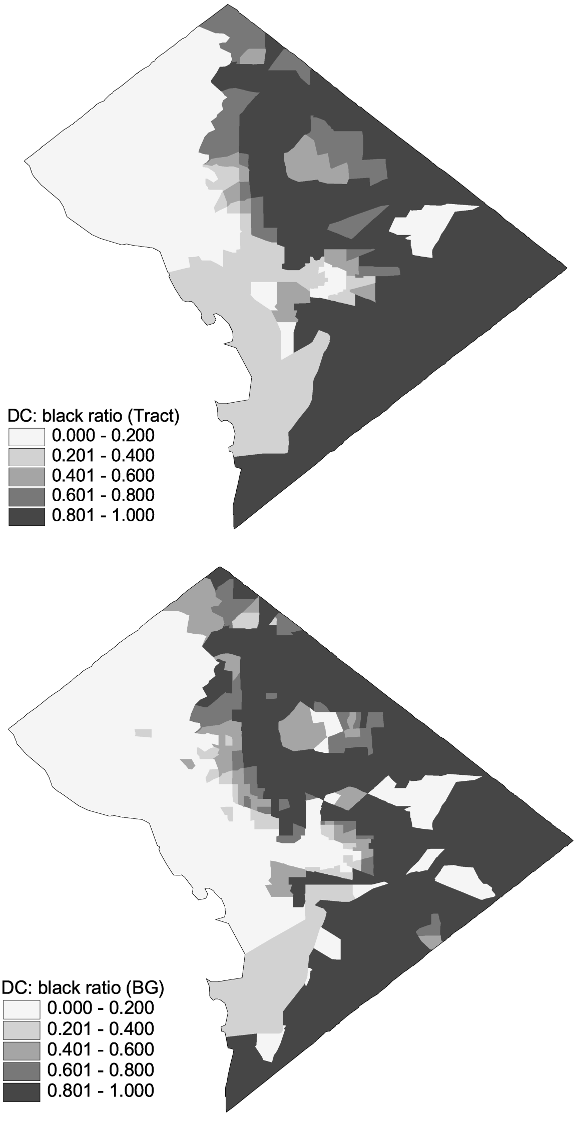
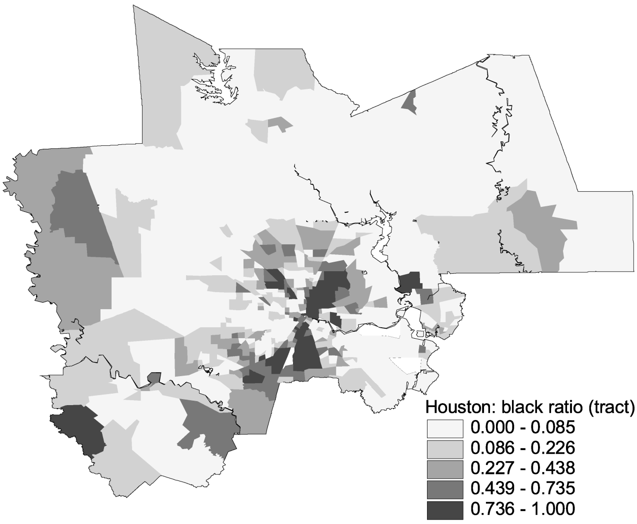


Fig. 4. Black ratios of Washington, DC at census tract and block group levels.

results from the ellipse-based measures. Houston had the lowest level of segregation and Washington, DC had the highest level of segregation based upon the tract level analysis (Fig. 3). When block group data were used, the rank of Houston shifted from the first to the twelfth, but Washington DC changed only slightly from the thirtieth to twenty-ninth. To explain the difference in scale effect on these two areas, we need to examine the distribution of the population in more detail.

Figures 4 and 5 show the ratios of Blacks in the two areas at both the census tract and block group levels. Even though, the S index included Asian in the analysis, Whites and



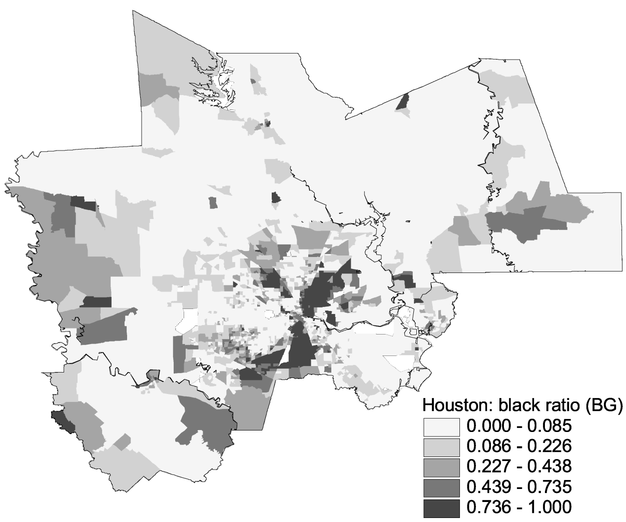


Fig. 5. Black ratios of Houston, TX at census tract and block group levels.

Blacks accounted for the majority of the population in both regions. In the Houston area, Blacks were mostly concentrated in the central area but with several intense clusters in the outskirts of the region, based upon block group level data. This distribution pattern intuitively should yield a moderate level of segregation as the Blacks were concentrated in several restricted areas but not in a large cluster. When the population data were "smoothed" at the census tract level, the clusters at the periphery of the region became less intense. This can be perceived as a dilution process, which makes the Blacks less dominant in the peripheral areas, and thus lowers the level of segregation. On the other hand, the distribution of Blacks in Washington, DC is highly concentrated in the eastern

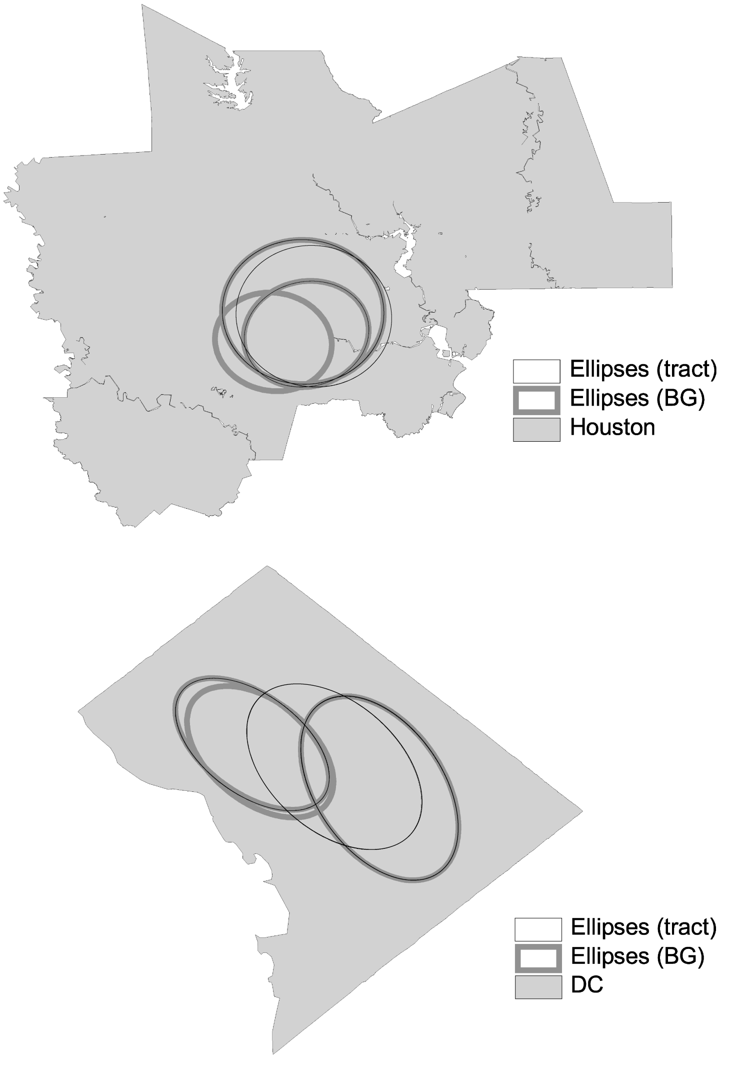


Fig. 6. Population ellipses for Houston, TX, and Washington, DC at the census tract and block group levels.

part of the city. This pattern remains more or less the same despite different levels of census data are used. When ellipses were generated to fit the distributions of the three groups at the two census levels in the two regions, the two sets of ellipses for Washington, DC overlap almost perfectly because the overall population distribution did not change between the two census levels (Fig. 6). However, because of the spatial "smoothing" at the census tract level, some ellipses exhibit shifts between the two levels in Houston, and thus the overlapping portions of the two sets of ellipses were also different. As a result, the levels of segregation at the two census levels were different.

Apparently, the impact of scale effect on segregation measures is related to the spatial distribution pattern of the population in the study area. Global scale summary measures of distribution characteristics are not effective to explain the scale effect. Population pattern at the local scale has to be taken into account in order to explain the scale effect on segregation measures, especially on those spatial measures. One may argue that different levels of segregation computed by data at different scale levels indicate that the segregation process operated at different magnitudes at different spatial scales. Then, comparing these results obtained at different scales essentially reflects how spatial scale modifies the magnitudes of segregation. In other words, the scale effect can be an indicator of how spatial scale alters spatial processes. As recommended by previous studies on the MAUP, it is desirable to conduct spatial analysis using multiple scale data to explore the scale-robustness of the results (Fotheringham, 1989; Fotheringham and Wong, 1991). In the case of segregation study, performing the analysis at multiple-scale levels can document the role of scale in modifying segregation. Because the current study indicates that spatial segregation measures can be more sensitive to scale effect than the aspatial measure D, this finding provides an additional reason to perform multiple-scale spatial segregation analysis.

REFERENCES

Anselin, L., 1995, Local indicators of spatial association—LISA. *Geographical Analysis*, Vol. 27, 93–115.

Duncan, O. D. and Duncan, B., 1955, A methodological analysis of segregation indexes. *American Sociological Review*, Vol. 20, 210–217.

Fotheringham, A. S., 1989, Scale-independent spatial analysis. In M. F. Goodchild and S. Gopal, editors, *Accuracy of Spatial Databases*. London, UK: Taylor and Francis, 221–228.

Fotheringham, A. S. and Wong, D. W. S., 1991, The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, Vol. 23, 1025–1044.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. E., 1998, Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, Vol. 30, 1905–1927.

Getis, A. and Ord, K., 1992, The analysis of spatial association by use of distance statistics. *Geographical Analysis*, Vol. 24, 189–206.

Holt, D., Steel, D. G., and Tranmer, M., 1996, Area homogeneity and the modifiable areal unit problem. *Geographical Systems*, Vol. 3, 181–200.

Jakubs, J. F., 1979, A consistent conceptual definition of the index of dissimilarity. *Geographical Analysis*, Vol. 11, 315–321.

Jakubs, J. F., 1981, A distance-based segregation index. *Journal of Socioeconomic Planning Sciences*, Vol. 15, 129–141.

King, G., 1997, *A Solution to the Ecological Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton, NJ: Princeton University Press.

Massey, D. S. and Denton, N. A., 1988, The dimensions of residential segregation. *Social Forces*, Vol. 67, 281–315.

Morgan, B. S., 1975, The segregation of socioeconomic groups in urban areas. *Urban Studies*, Vol. 12, 47–60.

Morgan, B. S., 1983, An alternative approach to the development of a distance-based measure of racial segregation. *American Journal of Sociology*, Vol. 88, 1237–1249.

Morrill, R. L., 1991, On the measure of geographical segregation. *Geography Research Forum*, Vol. 11, 25–36.

Newby, R. G., 1982, Segregation, desegregation, and racial balance: Status implications of these concepts. *The Urban Review*, Vol. 14, 17–24.

Openshaw, S. and Taylor, P. J., 1979, A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In N. Wrigley, editor, *Statistical Applications in the Spatial Sciences*. London, UK: Pion, 127–144.

Reardon, S. F. and Firebaugh, G., 2002, Measures of multigroup segregation. In R. Stolzenberg, editor, *Sociological Methodology*. Boston, MA, and Oxford, UK: Blackwell, 33–67.

Reynolds, H. and Amrhein, C. G., 1998, Some effects of spatial aggregation on multivariate regression parameters. In D. A. Griffith, C. G. Amrhein, and J-M. Huriot, editors, *Econometric Advances in Spatial Modelling and Methodology: Essays in Honour of Jean Paelinck*. Dordrecht, The Netherlands: Kluwer, 85-106.

Sakoda, J. N., 1981, A generalized index of dissimilarity. *Demography*, Vol. 18, 245–250.

Sui, D., 2000, New directions in ecological inference: An introduction. *Annals of the Association of American Geographers*, Vol. 90, 579–606.

Tobler, W. R., 1989, Frame-independent spatial analysis. In M. F. Goodchild and S. Gopal, editors, *Accuracy of Spatial Databases*. London, UK: Taylor and Francis, 115–122.

White, M. J., 1986, Segregation and diversity measures in population distribution. *Population Index*, Vol. 52, 198–221.

Wong, D. W. S., 1993, Spatial indices of segregation. *Urban Studies*, Vol. 30, 559–572.

Wong, D. W. S., 1997, Spatial dependency of segregation indices. *The Canadian Geographer*, Vol., 41, 128–136.

Wong, D. W. S., 1998, Measuring multiethnic spatial segregation. *Urban Geography*, Vol. 19, 77–87.

Wong, D. W. S., 1999, Geostatistics as measures of spatial segregation. *Urban Geography*, Vol. 20, 635–647.

Wong, D. W. S., 2001, Location-specific cumulative distribution function (LSCDF): An alternative to spatial correlation analysis. *Geographical Analysis*, Vol. 33, 76–93.

Wong, D. W. S., 2002. Implementing spatial segregation measures in GIS. (to appear in Computers, Environment and Urban Systems)

Wong, D. W. S., Lasus, H., and Falk, R. F., 1999, Exploring the variability of segregation index d with scale and zonal systems: an analysis of thirty U.S. cities. *Environment and Planning A*, Vol. 31, 507–522.