DISCUSSION: SEGREGATION INDICES AND THEIR FUNCTIONAL INPUTS

*Rick Grannis**

MULTIGROUP INDICES

Reardon and Firebaugh's (2002) advocacy of multigroup segregation indices is an important contribution to the index debate, a debate of central importance to understanding segregation. Indices guide and circumscribe all comparisons both over time and between cities, all correlations with other variables, and all classifications (Jahn, Schmid, and Schrag 1947). More importantly, they operationalize segregation theory itself and thus circumscribe all substantive understandings. Reardon and Firebaugh's (2002) review of multigroup indices highlights the fact that two-group indices have guided our thinking about segregation.

Indices reduce huge data arrays into simpler, more readily understandable numbers. Regardless of the formula one uses to reduce such arrays, the formula itself is merely a function of those arrays and only those arrays. Before choosing formulas, segregation researchers need to consider how their choice of inputs guides their theoretical development. Only when they have correctly determined which variables are appropriately included in the discussion of segregation can they proceed to develop indices and build definitions and theories.

Reardon and Firebaugh's (2002) formula-based criterion for segregation indices excluded other potentially important multigroup indices. They identified six measures that they evaluated against desirable properties of segregation indices adapted from Schwartz and Winship's (1980) and James and Taeuber's (1985) original four criteria: the principle of transfers, compositional invariance, size invariance, and organizational equivalence. These criteria, however, concern only the racial populations of each neighborhood and thus have no meaning for indices that include other variables. In fact, all of Reardon and Firebaugh's (2002) multigroup

*Cornell University and RAND

Direct all correspondence to Rick Grannis, Department of Sociology, Cornell University, Ithaca, New York 14853 (rdg25@cornell.edu).

measures can be derived from exactly the same set of data, a matrix whose rows consist of neighborhoods, whose columns consist of groups, and whose cell entries consist of the number of persons of the group represented by the column in the neighborhood represented by the row.

1. MULTIDIMENSIONAL INDICES

The distinction of considering multigroup segregation instead of dichotomous segregation is profound; it involves including new variables in the functions (one for each additional racial group) or adding new columns to the data matrix being reduced. However, the decision to include or to exclude other variables such as neighborhood contiguity, tract area, or proximity to the central city is as fundamental to the understanding of segregation data as is the choice of whether or not to restrict the scenario to a majority race and a single minority race or to allow for a multiple-race scenario. This is the issue of multidimensional inputs to segregation indices, and it reflects the multidimensional nature of spatial segregation. This is a topic that Reardon and Firebaugh (2002) do not address, and it is at least as important as their distinction between two-group and multigroup segregation.

Below, I consider some properties of multigroup segregation indices with multidimensional inputs. I show that indices can best be understood in terms of the variables they are functions of and that accounting for the location of neighborhoods with respect to each other is as important as analyzing multiple racial groups. I conclude by proposing a multigroup spatial proximity index.

2. INDICES AS FUNCTIONS

In addition to the racial population of each neighborhood, what other variables might we consider? In addition to Reardon and Firebaugh's (2002) six multigroup indices, Massey and Denton's (1988) classic survey of segregation indices identified 20 more distinct formulas used to measure aspects of segregation. We can reduce these 26 indices to a set of simple functions and illustrate them in terms of the variables they use as inputs to show how our choice of indices defines segregation and profoundly influences our understanding of this phenomenon.[1] Table 1 lists the indices

[1]An appendix, available from the author on request, discusses how analyzing functional inputs relates to factor analyses.

TABLE 1
Original Citation of Segregation Indices

Segregation Index	Citation
Absolute centralization	Massey and Denton (1988)
Absolute clustering	Massey and Denton (1988) adapted from Geary (1954) and Dacey (1968)
Absolute concentration	Massey and Denton (1988)
Atkinson	Atkinson (1970)
Correlation ratio	Bell (1954); White (1986)
Delta	Duncan, Cuzzort, and Duncan (1961) adapted from Hoover (1941)
Dissimilarity	Duncan and Duncan (1955)
Distance-decay isolation	Morgan (1983)
Entropy	Theil and Finezza (1971); Theil (1972)
Gini	Duncan and Duncan (1955)
Isolation	Bell (1954); Lieberson (1981)
Multigroup dissimilarity	Morgan (1975); Sakoda (1981)
Multigroup Gini	Reardon and Firebaugh (2002)
Multigroup information	Theil and Finezza (1971); Theil (1972)
Multigroup normalized exposure	James (1986)
Multigroup relative diversity	Goodman and Kruskal (1954); Carlson (1992)
Multigroup squared coefficient of variation	Reardon and Firebaugh (2002)
Proportion in central city	Massey and Denton (1988)
Relative centralization	Duncan and Duncan (1955)
Relative clustering	Massey and Denton (1988)
Relative concentration	Massey and Denton (1988)
Spatial proximity	White (1986)

and cites their original appearance in the literature.[2] Table 2 displays their computational formulas; Table 3 lists the indices as functions of the variables they use as inputs.[3]

Each index is a function of between two and 11 variables. Through inspection alone, some patterns are obvious and algebra makes

[2] For simplicity's sake, I use only the generalized version of the Atkinson index. The three versions of it used by Massey and Denton (1988) are easily derivable by substitution. I also use only the interaction index, and its distance-decay counterpart. The isolation index and its distance-decay counterpart are corollaries equaling unity minus the interaction indices.

[3] An appendix, available from the author on request, discusses how these variables were identified and summarizes the algebraic transformations.

TABLE 2
Computational Formulas for Segregation Indices

Segregation Index	Formula
Absolute centralization	$(\Sigma X_{i-1} A_i) - (\Sigma X_i A_{i-1})$
Absolute clustering	$\{[\Sigma(x_i/X)\Sigma(c_{ij}x_j)] - [X/n^2\Sigma\Sigma c_{ij}]\}/$ $\{[\Sigma(x_i/X)\Sigma c_{ij}t_j] - [X/n^2\Sigma\Sigma c_{ij}]\}$
Absolute concentration	$1 - \{[\Sigma(x_i a_i/X) - \Sigma_{i=1}^{n1}(t_i a_i/T_1)]/$ $[\Sigma_{i=n2}^{n}(t_i a_i/T_2) - \Sigma_{i=1}^{n1}(t_i a_i/T_1)]\}$
Atkinson	$1 - [\mathrm{P}/(1-\mathrm{P})]\lvert\Sigma[(1-p_i)^{(1-k)}p_i^k t_i/PT]\rvert^{(1/(1-k)}$
Correlation ratio	$({}_xP_x^* - P)/(1-P)$, where ${}_{\mathrm{x}}\mathrm{P}_{\mathrm{x}}^* = \Sigma[x_i/X][x_i/t_i]$
Delta	$\frac{1}{2}\Sigma\lvert[x_i/X - a_i/A]\rvert$
Dissimilarity	$\Sigma[t_i\lvert p_i - P\rvert/2TP(1-P)]$
Distance-decay isolation	$\Sigma x_i/X\Sigma K_{ij}y_j/t_j$, where $K_{ij} = c_{ij}t_j/\Sigma c_{ij}t_j$
Entropy	$\Sigma[t_i(E - E_i)/ET]$, where $E = (P)\log[1/P] + (1-P)log[1/(1-P)]$ $E_i = (p_i)\log[1/p_i] + (1-p_i)\log[1/(1-p_i)]$
Gini	$\Sigma\Sigma[t_i t_j\lvert p_i - p_j\rvert/2T^2P(1-P)]$
Isolation	$\Sigma[x_i/X][x_i/t_i]$
Multigroup dissimilarity	$(1/2TI)\Sigma\Sigma t_j\lvert\pi_{jm} - \pi_m\rvert$
Multigroup Gini	$(1/2T^2I)\Sigma\Sigma\Sigma t_i t_j\lvert\pi_{im} - \pi_{jm}\rvert$
Multigroup information	$(1/TE)\Sigma\Sigma t_j\pi_{jm}\ln[\pi_{jm}/\pi_m]$
Multigroup normalized exposure	$(1/T)\Sigma\Sigma t_j(\pi_{jm} - \pi_m)^2/(1-\pi_m)$
Multigroup relative diversity	$(1/TI)\Sigma\Sigma t_j(\pi_{jm} - \pi_m)^2$
Multigroup squared coefficient of variation	$(1/T(M\text{-}1))\Sigma\Sigma t_j(\pi_{jm} - \pi_m)^2/\pi_m$
Proportion in central city	X_{cc}/X
Relative centralization	$(\Sigma X_{i-1}Y_I) - (\Sigma X_i Y_{i-1})$
Relative clustering	$P_{xx}/P_{yy} - 1$, where $P_{xx} = \Sigma\Sigma x_i x_j c_{ij}/X^2$ $P_{yy} = \Sigma\Sigma y_i y_j c_{ij}/Y^2$
Relative concentration	$\{[\Sigma(x_i a_i/X)]/[\Sigma(y_i a_i/Y)] - 1\}/$ $\{[\Sigma_{i=1}^{n1}(t_i a_i/T_1)]/[\Sigma_{i=n2}^{n}(t_i a_i/T_2)] - 1\}$
Spatial proximity	$(XP_{xx} + YP_{yy})/TP_{tt}$, where $P_{tt} = \Sigma\Sigma t_i t_j c_{ij}/T^2$ $P_{xx} = \Sigma\Sigma x_i x_j c_{ij}/X^2$ $P_{yy} = \Sigma\Sigma y_i y_j c_{ij}/Y^2$
Notation	
a_i	Each neighborhood *i*'*s* land area
A	Urban region's land area
c_{ij}	Dichotomous variable that equals 1 when neighborhoods *i* and *j* are contiguous and 0 otherwise*

continued

TABLE 2
Continued

Segregation Index	Formula
Notation (continued)	
E	$\Sigma_{m=1}^{M} \pi_m(1 - \pi_m)$
I	$\Sigma_{m=1}^{M} \pi_m \ln[1/\pi_m]$
M	Number of groups being considered
n	Number of neighborhoods i in the urban area
p_i	Neighborhood i's minority proportion
P	Urban region's minority proportion
t_i	Neighborhood i's total population
T	Urban region's total population
x_i	Neighborhood i's minority population
X	Urban region's minority population
X_{cc}	Number of minorities living within the boundaries of the central city
y_i	Neighborhood i's majority group population
Y	Urban region's majority population
π_{jm}	Proportion in group m
π_m	Proportion in group m, of those in unit j
When Neighborhoods Are Ordered by Land Area, from Smallest to Largest	
n_1	Rank of the neighborhood where the cumulative population of neighborhoods equals X, the study area's minority population, summing from the smallest neighborhood up
n_2	Rank of the neighborhood where the cumulative population of neighborhoods equals Y, the study area's majority group population, summing from the largest neighborhood down
T_1	Cumulative population of neighborhoods 1 to n_1
T_2	Cumulative population of neighborhoods n_2 to n
When Neighborhoods Are Ordered by Increasing Distance from Central Business District	
A_i	Cumulative proportion of land area through neighborhood i
X_i	Cumulative proportion of minorities through neighborhood i
Y_i	Cumulative proportion of majority group members through neighborhood i

*Massey and Denton (1988) used $\exp(-d_{ij})$, or the negative exponential of the distance between the centroids of i and j, to estimate contiguity.

TABLE 3
Input Variables for Segregation Indices

Segregation Index	Input Variables: Constants	Vectors	Matrices	Orderings
Absolute centralization		**a, x**		O_2
Absolute clustering	n, X	**t, x**	**C**	
Absolute concentration	n_1, n_2, $T_1 T_2$, X	**a, t, x**		O_1
Atkinson	P, T	**p, t**		
Correlation ratio	P, X	**t, x**		
Delta	A, X	**a, x**		
Dissimilarity	P, T	**p, t**		
Distance-decay isolation	X	**t, x, y**	**C**	
Entropy	P, T	**p, t**		
Gini	P, T	**p, t**		
Isolation	X	**t, x**		
Multigroup dissimilarity	T, I	**t**, $\boldsymbol{\pi}$		
Multigroup Gini	T, I	**t**, $\boldsymbol{\pi}$		
Multigroup information	T, E	**t**, $\boldsymbol{\pi}$		
Multigroup normalized exposure	T	**T**, $\boldsymbol{\pi}$		
Multigroup relative diversity	T, I	**T**, $\boldsymbol{\pi}$		
Multigroup squared coefficient of variation	M, T	**t**, $\boldsymbol{\pi}$		
Proportion in central city	X, X_{cc}			
Relative centralization		**x, y**		O_2
Relative clustering	X, Y	**x, y**	**C**	
Relative concentration	n_1, n_2, $T_1 T_2$, X, Y	**a, t, x, y**		O_1
Spatial proximity	T, X, Y	**t, x, y**	**C**	

Notation
$\mathbf{a} = [a_i]$
$\mathbf{cc} = [cc_i]$
$\mathbf{t} = [t_i]$
$\boldsymbol{\pi} = [\pi_{im}] = [\pi_m]$
$\mathbf{x} = [x_i]$
$\mathbf{y} = [y_i]$
$\mathbf{C} = [c_{ij}]$
n_1 (see Table 1)
n_2 (see Table 1)
O_1 Ordering of the neighborhoods by land area, from smallest to largest
O_2 Ordering of the neighborhoods by increasing distance from the central business district

numerous other patterns apparent. Table 4 lists the indices as functions of this reduced set of input variables. Table 5 categorizes indices by their common input variables and describes their inputs more fully.

TABLE 4
Reduced Set of Input Variables for Segregation Indices

	Input Variables		
Segregation Index	Vectors	Matrices	Orderings
Absolute centralization	$\mathbf{a}, \mathbf{x}$		O_2
Absolute clustering	$\mathbf{x}, \mathbf{y}$	$\mathbf{C}$	
Absolute concentration	$\mathbf{a}, \mathbf{x}$, y		
Atkinson	$\mathbf{x}, \mathbf{y}$		
Correlation ratio	$\mathbf{x}, \mathbf{y}$		
Delta	$\mathbf{a}, \mathbf{x}$		
Dissimilarity	$\mathbf{x}, \mathbf{y}$		
Distance-decay isolation	$\mathbf{x}, \mathbf{y}$	$\mathbf{C}$	
Entropy	$\mathbf{x}, \mathbf{y}$		
Gini	$\mathbf{x}, \mathbf{y}$		
Isolation	$\mathbf{x}, \mathbf{y}$		
Multigroup dissimilarity	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Multigroup Gini	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Multigroup information	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Multigroup normalized exposure	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Multigroup relative diversity	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Multigroup squared coefficient of variation	$\mathbf{t}_1, \mathbf{t}_2, \ldots$		
Proportion in central city	$\mathbf{cc}, \mathbf{x}$		
Relative centralization	$\mathbf{x}, \mathbf{y}$		O_2
Relative clustering	$\mathbf{x}, \mathbf{y}$	$\mathbf{C}$	
Relative concentration	$\mathbf{a}, \mathbf{x}, \mathbf{y}$		
Spatial proximity	$\mathbf{x}, \mathbf{y}$	$\mathbf{C}$	

This new list represents those variables, and only those variables, that a segregation researcher would have to find values for in order to calculate these indices.[4]

Some indices (dissimilarity, Gini, entropy, Atkinson, isolation, and the correlation ratio) are functions of the minority population and the majority population of each neighborhood, and only those variables. These indices are more widely used than any other segregation measures and many of

[4] For example, one counts the number of minorities in a neighborhood or the number of majority group members in a neighborhood, but one does not count a proportion. One calculates a proportion from counted values. Similarly, the sums of neighborhood populations are just that, sums. One had to identify neighborhoods and count their populations.

TABLE 5
Segregation Indices Grouped by Shared Variables

Input Variables	Indices
Minority and majority populations of individual neighborhooods (x, y)	Atkinson Correlation ratio Dissimilarity Entropy Gini Isolation
Racial populations of individual neighborhoods $(t_1, t_2, \ldots)$	Multigroup dissimilarity Multigroup Gini Multigroup information Multigroup squared coefficient of variation Multigroup relative diversity Multigroup normalized exposure
Minority and majority populations of individual neighborhoods and neighborhood contiguity (C, x, y)	Absolute clustering Distance-decay isolation Relative clustering Spatial proximity
Minority and majority populations and land area of individual neighborhoods (a, x, y)	Absolute concentration Relative concentration
Minority population and land area of individual neighborhoods (a, x)	Delta
Minority population and land area of individual neighborhoods and neighborhood ordinal distance from central business district (a, O_2, x)	Absolute centralization
Minority and majority populations of individual neighborhoods and neighborhood ordinal distance from central business district (x, O_2, y)	Relative centralization
Minority population of neighborhood and whether neighborhood is in central city (cc, x)	Proportion in the central city

the debates about indices refer only to these. By considering only members of this group, most segregation researchers have implicitly argued that only two arrays of numbers—the majority population of the neighborhoods and the population of a single minority group of each neighborhood—are important for understanding segregation and that no other information is necessary. Reardon and Firebaugh's (2002) multigroup indices are extensions of these indices to the multigroup case.

3. SPATIAL PROXIMITY

Some indices (absolute and relative clustering, distance-decay isolation, and spatial proximity) are primarily concerned not with the distribution of the minority and majority populations across neighborhoods but with the distribution of minority and majority neighborhoods with respect to each other. White (1983) termed this the “checkerboard problem.” If one allows the squares on a checkerboard to represent neighborhoods, once the composition of each square is given, any spatial rearrangement of them will result in the same calculation for most segregation indices. Thus, “A city in which all the nonwhite parcels were concentrated into one single ghetto would have the same calculated segregation as a city with dispersed pockets of minority residents” (White 1983, pp. 1010–11).

These “clustering” indices can all be rewritten in terms of a single function:

$$f(x_i, y_j) = \Sigma_{i=1} \Sigma_{j=1} c_{ij} x_i y_j \tag{1}$$

This function focuses on the potential for interactions and is explicitly defined in terms of the product of the number of individuals of the specified groups in the set of neighborhoods defined by the contiguity matrix **C**. All of these indices can be thought of as interaction measures, except that instead of assuming the potential for interaction exists only among residents of a single neighborhood (e.g., a census tract), they assume the potential for interaction exists between residents of neighborhood *i* and all other neighborhoods defined as contiguous to *i* by the contiguity matrix **C. C** could also be defined as a weight such as inverse distance so that neighborhoods that were closer were weighted more heavily while neighborhoods that were farther away were weighted less heavily.

The formulas for the clustering indices then translate as follows:

$$\text{Distance-decay isolation} = f(x_i/X, y_j)/f(t_j, 1) \tag{2}$$

$$\text{Absolute clustering} = \{f(x_i, x_j) - f(X^2, n^{-2})\}/\{f(x_i, t_j) - f(X^2, n^{-2})\} \tag{3}$$

$$\text{Relative clustering} = Y^2 f(x_i, x_j)/X^2 f(y_i, y_j) - 1 \tag{4}$$

$$\text{Spatial proximity} = \{f(x_i, x_j)/X + f(y_i, y_j)/Y\}/\{f(t_i, t_j)/T\} \tag{5}$$

The distance-decay isolation index can be interpreted as the probability that the next person a group *X* member meets is from group *Y*. The absolute clustering index measures how much group *x* members are clustered so that they interact with each other more than one would expect as a proportion of how much opportunity group *x* members have to interact with anyone more than one would expect. The relative clustering index compares the average proximity of members of group *X* to the average proximity of members of group *Y*. The spatial proximity index is the average of intragroup proximities weighted by each group's fraction in the population. In short, the distance between groups, or the geographic level at which they are segregated, as well as the fact that they are separated, has been a concern of previous research. It would be useful to combine Reardon and Firebaugh's (2002) focus on multiple groups with this concern for the dimension of distance.

4. AN EXAMPLE

Although segregation researchers are certainly aware of the presence and importance of ghettos and the spatial patterning of neighborhoods, there is a tendency to revert to more simplistic measures when analyzing more complex relationships. I illustrate one such complexity by reexamining some data from an article by Farley et al. (1994) that began by examining "residential segregation scores" (using the dissimilarity index) for metropolitan Detroit in 1990, controlling separately for household income and for educational attainment. Since Farley et al. used only the dissimilarity index, their analysis utilized only two arrays of numbers to compute their index scores for each subset. Using their data, I computed the dissimilarity index (to replicate their results), the isolation index, and the spatial proximity index for each of their income and educational subsets. The isolation index is perhaps the second most commonly cited index and, like the dissimilarity index, uses only information about the number of white and black households in the tract. Computing the spatial proximity index required using one additional array of numbers: those tracts in the Detroit metropolitan area that were contiguous and those that were not. Figures 1 and 2 display the results.

In both figures, the line represented by long dashes and triangles represents the dissimilarity index scores, the line with short dashes and circles represents the isolation index, and the solid line with squares represents the spatial proximity index. As Farley et al. (1994) noted, the

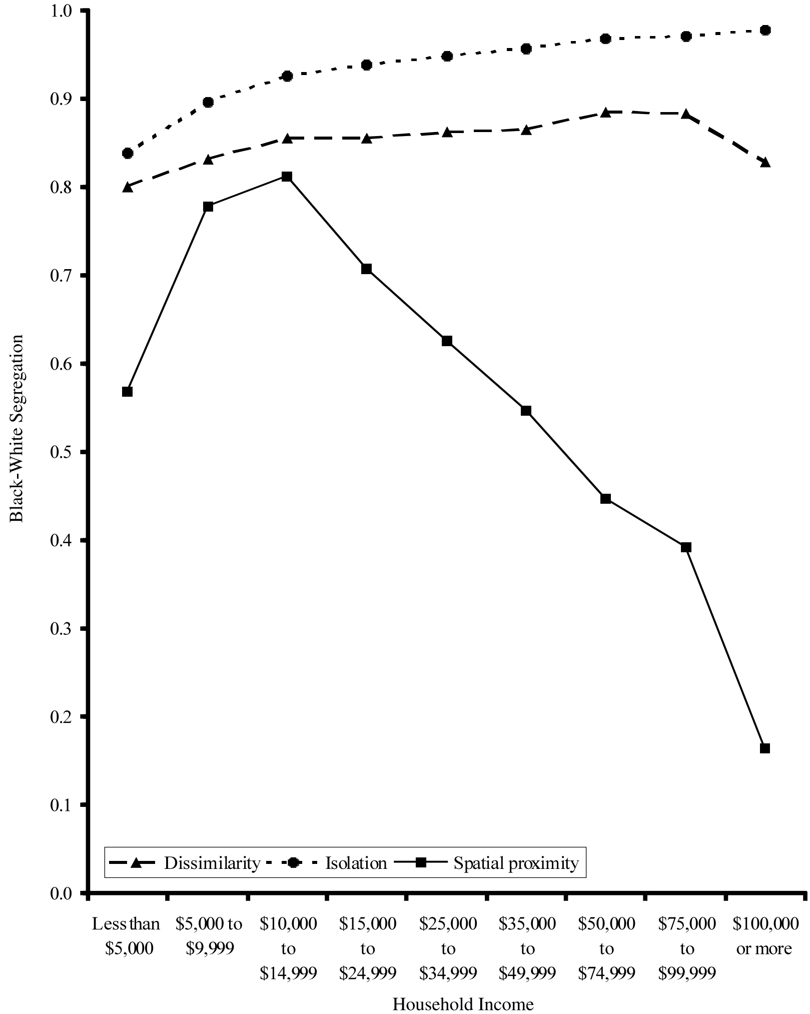

FIGURE 1. Black-white segregation in Detroit metropolitan area by income.

dissimilarity index is essentially unaffected by changes in income; the isolation index behaves similarly. Using one or both of these indices, which only use information about the numbers of white and black households in each tract, one might reasonably conclude that segregation is unaffected by changes in income or educational attainment, as did Farley et al. (1994). Segregation as measured by the spatial proximity index,

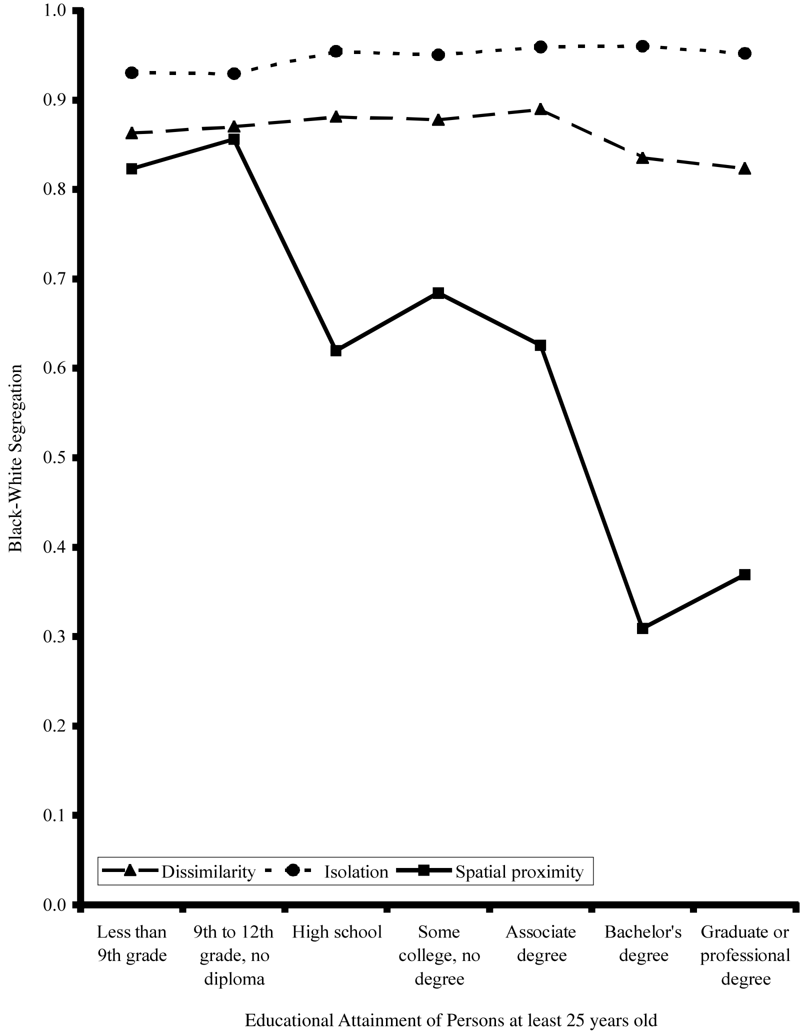

FIGURE 2. Black-white segregation in Detroit metropolitan area by educational attainment.

however, drops dramatically with increases in income or educational attainment.

Thus two findings emerge from these figures. First, in Detroit, blacks with high incomes and high education are just as likely to be segregated at the tract level as blacks with low incomes and little education (as shown by the dissimilarity or isolation indices). Second, the tracts that blacks with high incomes and high education live in are much more likely to be adjacent to white tracts than are the tracts occupied by blacks with low incomes and little education. While income and education do not allow blacks in Detroit easy access into white tracts, they do allow them access into tracts in larger white areas. Segregated neighborhoods of 1300 households (the average tract size in this region) may be sufficiently large to segregate most face-to-face interactions but may not be sufficiently large to segregate supermarkets, school districts, or churches. Upper-class blacks live in smaller "mostly black" communities than do their less well-to-do counterparts. Using indices that only account for the distribution of whites and blacks at the tract level misses this second important and powerful finding.

5. MULTIGROUP SPATIAL PROXIMITY

While Reardon and Firebaugh (2002) use "organizational units" instead of neighborhoods to maximize generality (for example, with school districts, etc.), this simplification ignores important differences between neighborhoods and schools, not the least of which is that school districts are often much bigger than the neighborhood equivalents by which we typically measure residential segregation (e.g., a high school feeder zone may include several census tracts and numerous census block groups). Much of the theorizing done about residential segregation would not have any meaning at the micro-level at which Reardon and Firebaugh's (2002) indices (as well as many traditional segregation indices) would measure it, concerning only very small communities of a few hundred or a thousand households, oblivious to larger patterns. This has not been a critical issue in the past since segregation has primarily been measured in a two-group scenario, black and white. Given the ghettoization of blacks throughout America, segregation at the neighborhood level typically implied larger-scale segregation (Massey and Denton 1993). As we begin to consider more complex, multigroup cases, however, the disparity between micro-level segregation and larger patterns may become more acute.

As we begin to consider more widespread use of multigroup measures, it is appropriate that we simultaneously consider using spatial proximity measures more. The spatial proximity index has an obvious adaptation to the multigroup case:

$$\text{Multigroup spatial proximity} = \Sigma_m\{f(x_{im}, x_{jm})/X_m\}/\{f(t_i, t_j)/T\} \quad (6)$$

$$= \Sigma_m(1/X_m)(\Sigma_i \Sigma_j x_{im} x_{jm} c_{ij})/(\Sigma\Sigma t_i t_j c_{ij}/T) \quad (7)$$

This new index is the average of intragroup proximities weighted by each group's fraction of the population. The numerator is the sum of several terms, one for each racial group. Each term is the total number of potential interactions among members of each racial group averaged over the number of members of that racial group. The denominator equals the total number of potential interactions between all individuals averaged over the total number of people in the study area. Since this index is a sum of an independent term for each racial group, this index could be used effectively to measure spatial proximity among a single racial group, two racial groups, or several. It could also be expanded to measure spatial proximity among different subcategories (e.g., income, educational level) within a single racial group or among multiple racial groups.

Researchers need to consider indices that account for nearby neighborhoods unless they are certain the boundaries of census tracts, block groups, or other neighborhood proxies completely confine the experience of segregation. Using clustering indices, such as the multigroup spatial proximity index proposed above, also allows one to deal with the Modifiable Areal Unit Problem (MAUP), which results from the imposition of artificial boundaries on a geographically correlated phenomenon (Openshaw and Taylor 1979). Analytical results may be highly sensitive to the size and boundaries of the zones used. Dramatically different results may be obtained from the same set of data when the information is grouped in different levels of spatial resolution (scale effect) or by merely altering the boundaries or configurations of the zones at a given scale of analysis (zone effect). Using the smallest possible neighborhood equivalent in combination with a clustering index would allow one to account for segregation patterns at or above the level of analysis. While the availability of data limits our analyses, we should not allow our analytic tools to be more limiting than our data.

6. CONCLUSION

In choosing the indices by which they measure segregation, sociologists define segregation itself and induce particular substantive understandings. When we use indices, we are trying to reduce huge data arrays into simpler, more readily understandable numbers. While no index will be useful for every study of segregation, the importance of measuring both neighborhood-level and larger-level segregation patterns clearly shows that indices that include distance or contiguity as an input should be used more often in future segregation research. Furthermore, the definition of contiguity (or distance), as used in these measures, needs to remain adaptable, to account not only for geographic distance or adjacency but also for the actual probabilities of contact.[5]

When researchers choose to ignore information about the racial composition of nearby neighborhoods, they are making assertive theoretical claims that segregation only occurs at (or below) the level of the neighborhood they are using (e.g., a census tract) and that life in a minority community of a few hundred or a thousand households in the midst of a majority population is not conceptually different from life in a minority community of a hundred thousand households since they fail to incorporate data that would allow them to distinguish these two situations. Just as multigroup segregation indices are important for understanding complex demographic patterns, so indices that include data on neighborhood contiguity are important for understanding segregation patterns occurring at levels larger than a census tract.

REFERENCES

Atkinson, A. B. 1970. "On the Measurement of Inequality." *American Sociological Review* 43:865–80.

Bell, Wendell. 1954. "A Probability Model for the Measurement of Ecological Segregation." *Social Forces* 32:357–64.

Carlson, Susan M. 1992. "Trends in Race/Sex Occupational Inequality: Conceptual and Measurement Issues." *Social Problems* 39:269–90.

Dacey, Michael F. 1968. "A Review on Measures of Contiguity for Two and K-Color Maps." Pps. 479–95 in *Spatial Analysis: A Reader in Statistical Geography*, edited by Brian J. L. Berry and Duane F. Marble. Englewood Cliffs, NJ: Prentice-Hall.

Duncan, Otis D., and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indices." *American Sociological Review* 20:210–17.

[5] For example, Grannis (1998) has argued that, in addition to mere geographic contiguity, it is important to note whether residential streets connect neighborhoods to each other. Therefore, one could let the contiguity or distance variables in these indices refer to residential street connectivity or distance, or any other appropriate measure of the actual potential for interaction.

Duncan, Otis D., Ray P. Cuzzort, and Beverly Duncan. 1961. *Statistical Geography: Problems in Analyzing Area Data*. Glencoe, IL: Free Press.

Farley, Reynolds, Charlotte Steeh, Maria Krysan, Tara Jackson, and Keith Reeves. 1994. "Stereotypes and Segregation: Neighborhoods in the Detroit Area." *American Journal of Sociology* 100:750–80.

Geary, R. C. 1954. "The Contiguity Ratio and Statistical Mapping." *Incorporated Statistician* 5:115–41.

Goodman, Leo, and William H. Kruskal. 1954. "Measures of Association for Cross-Classifications." *Journal of the American Statistical Association* 49: 731–64.

Grannis, Rick. 1998. "The Importance of Trivial Streets: Residential Streets and Residential Segregation." *American Journal of Sociology* 103:1530–64.

Hoover, Edgar M. 1941. "Interstate Redistribution of Population, 1850–1940." *Journal of Economic History* 1:199–205.

Jahn, Julius A., Calvin F. Schmid, and Clarence Schrag. 1947. "The Measurement of Ecological Segregation." *American Sociological Review* 103:293–303.

James, David R., and Karl E. Taeuber. 1985. "Measures of Segregation." *Sociological Methodology* 14:1–32.

James, Franklin J. 1986. "A New Generalized "Exposure-Based" Segregation Index." *Sociological Methods and Research* 14:301–16.

Lieberson, Stanley. 1981. "An Asymmetrical Approach to Segregation." Pp. 61–82 in *Ethnic Segregation in Cities*, edited by Ceri Peach, Vaughn Robinson, and Susan Smith. London: Croom Helm.

Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67:281–305.

Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University.

Morgan, Barrie S. 1975. "The Segregation of Socioeconomic Groups in Urban Areas: A Comparative Analysis." *Urban Studies* 12:47–60.

———. 1983. "An Alternate Approach to the Development of a Distance-Based Measures of Racial Segregation." *American Journal of Sociology* 88:1237–49.

Openshaw, S., and P. Taylor. 1979. "A Million or so Correlation Coefficients: Three Experiments on the Modifiable Area Unit Problem." Pp. 127–44 in *Statistical Applications in the Spatial Sciences*, edited by Neil Wrigley. London: Pion.

Reardon, Sean F., and Glenn Firebaugh. Forthcoming. "Measures of Multigroup Segregation." *Sociological Methodology*.

Sakoda, J. 1981. "A Generalized Index of Dissimilarity." *Demography* 18:245–50.

Schwartz, J., and C. Winship. 1980. "The Welfare Approach to Measuring Inequality." *Sociological Methodology* 9:1–36.

Theil, Henri. 1972. *Statistical Decomposition Analysis*. Amsterdam: North-Holland.

Theil, Henri, and Anthony J. Finezza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts." *Journal of Mathematical Sociology* 1:187–94.

White, Michael J. 1983. "The Measurement of Spatial Segregation." *American Journal of Sociology* 98:1008–19.

———. 1986. "Segregation and Diversity Measures in Population Distribution." *Population Index* 52(2):198–221.