AN ASYMMETRICAL APPROACH TO MEASURING RESIDENTIAL SEGREGATION

by

Stanley Lieberson*

Research Paper No. 115

*Professor of Sociology, University of Arizona
The Claude Bissell Distinguished Visiting Professor,
University of Toronto, 1979 - 80

The Centre for Urban and Community Studies
University of Toronto

March 1980

ISSN: 0316-0068

ISBN: 0-7727-1274-3

Abstract

The index of dissimilarity is the most widely used measure of residential segregation employed by sociologists. Since a single index value is used to describe the segregation between a pair of groups or between one group and the remainder of the population, its usage without taking the population composition into account implies a certain symmetry. A different way of viewing residential segregation is suggested here, in which the isolation of a pair of groups from one another is not assumed to be equivalent. This occurs when the relative numbers of each group is taken into account — a matter intentionally avoided with the index of dissimilarity.

This paper describes a set of basic asymmetrical measures for determining the isolation of a group from either the remainder of the population or another specific group. The segregation of blacks and whites in a number of cities in the United States is then examined for several decades. The results generated by these asymmetrical indexes are quite different from the conclusions that would otherwise be obtained through the index of dissimilarity. In turn this leads to a novel interpretation of shifts in black-white residential segregation, one which could not possibly be observed through a symmetrical measure.

The index of dissimilarity, D, has been the most widely used measure of residential segregation in sociology for nearly 25 years. Even now its main challenge is only in terms of statistical significance and biases introduced by differences in composition (witness, for example, Cortese, Falk, and Cohen, 1976; 1978). There is much to be said for this index: it is easy to compute; has a simple and clear operational meaning; ranges from 0 to 100; and avoids the influence of population composition since it is confined to purely describing differences between groups in their proportional distributions among specified subareas of a city or metropolitan area. The author has employed the measure extensively in studying residential segregation of racial and ethnic groups, language groups, and differences in other spatially based phenomena such as banking activities. To be sure, D has been used in situations where it is inappropriate, such as in settings where the data have some rank ordering to them and where it is unwise to ignore that ordering. Examples are using D to summarize differences between two groups with respect to their educational distributions, occupational distributions, or income distributions. All of these being instances where D fails to distinguish differences between adjacent categories from differences between distant categories (see Johnson, 1973; Lieberson, 1975).

The computation and interpretation of D is well-known, having been described in many sources (for example, Duncan and Duncan, 1955a; Lieberson, 1963; Taeuber and Taeuber, 1965; Peach, 1975). Consider columns 1 and 2 of Table 1, which provide hypothetical population distributions of blacks and whites in the subareas of a city in which there are no other groups. The numbers in each column are converted to percentage distributions (see columns 4 and 5). Of the 200 blacks living in the city, 60 reside in subarea A. Hence 60/200 = .30, or 30 percent. Likewise, 20/200 of blacks are found in subarea B; hence 10 percent of the black population are located therein, etc. In similar fashion, the percentage distribution for the 1,000 whites living in the city may be calculated (shown in column 5, based on numbers in column 2). By converting to percentage distributions, D intentionally ignores the absolute numbers involved in each group, but simply compares the two percentage distributions to determine how similar they are to each other. This is done by summing differences between the black and white percentages in each subarea (ignoring signs). D is one-half the sum of these differences (or alternately and ignoring rounding errors, the sum of the positive differences, or the sum of the negative differences). In the example at hand $D = \left(|30-37| + |10-29| + |0-30| + |60-4| \right) /2 = .56$. The index ranges from 0 (in which case there is no segregation because the percentage distribution is identical for each group) to 100 (which would occur in the maximum case of segregation such that blacks would be found only in subareas where whites were absent and vice versa.

The index is not only easily computed, but has the added attraction of being readily interpreted. It describes the percentage of one group or the other which would have to move if there was to be no segregation between the groups. A symmetrical variant of this, which takes composition into account in order to determine the minimum percentage of the population that would have to move, is the replacement index (Walker, Stinchombe, and McDill, 1967). D can also be used in multiple group situations, for example, to determine the segregation between pairs of occupational or ethnic groups or between a given group and the total population (even when the total consists of many different groups). A procedure also exists for avoiding the part-whole problem that occurs if group X is compared with the spatial distribution figures for the total population (a set of figures which also includes group X). 1/

(Table 1 about here)

The dominance of D is due to a justifiably influential paper by the Duncans (1955b) which, along with their empirical study of occupational segregation published in the same year (1955a), and later work, played a key role in settling what Peach (1975, p.3) has labelled as the "index war" between 1947 and 1955. A wide variety of measures had been proposed and evaluated during that period (see the literature reviewed in Duncan and Duncan, 1955b and Taeuber and Taeuber, 1965, Appendix A), but the Duncan and Duncan study (1955b) raised a number of serious objections about many of them. Of particular importance for the issue at hand is their consideration of several indexes which were affected by population composition. Turning back to Table 1 again, it can be readily shown that if there were only 100 whites in the city instead of 1,000, but if they still had the same percentage distribution among the subareas (in other words, if the number of whites in each subarea was divided by 10), then the index of dissimilarity would remain unchanged. By contrast, there were a number of indexes that would be affected by such changes in population composition since the groups' relative numbers affected the index value.

Duncan and Duncan showed that the correlations reported by Jahn, Schmid, and Schrag (1947) between intercity differences in black-white segregation measured by Gh, an index affected by composition, and three dependent variables were grossly misleading. The correlations originally found between intercity differences in Gh and variation in the tuberculosis death rates, degree of overcrowded housing, and Thorndike's "G" measure all declined sharply when composition was directly taken into account (1955, p.215). For example, the correlation between Gh and the percent of overcrowded housing in a city went from .40 to .02 after partialling for composition.

Following up on the Duncans' lead, Taeuber and Taeuber (1965, p.218) correlated various measures of segregation obtained for 188 cities in 1950 with percent black. For the three segregation indexes that use composition in their formulas, Taeuber and Taeuber found "moderately high" correlations between the indexes and composition, ranging from .69 to .79. By contrast, there was no association between composition and either D or the closely related Gini index. None of this was likely to encourage further use of indexes which take composition into account. After distinguishing between those affected and not by composition, the Duncans had been cautious about which was a more desirable property for an index (1955b, pp.216 - 217). Indeed, they then went on to suggest that no single index would be sufficient for all purposes. To my knowledge, however, they did not use composition based measures of segregation in their work. The Taeubers reached a stronger conclusion in favor of a symmetric approach towards the degree of unevenness of two groups' spatial distributions and in opposition to measures which create "distortions" due to composition (1965, p.214). Indeed, they go on to argue against any asymmetrical approach, "The basic notion of unevenness is symmetrical: if Negroes are segregated from whites, whites are equally segregated from Negroes." (p.243).

A modification by Bell (1954) in the index of isolation proposed by Shevky and Williams in their social area analyses (1959) is relevant to both developments later in this paper and to the Duncans' findings with respect to Gh. On the one hand, Bell's measure was found by Duncan and Duncan to be identical with the square of a standard statistic, <u>eta</u> (1955b, p.213). In turn, <u>eta</u> was found to have a close linkage to the Nonwhite ghetto index, Gh. On the other hand, one component of the Bell formula is the basis for the asymmetrical measures described below.

AN ASYMMETRICAL APPROACH

Despite all of these objections, I have chosen to reconsider and elaborate on one component of a composition-linked segregation measure, namely what can be called the P*-type indexes. The resurrection under consideration here is not entirely due to chance, but rather reflects some concerns I have had when using the index of dissimilarity. The apparent immunity of D from compositional influences is not always a desirable feature because the effect of a given D on other social phenomena will be influenced by the relative numbers of the groups involved. The implications of the same index value will depend on the percentage of the population in each group. In a city with relatively few blacks, for example, the probability of black interaction with whites will be much greater than in another city with a bigger black component were D to be indentical in the two cities (Lieberson, 1969, p.859). If one is examining segregation in order to consider its impact on group patterns, then the D index by itself does not tell enough. Observe that for other purposes, D could be ideal, but not in this case. In dealing with language behavior, for example, suppose one wishes to determine the influence of segregation on the propensity of immigrant groups to learn the host society's language. If group X has a higher D than group Y, say 80 and 40 respectively, but the latter is 14 percent of the population whereas X is 3 percent, how do we combine the two factors influencing language behavior? If there are enough groups, then perhaps a multiple regression might be possible, although it is not necessarily entirely clean. If there are two groups then such steps are not possible even though we still want to know about the relative isolation of groups X and Y. Moreover, multiple regressions— even when appropriate — do not come up with a parameter which combines the two factors in an operational manner such as is possible with asymmetrical measures.

Also, there are times when one would want to take into account the asymmetric quality of group interaction. Namely, the probability of a given member of group X interacting with a member of Y is not the same as the probability of a given member of group Y interacting with an X in the usual situation where the size of the two groups are different. (This difference in probabilities is similar to that developed theoretically by Blau, 1977.) For example, if 100 out of 1,000 married Protestants have Catholic mates and 100 out of 400 Catholics are married to Protestants, then 10 percent of Protestants are intermarried whereas 25 percent of Catholics are married across religious lines. The meaning and consequences for the two groups are different. If one recognizes that there is not one best measure for all possible circumstances, then clearly D is at a relative disadvantage when compared with composition based measures for some problems.

In a recently completed monograph on blacks and South-Central-Eastern Europeans in the United States, the author found it fruitful to use measures of segregation that simultaneously take into account differences in group size as well as spatial patterns (Lieberson, in press, chapter 9). By not assuming symmetry in the segregation between two groups, such as is the case with D, but by viewing X's isolation from Y as not necessarily the same as Y's isolation from X, a somewhat different perspective was gained on the patterns of segregation. As will be shown below, the average black's isolation from whites was found to increase during a period when the average white's isolation from blacks decreased. Such a conclusion is perfectly possible with an asymmetrical approach but is not likely to be reached under the bilateral assumptions inherent in D.

The aforementioned monograph was primarily concerned with the substantive results obtained with these asymmetrical indexes rather than with the methodological issues pertaining to the measurement of segregation. However, if these asymmetrical measures do provide a way of thinking about residential segregation and measuring it that could alter existing conclusions regarding the actual direction and magnitude of residential segregation in cities and its causal linkages with other social phenomena, then a direct comparison is in order. Because the index of dissimilarity has been used so widely and for such a long time, receptivity to a different measure should be minimal if only because there are so many cities and periods for which segregation has been measured through D. Indeed, unless there are significant and very clear advantages to using an alternative, the arguments for simplicity, comparability, and standardization of measurement would all weigh strongly for continuation of D. There are three goals to this paper: an exposition of the asymmetrical measures of segregation; a comparison between the proposed measures and D based on the inferences obtained with actual data; and a formal mathematical comparison between D and the asymmetrical measures.

P* Indexes

The index of isolation (I.I.) proposed by Shevky and Williams (1949) and later modified by Bell (1954) can be viewed as the quotient of two measures. "The numerator is an approximation to the probability P that the next person a random individual from group 1 will meet is also from group 1. The denominator is an approximation to the hypothetical probability P_h that the next person a random individual from group 1 will meet is also from group 1 assuming group 1 is homogeneously mixed in all the census tracts of the city." (Bell, 1954, pp.357 - 358). The numerator of this ratio, what Bell then calls P*, is of special interest here since it can provide a direct and simple way of describing the combined effects of composition and non-random residential patterns on the isolation of ethnic or racial groups from either the remainder of the population or specific subgroups. It is, of course, false to assume that interaction within subareas is a random event. First, the groups are not likely to be randomly distributed within the area. Beyond this, it is unlikely that actual contacts between people within the areas will ignore the origins of the persons involved. Nevertheless, if these facts are kept in mind, then the operational quality of the P* measure is appealing, to wit, describing for a specified group in the city the average probability of interacting with some specfied population based on the distribution of persons by subareas and the assumption that interaction is with someone in the same subarea. The measures described below are essentially an elaboration and extension of the initial P*-type index.[2/]

Let the subscripts preceding and following P* indicate, respectively, the group from whom and to whom interaction is directed. Thus, for a randomly selected member of group X in a city, the probability that someone else selected from the same residential subarea will be a member of group Y is denoted by ${}_xP^*_y$. Only in the exceptional case, where the number of X and Y in the city is identical will ${}_xP^*_y = {}_yP^*_x$; normally the isolation of the average X from Y will not be of the same magnitude as the isolation of the average Y from X. This is due, of course, to the fact that the numbers of X's and Y's in the city are taken into account along with the dissimilarity of their spatial distributions. ${}_{x}P^*_{x}$ refers to the isolation experienced by members of group X in the city, that is, for a member of group X randomly selected in the city, ${}_{x}P^*_{x}$ gives the probability that someone else chosen from the same residential area will also be a member of group X. In other words, it gives the average proportion X is of the population in each subarea weighted by the number of X's residing in each of these subareas (Farley and Taeuber, 1968, p.956). Likewise, ${}_{y}P^*_{y}$ refers to the average isolation experienced by members of group Y in the city. One can not only measure the isolation of a given group, say X, from other specific groups such as Y and Z, but also its total isolation from all others.

In its most general form, the equation for the index is:

$${}_{x}P^*_{y} = \sum_{i=1}^{n} \left(\frac{x_i}{X}\right)\left(\frac{y_i}{t_i}\right),$$

where X is the total number of members of group X in the city, x_i is the number of group X in a given subarea, y_i is the number of group Y in the subarea, and t_i is the total population of the subarea. (It is possible for y_i to refer to group X when the simple isolation of the group is being measured.)[3/]

The computation of P* is quite simple and an illustration will help the reader gain a conceptual understanding as well. Suppose one wishes to know about the probability of blacks interacting with blacks. One determines the total population in each subarea, t_i (Table 1, column 3), and then divides the number of blacks in each subarea, b_i, by this total. The resulting black proportions in each subarea

the absence of any segregation.

An important quality of P* is that the results are asymmetrical for the groups, that is , ${}_bP^*_w \neq {}_wP^*_b$, ${}_bP^*_b \neq {}_wP^*_w$. Hence, if we saw above that the average black in the city lives in a residential area where .5017 of the residents are white, this does not mean the probability for a randomly selected white having contact with blacks is also .5017. In point of fact, except for the rare case when the two groups have exactly the same number of residents in the city, it is certain that the other index will be different. In the present case, one would weight the black proportion in each subarea population (Table 1, column 6) by the number of whites (column 2) rather than the number of blacks (column 1) if the goal is the proportion of blacks living in the average white's subarea (${}_wP^*_b$ = . 1003). (For those oriented to demographic standardization methods, these differences are analogous to the effect of using different population weights for a given set of rates.) We shall make good use of this diffence in probabilities.

In the course of this generalization of P*, an exact equation was developed for determining ${}_wP^*_b$ if ${}_bP^*_w$ is known. This is:

$${}_wP^*_b = ({}_bP^*_w) \quad (B/W).$$

From this equation, which can be used to derive either index from the other as long as the relative size of B and W are known for the city, it follows that a change in one P* of a given magnitude has a different effect on the other index, the difference being a function of the ratio between W and B. Thus, if ${}_w\Delta_b$ represents a change in ${}_wP^*_b$, it follows that the change in the ${}_bP^*_w$ index, ${}_b\Delta_w$, will be function of B/W as such:

$$ {}_b\Delta_w = \frac{{}_w\Delta_b}{\frac{B}{W}} . $$

This is extremely important because it means that the greater the difference between two groups in their total numbers in the city, the greater will be the differential consequence of a change in segregation. This property has important substantive implications.

Finally, given a group's isolation index, a conversion exists for determining the other group's isolation in the simple case where only two groups are present:

$$ {}_bP^*_b = 1 - \left[\left(\frac{W}{B}\right)\left(1 - {}_wP^*_w\right)\right] . $$

There are, of course, a number of other technical issues in the computation of segregation indexes, for example, the spatial grid used, the number of spatial units, and the like, but these have been carefully considered elsewhere and there is no point rehashing the discussion at this point (see Taeuber and Taeuber, 1965; Lieberson 1963). The key issue is determining whether these P* indexes have a significant function in research on segregation that cannot be adequately met with D.

EMPIRICAL APPLICATIONS

Rather than speculate on these matters, both the D and the ${}_iP^*_i$ measures are applied to some actual data sets. The first of these sets are data initially reported in my forthcoming monograph (Lieberson, in press) that had previously been analyzed exclusively with P* indexes. In this case, the results are first examined in terms of D to see what the inferences would normally be and then considered in terms of the P* values obtained earlier.[4/]

Indexes of dissimilarity were computed between blacks and whites for 17 large non-Southern cities in the United States in each decade between 1890 and 1930.[5/] The results, summarized in Table 2, indicate that the average level of segregation at first increased only slightly and then more rapidly during the period at hand. Between 1890 and 1900, the average increased just under 2 points from 44.06 to 45.96. Indeed, there were 8 cities experiencing declines compared with 9 increases (the average increase for the latter outweighed the average decline for the former to generate this change). In each succeeding decade, an increasing number of the 17 cities experienced gains in segregation, with the average increment going up. It seems reasonable to conclude that residential segregation between blacks and whites increased during the span, possibly reflecting the growing numbers of new blacks migrating to these Northern cities. But, whatever the reason, certainly one would conclude that a jump in the average index of residential segregation from 44 to 62, a change of about 40 percent, represents increased residential segregation between the two groups.

By contrast, what do the P*-type measures tell us about the same period? The mean value of ${}_bP^*_b$ increased during this period from about .07 to .30. In 1890, the average black in these cities lived in a ward where 7 percent of the residents were black; in 1930, the average black resided in a ward in which 30 percent were black. The much more rapid increase reflects the fact that two forces were working in the same direction. Namely, the spatial dissimilarity between blacks and whites increased, as indicated by D, and the average percent blacks also increased. The results are a four-fold increase in black isolation. The change for whites is quite different however. Although the D index went up, white isolation declined slightly during the same period from .9720 to .9540. The reason is simple enough; the isolation of whites declined because the increases in spatial dissimilarity was not of a magnitude sufficient to compensate for the increase in the black proportion of the cities' populations. Hence the isolation trends (to wit, the combined product of spatial dissimilarity and population composition) are different for whites and blacks in the North in the early decades of this century. Blacks were becoming more isolated, but whites were becoming slightly less isolated even though the differences in spatial distribution were increasing.

(Table 2 about here)

There is no immediate way such a distinctive interpretation could occur if we knew changes in population composition and added that to patterns of change in D. The black average proportion in these 17 cities increased from .0295 (1890) to .0331 (1900) to .0333 (1910) to .0458 (1920) to .0648 (1930). Would that be enough to decrease the average isolation of whites in these cities even though D increased? It is hard to answer by just combining the D values with the compositional information. To be sure, one could correlate composition with D, either in static fashion or in terms of changes over time. But if one wishes to draw some inferences about the levels of isolation and/or changes in these levels over time, the P*-type provide a distinctive perspective.

Consider some actual cases. The index of dissimilarity between blacks and whites in Philadelphia increased from 47.11 to 51.03 between 1920 and 1930. During the same period the black proportion of the population also went up, from .0736 to .1126. In terms of the actual isolation of blacks, we know that the changes are in the same direction, that is, there is both greater dissimilarity in spatial distribution and a larger black proportion of the population. But what is the combined effect? On the other hand, from the white perspective the increase in the black-white D measure will be counterbalanced to some unknown degree by the decrease in the white proportion of the population. Again, it is rather hard to put these facts together except through the use of P* measures. In this case, ${}_bP^*_b$ increased from .2078 to .2727 during the decade; ${}_wP^*_w$ dropped from .9371 to .9078, even though the index of dissimilarity increased.

Use of P* indexes suggests a rather different interpretation of the events transpiring during the early part of the twentieth century in the urban North. With the movement of blacks to the urban North, the upward shifts in the index of dissimilarity are not necessarily a function of a change in white attitudes towards blacks, but rather the D increases represent nothing more than an effort to maintain the same level of white isolation from blacks at a time when the compositional shifts are working in the

extremely valuable for some problems, not the least of which is understanding the actual consequences for each group. After all, it is easier to assume that people are responding to the relative numbers of blacks and whites in their home turf rather than to the inequality of the spatial distributions between the groups. Massive differences between cities in their indexes of dissimilarity during this period are far more comprehensible when the degree of white isolation is considered (to wit, jointly taking into account composition and differences in spatial distributions). Table 4 gives information for the 17 specific Northern cities in both 1930 and 1890. In both cases, observe the substantial inter-city variation in the D obtained between blacks and whites. In 1890 the index ranges from 17.64 and 21.44 for Kansas City and Los Angeles, respectively, to the mid-60's for Chicago and Milwaukee. Likewise, there is a wide range in D for 1930, ranging from 33.64 in Minneapolis and 40 in Indianapolis to the mid-80's for Chicago and Milwaukee again. The average D between blacks and whites has gone up considerably during this period, from 44 to 62.

(Table 4 about here)

Consider, however, the conclusions generated by a separate analysis of black and white P* indexes. As is the case for D, there is considerable variation between cities in the isolation of blacks. In 1930, ${}_bP^*_b$ ranges from less than .02 for Minneapolis and San Francisco to .51 for Cleveland and .70 for Chicago. Observe by contrast the steady nature of white isolation, both with respect to changes between 1890 and 1930 (a decline in the mean $_{w}P^*_{w}$ from .97 to .95) and in terms of the rather small variation between cities in each period. In 1890, the indexes were rather uniformly high, ranging from lows of .8994 and .92 in Kansas City and Indianapolis to highs in excess of .99 in four cities. In 1930, the lows again were about .90 (for Indianapolis and Philadelphia), with $_{w}P^*_{w}$ exceeding .99 in two cities (Minneapolis and San Francisco). The enormous stability in the magnitude of white isolation, in the face of increases in D, black isolation, and the black proportion of the population, is certainly consistent with the hypothesis that at least part of the fluctuations in segregation are nothing more than the dominant white population attempting to maintain the same high level of isolation from blacks as before. The point of all this here is to illustrate the kinds of problems that are attratctive to study through an asymmetrical approach to segregation.

Another Example

One might object to the above example on the grounds that the analysis was necessarily confined to rather large spatial units, wards. It may be that the increases over time in both D and black isolation simply reflected the filling in of the ward units over time. If a group's total population in the city is smaller than the population of most of the spatial units used in the analysis, it is difficult to adequately measure segregation. Even if all members of the group are in one spatial unit, the index of dissimilarity will necessarily be less than 100 (see Lieberson, 1963, pp.36 - 37). Although these spatial units affect both the P* and D indexes, a certain reservation is healthy. Particularly when the period from 1890 to 1930 may have covered some important shifts in urban spatial structure due to changes in transportation. Another data set is available which allows one to overrule such issues. Based on data for 13 cities experiencing special censuses in the 1960's, Farley and Taeuber (1968) compared the segregation between blacks and whites in 1960 with the data for later in the decade. Of particular value for the purposes at hand, they present both the index of dissimilarity and what they call the "Homogeneity index", the latter being a P* index computed separately for both blacks and whites, ${}_bP^*_b$ and ${}_wP^*_w$ respectively.

Table 5 indicates the indexes of dissimilarity between blacks and whites, the P* indexes for each group, and the proportion of the population black. Since computations were based on the census tract distributions of the groups, the ward objection is eliminated. Here too we see the D index goes up in all but two of the cities in the 1960's. Again, the patterns are quite different for the specific racial groups. Although black isolation increases in all cities, a very mixed pattern exists for whites: increases in 4 cities, declines in 3, and no change in 6. The pattern is very much in keeping with the notion that whites are simply trying to maintain a certain constant high level of isolation from blacks. White isolation varies only narrowly in both periods, ranging from .88 to .97 in 1960 and from .89 to .97 later in the decade. By contrast, there is much greater variation between cities in D (or in ${}_bP^*_b$ for that matter). In both periods Sacramento has the lowest D, in the high 50's; Cleveland has the highest D in the first period and second highest later in the decade, 85 and 87 respectively. Does it make sense trying to understand massive differences between such cities in their D indexes when the levels of white isolation are actually so similar? Indeed, whites are slightly more isolated in Sacramento than in Cleveland.

(Table 5 about here)

To be sure, there are cases where the levels of isolation may indeed be truly different. An analysis of Southern cities, for example, indicates sharply lower white isolation than in the North throughout the 1890 - 1930 period. It is not so crucial, for the purposes at hand, to draw any definite conclusions about attempts by the dominant white population to maintain their distance from blacks in the face of the latter's increasing proportions. Of concern here is that these indexes provide a distinctive approach to segregation—one that cannot be matched adequately by dealing with spatial dissimilarity and population compositon as two separate variables.

LINKAGE BETWEEN D AND P*

It is therefore very important to consider closely the linkage between P*-type indexes and the combined effect of D and population composition. If knowledge of D and population composition would allow the investigator to determine the value of P* or even closely estimate its value, then the P* measures could be generated with existing data on D and population composition. At the extreme values of D such an exact linkage exists with P*. When D between groups I and J is 100, then ${}_iP^*_i = 1.0$ and ${}_jP^*_j = 1.0$. When $D_{ij} = 0$, then ${}_iP^*_i = I/T$ and ${}_jP^*_j = J/T$. However, the measures are not so closely linked in the situations commonly encountered when segregation is neither absent nor total. Through trial and error processes the following bounds are inferred for a given index of dissimilarity, D_{ij}:

$$\text{Maximum } {}_iP^*_i = I/T + \left[(D_{ij}/100) \quad (1 - I/T)\right]$$

$$\text{Maximum } {}_jP^*_j = J/T + \left[(D_{ij}/100) \quad (1 - J/T)\right]$$

The minimum P* for a given D, is somewhat more complicated

$$\text{Minimum } {}_iP^*_i = \frac{(.50 + D_{ij}/200)^2 \quad (I)}{\left[(.50 + D_{ij}/200) \quad (I)\right] + \left[(.50 - D_{ij}/200) \quad (T - I)\right]} + \frac{(.50 - D_{ij}/200)^2 \quad (I)}{\left[(.50 - D_{ij}/200) \quad (I)\right] + \left[(.50 + D_{ij}/200) \quad (T - I)\right]}$$

Two important conclusions follow from these equations. Holding D_{ij} constant, the range of possible values for ${}_iP^*_i$ increases as I/T gets smaller. If I/T is very large, then the range of possible values for ${}_iP^*_i$ is relatively small. Second, the range of ${}_iP^*_i$ is affected by the level of D_{ij}, holding I/T constant, but in a more complicated manner. If I/T = .5, then the range of ${}_iP^*_i$ is greatest when D is 50, becoming progressively smaller as D approaches either extreme of 100 or 0. As I/T moves towards either a smaller or larger proportion, the peak range for ${}_iP^*_i$ occurs at increasingly large values of D_{ij}. This is relatively insignificant as I/T approaches 1.0 because, as pointed out above, the range is so constrained anyway.

Table 6 converts these formulas into actual values, giving the reader some understanding of the ${}_iP^*_i$ range that can be encountered with a given index of dissimilarity and when group I is some specific proportion of the population. The range of ${}_iP^*_i$ is narrow, in both an absolute and relative sense, when I/T is large (especially when D is relatively small). Hence, when I/T is .9 and D is 10, ${}_iP^*_i$ ranges from .90 to .91. But the range is considerably wider in a two-group situation where the other group, J, is one-tenth. With an index of dissimilarity of 10 ${}_jP^*_j$ ranges from .10 to .19 (nearly double the bottom limit). A quick survey of Table 3 shows many instances with much wider ranges in situations that could easily be encountered in occupational, educational, or racial and ethnic situations in a variety of settings. A group that is 30 percent of the population and has an index of dissimilarity of 70 could range from an average level of isolation of .61 to .79; a similar index for a group amounting to 10 percent of the population could mean a ${}_iP^*_i$ anywhere from .33 to .73.

(Table 6 about here)

In short, there are certain extreme cases when the P* index of isolation will have a very narrow possible range, regardless of the level of the index of dissimilarity. But under those very conditions, there will have to be one or more other groups in the community whose members make up a small proportion of the population and hence for whom a wide range of possible P* values will exist. It is therefore almost certain that in any segregation study, the values of D, coupled with information on composition, will not permit a sufficiently exact determination of the P* indexes without a direct computation.

The range is quite substantial, given the composition and indexes of dissimilarity encountered with actual data. In the late 1960's data set for 13 cities reported in Table 5, the range in ${}_bP^*_b$ for the average city was .27, hardly a trivial matter. Since the black proportions of the population were even smaller in the 1930 data set reported in Table 4, the range of possible black isolation indexes was even greater, running from .29 and .33 in Indianapolis and Minneapolis respectively to .64 in Buffalo and .72 in Milwaukee. Hence, at the moment, the information obtained through a compositional variable and the index of dissimilarity cannot provide a substitute for measurement of the P* index (except where the formulas provided above indicate that the possible range of P* is acceptably narrow). To be sure, more needs to be done with this issue since there is no knowledge about the distributions actually occurring across the possible range of values.

CONCLUSION

What are the arguments for and against the P*-type index? The existence of a large body of existing data with D values is not a trivial matter, if only because standard data are not that common in the social sciences and there is enough difficulty comparing segregation between places and across periods as it is, without complicating the matter by introducing a different segregation measure. The second objection stems from the fact that the measure is correlated with population composition. All things being the same, differences between cities in P* indexes will be correlated with differences in population composition. Hence the investigator is apt to run into serious and sticky problems when correlating P*-type measures with composition. (Even here, it is possible to analyze the linkage with composition in different periods and places by considering the regression of P*-type indexes on composition, such as was described earlier in this paper and reported in Table 3.)

On the other hand, a P* index describes the actual isolation of a group from either all others or from a specific group in a manner which takes into account the joint influence of population composition and spatial distribution. The operational definitions possible with these measures are appealing and readily understood. Hence, by taking composition into account, it is possible with one number to pull together the two forces most likely to mean isolation; namely differences in spatial distribution and relative numbers. In so far as the investigator wishes to know this, we have seen that combining the index of dissimilarity with a separate measure of composition is an inadequate substitute in many cases. Considering the indexes of dissimilarity alone can easily be misleading since it is all too easy to assume that this measure of spatial differences in distribution is describing the actual spatial isolation of the groups. I believe, also, that the asymmetrical approach is full of interesting research and substantive possibilities. Even in some of the studies where these measure have been used, the investigators have not taken full advantage of the implications of looking at the event from the perspective of each of the two groups (even more perspectives are possible when segregation is analyzed in a multi-group setting). At the very least, it seems reasonable in many research projects to expect researchers to compute both the index of dissimilarity and P*-type measures—particularly since the same data inputs are necessary. As one moves away from describing dissimilarity in spatial patterns and gets into the question of the actual isolation of the groups, it is likely that P* will prove to be a more complete and adequate measure.

REFERENCES

BELL, Wendell (1954), "A Probability Model for the Measurement of Ecological Segregation" Social Forces, 32: 357-364.

BLAU, Peter M. (1977), Inequality and Heterogeneity. New York: Free Press.

COLEMAN, James S., KELLY, Sara D., and MOORE, John A. (1975), Trends in School Segregation, 1968-73. Washington, D.C.: The Urban Institute.

CORTESE, Charles F., FALK, R. Frank and COHEN, Jack (1976) "Further Considerations on the Methodological Analysis of Segregation Indices" American Sociological Review, 41: 630-637

CORTESE, Charles F., FALK, R. Frank and COHEN, Jack (1978) "Understanding the Standardized Index of Dissimilarity: Reply to Massey" American Sociological Review, 43: 590-592

DUNCAN, Otis Dudley and DUNCAN, Beverly (1955a), "Residential Distribution and Occupational Stratification" American Journal of Sociology, 60: 493-503.

DUNCAN, Otis Dudley and DUNCAN, Beverly (1955b), "A Methodological Analysis of Segregation Indexes" American Sociological Review, 20: 210-217.

FARLEY, Reynolds and TAEUBER, Karl E. (1968) "Population Trends and Residential Segregation since 1960" Science, 159: 953-956.

JAHN, Julius, SCHMID, Calvin F. and SCHRAG, Clarence (1947), "The Measurement of Ecological Segregation" American Sociological Review, 12: 293-303.

JOHNSON, M. (1973) "A Comment on Palmore and Whittington's Index of Similarity" Social Forces, 51: 490-492.

LIEBERSON, Stanley (1963), Ethnic Patterns in American Cities. New York: The Free Press.

LIEBERSON, Stanley (1969), "Measuring Population Diversity" American Sociological Review, 34: 850-862

LIEBERSON, Stanley (1975) "Rank-sum Comparisons between Groups" in David Heise, ed., Sociological Methodology 1976. San Francisco: Jossey-Bass.

(References cont'd)

LIEBERSON, Stanley (In Press), A Piece of the Pie: Blacks and White Immigrants Since 1880. Berkeley: University of California Press.

PEACH, Ceri (1975), Urban Social Segregation. London: Longman.

SHEVKY, Eshref and WILLIAMS, Marilyn (1949), The Social Areas of Los Angeles, Analysis and Typology. Berkeley: University of California Press.

TAEUBER, Karl E and TAEUBER, Alma F. (1965), Negroes in Cities. Chicago: Aldine.

WALKER, D., STINCHCOMBE, A. L. and McDILL, M. J. (1967), School Desegregation in Baltimore. Baltimore: Johns Hopkins Center for the Study of Social Organization of Schools.

FOOTNOTES

1. If X/T is the proportion X is of the total population, then the index of dissimilarity between X and the total population (thus including X) is divided by 1 - (X/T) to obtain the D that would occur between X and non-X.

2. Apparently unaware of the residential segregation literature, Coleman, Kelly, and Moore (1975, pp.7 - 9) independently rediscovered both the revised I.I. index and P* some 20 years later, calling them r_{ij} and s_{ij} respectively. In the same way that I.I. was found to inadequately take composition into account, one could make the same conclusion with respect to r_{ij} in the controversial study of school desegregation.

3. In the case where one computes a group's probability of interacting with fellow members of the same group, that is, where the subscripts before and after P* are identical, it is convenient to alter the true probability by in effect assuming sampling with replacement (Bell, 1954, p.358).

4. The figures for P* in some cases differ slightly from those reported in Chapter 9 of Lieberson, in press. With one small exception, this reflects rounding.

5. Nonwhites other than blacks are included with whites. This is of no consequence in nearly all of the cities, but facilitates the analysis and illustrations by avoiding a multiple group situation.

Table 1

COMPUTATION OF D and P* INDEXES

Subareas	Number: Blacks (1)	Number: Whites (2)	Number: Total (3)	Percent: Black (4)	Percent: White (5)	Black Proportion of Subarea Total (6)
A	60	370	430	30	37	.140
B	20	290	310	10	29	.065
C	0	300	300	0	30	0
D	120	40	160	60	4	.750
Sum	200	1,000	1,200	100	100	

Table 2

BLACK-WHITE SEGREGATION, 17 LEADING NORTHERN UNITED STATES CITIES, WARDS, 1890 - 1930

Year	Index of Dissimilarity		$_bP^*_b$		$_wP^*_w$	
	Mean	Number of Cities increasing over preceding year	Mean	Number of Cities increasing over preceding year	Mean	Number of Cities increasing over preceding year
1890	44.06		.0666		.9720	
1900	45.96	9	.0760	12	.9686	4
1910	49.72	11	.0972	13	.9698	10
1920	54.37	12	.1677	17	.9614	1
1930	62.14	16	.2992	16	.9540	5

For list of cities, see Table 4.

Table 3

COMPARISONS BETWEEN ADJACENT DECADES IN THE REGRESSION OF ${}_{b}P_{b}^{*}$
ON THE BLACK PROPORTION OF THE POPULATION,
17 LEADING NONSOUTHERN CITIES

Year	r	b	a
1900	.81	1.2621	.0365
1910	.87	1.8393	.0373
1910	.88	1.8887	.0343
1920	.69	2.4384	.0618
1920	.67	2.0846	.0722
1930	.63	3.3280	.1127

Note: First year shown in a pair includes all 17 cities; second row is for those cities 10 years later with a black proportion of the population falling within the range existing in the first decade of the pair.

Table 4

BLACK-WHITE SEGREGATION IN 17 SPECIFIC CITIES, BY WARDS, 1890 AND 1930

City	D 1890	D 1930	$_bP^*_b$ 1890	$_bP^*_b$ 1930	$_wP^*_w$ 1890	$_wP^*_w$ 1930
Boston	56.50	67.76	.0849	.1923	.9831	.9782
Buffalo	38.10	79.91	.0107	.2420	.9957	.9816
Chicago	63.65	84.90	.0811	.7041	.9879	.9780
Cincinnati	45.06	65.78	.0943	.4449	.9630	.9342
Cleveland	60.70	80.16	.0467	.5094	.9890	.9574
Detroit	57.02	59.52	.0553	.3115	.9840	.9429
Indianapolis	41.95	40.07	.1727	.2601	.9215	.8984
Kansas City	17.64	59.83	.1262	.3159	.8994	.9269
Los Angeles	21.44	68.20	.0329	.2562	.9752	.9759
Milwaukee	66.99	83.71	.0143	.1636	.9978	.9890
Minneapolis	40.02	33.64	.0163	.0165	.9921	.9911
Newark	34.59	46.55	.0404	.2282	.9776	.9256
New York City	42.35	64.08	.0349	.4177	.9851	.9711
Philadelphia	42.74	51.03	.1170	.2727	.9655	.9078
Pittsburgh	45.22	51.51	.0814	.2681	.9688	.9346
St. Louis	33.50	75.55	.1089	.4655	.9437	.9313
San Francisco	41.58	44.14	.0134	.0179	.9939	.9941

Table 5

BLACK-WHITE SEGREGATION IN 13 CITIES, BY TRACTS, 1960 AND LATER

City	D		$_bP^*_b$		$_wP^*_w$		Percent Black	
	1960	later	1960	later	1960	later	1960	later
Buffalo	84.5	85.1	.65	.74	.95	.95	13.2	17.0
Providence	64.2	70.3	.23	.30	.96	.94	5.4	7.4
Rochester	76.7	79.3	.44	.53	.96	.95	7.4	10.4
Cleveland	85.2	87.2	.81	.86	.92	.92	28.6	34.1
Des Moines	76.7	77.3	.35	.40	.97	.97	4.9	5.3
Evansville	76.9	80.5	.54	.61	.97	.97	6.6	6.9
Fort Wayne	79.8	79.2	.38	.52	.95	.95	7.5	10.2
Greensboro	83.8	89.1	.83	.88	.94	.96	25.8	26.7
Louisville	78.6	81.2	.68	.73	.93	.93	17.9	20.2
Memphis	79.3	83.7	.79	.86	.88	.89	37.6	42.6
Raleigh	75.0	78.0	.72	.74	.92	.93	23.4	23.4
Shreveport	82.5	85.1	.81	.85	.90	.92	33.1	34.7
Sacramento	58.2	57.2	.24	.29	.95	.94	6.5	8.3

Source: Farley and Taeuber, 1968, Tables 1 and 3.

Table 6

LIMITS IN P* INDEXES, GIVEN D AND POPULATION COMPOSITION

I/T	Δ =10 Max.	Δ =10 Min.	Δ =10 Range	Δ =30 Max.	Δ =30 Min.	Δ =30 Range	Δ =50 Max.	Δ =50 Min.	Δ =50 Range	Δ =70 Max.	Δ =70 Min.	Δ =70 Range	Δ =90 Max.	Δ =90 Min.	Δ =90 Range
.05	.14	.05	.09	.34	.07	.27	.52	.11	.41	.72	.20	.52	.91	.48	.43
.10	.19	.10	.09	.37	.13	.24	.55	.20	.35	.73	.33	.40	.91	.64	.27
.30	.37	.31	.06	.51	.35	.16	.65	.45	.20	.79	.61	.18	.93	.85	.08
.50	.55	.50	.05	.65	.54	.11	.75	.62	.13	.85	.74	.11	.95	.91	.04
.70	.73	.70	.03	.79	.72	.07	.85	.77	.08	.91	.83	.08	.97	.93	.04
.90	.91	.90	.01	.93	.90	.03	.95	.91	.04	.97	.93	.04	.99	.96	.03
.95	.96	.95	.01	.97	.95	.02	.98	.95	.03	.99	.96	.03	.995	.95	.045

REFERENCES

Bell, Wendell. 1954. "A Probability Model for the Measurement of Ecological Segregation" Social Forces. 32: 357-364.

Blau, Peter M. 1977. Inequality and Heterogeniety. New York: Free Press.

Coleman, James S., Kelly, Sara D., and Moore, John A. 1975. Trends in School Segregation, 1968-73. Washington, D.C.: The Urban Institute.

Cortese, Charles F., Falk, R. Frank and Cohen, Jack. 1976. "Further Considerations on the Methodological Analysis of Segregation Indices" American Sociological Review. 41: 630-637.

______. 1978. "Understanding the Standardized Index of Dissimilarity: Reply to Massey" American Sociological Review. 43: 590-592.

Duncan, Otis Dudley and Duncan, Beverly. 1955a. "Residential Distribution and Occupational Stratification" American Journal of Sociology. 60: 493-503.

______. 1955b. "A Methodological Analysis of Segregation Indexes" American Sociological Review. 20: 210-217.

Farley, Reynolds and Taeuber, Karl E. 1968. "Population Trends and Residential Segregation Since 1960" Science. 159: 953-956.

Jahn, Julius, Schmid, Calvin F. and Schrag, Clarence. 1947. "The Measurement of Ecological Segregation" American Sociological Review. 12: 293-303.

Johnson, M. 1973. "A Comment on Palmore and Whittington's Index of Similarity" Social Forces. 51: 490-492.

Lieberson, Stanley. 1963. Ethnic Patterns in American Cities. New York: The Free Press.

______. 1969. "Measuring Population Diversity" American Sociological Review. 34: 850-862.

______. 1975. "Rank-sum Comparisons Between Groups" in David Heise, ed., Sociological Methodology 1976. San Francisco: Jossey-Bass.

______. In Press. A Piece of the Pie: Blacks and White Immigrants Since 1880. Berkeley: University of California Press.

Peach, Ceri. 1975. Urban Social Segregation. London: Longman.

Shevky, Eshref and Williams, Marilyn. 1949. The Social Areas of Los Angeles, Analysis and Typology. Berkeley: University of California Press.

Taeuber, Karl E. and Taeuber, Alma F. 1965. Negroes in Cities. Chicago: Aldine.

Walker, D., Stinchcombe, A.L. and McDill, M.J. 1967. School Desegregation in Baltimore. Baltimore: Johns Hopkins Centre for the Study of Social Organization of Schools.