FURTHER CONSIDERATIONS ON THE METHODOLOGICAL ANALYSIS OF SEGREGATION INDICES *

CHARLES F. CORTESE, R. FRANK FALK and JACK K. COHEN
University of Denver

American Sociological Review 1976, Vol. 41 (August): 630–637

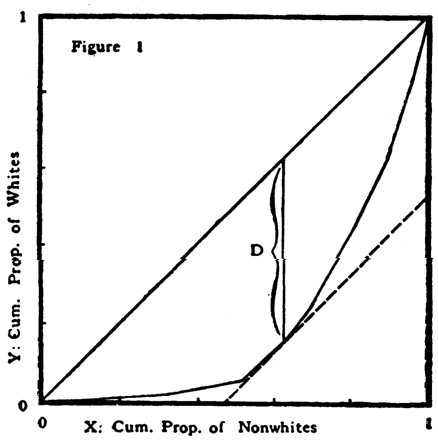
The process of developing an adequate measure of segregation occupied the literature for over a decade and culminated in the widespread use of the Index of Dissimilarity. The inadequacies of this index, although identified by the Duncans (1955), remain with us and largely have come to be ignored. This research further explores the difficulties pertaining to limitations in the use and interpretation of the Index of Dissimilarity, demonstrates some of the systematic biases resulting from these inadequacies and provides a mathematical refinement which overcomes some of the major problems inherent in the use of this index.

The concept of ecological segregation has never been dealt with adequately in definitional terms (cf. Duncan and Duncan, 1955:217). Instead of defining *segregation,* most work has considered how the opposite of segregation—often called *assimilation*—should be defined. A clear and proper definition of assimilation is especially necessary since, in fact, most attempts (including ours) to measure relative segregation are based on *a priori* "ideal" distributions. That is, the complicated details of geographical distribution and clustering are ignored in an effort to get a simple overall index of segregation. Heretofore, the "ideal" distribution has usually been taken to be the even distribution. Jahn et al. (1947) were concerned with the questions inherent in such operationalization of the concept of "ecological segregation." In that early article, they formulated their index of dissimilarity on the expected value of the central tendency of a random distribution. They stated: "This means that if there is no segregation, then members of a minority racial group will be distributed randomly throughout the various census tracts of a city. For example, if ten percent of the population of a city is Negro, then each census tract would be expected to have a Negro population of approximately ten percent" (Jahn et al., 1947:293). Their article was followed by criticism as well as the development of different measures (Hornseth, 1947; Jahn et al., 1948; Jahn, 1950; Williams, 1948; Cowgill and Cowgill, 1951) which eventually led to Duncan and Duncan (1955) demonstrating the mathematical relationships between the segregation indices previously presented.

The Duncans suggested that the Index of Dissimilarity (D) was the most useful of these indices (1955:214–5) and it has, in fact, achieved preeminence in the measurement of segregation. Taeuber and Taeuber (1965) have presented what is probably the most detailed discussion on the measurement of segregation to date. Figure 1, reproduced from Duncan and Duncan (1955), shows a geometrical definition of D as the maximum vertical distance between the "segregation curve" and the curve $Y = X$. The segregation curve is the plot of the cumulative proportion of whites versus the cumulative proportion of nonwhites, where the respective cumulations are obtained by taking the census tracts in order of increasing nonwhite percentages.

The Duncans provided a set of inadequacies and precautionary comments about the interpretation of segregation indices. Our article builds upon the Duncans' comments and further explores the "mathematical properties [of the indices] of which their proponents were unaware, and which lead to difficulties of interpretation" (Duncan and Duncan, 1955:210). The objections to D have resulted in earlier attempts

* This research was funded in part by a grant from N.S.F. through the Undergraduate Research Participation Program, Grant Number GY-9959. The position of authors' names does not imply ordinal properties.



to correct the measure (cf. Leasure and Stern, 1968), but thus far it seems the major problems are yet to be dealt with adequately.

Objections to D

The major objections to the Index of Dissimilarity follow:

1. The expectation of "evenness" as the opposite of segregation is not as useful in most cases as the concept of "randomness."
2. D is affected by differences in the proportion of the minority in the population, thus preventing intercity comparisons.
3. D is affected by the size (number of households) of the areal unit of analysis.
4. The present interpretation of **D** as the proportion of nonwhites who would have to change their tract of residence to make the distribution of the minority even throughout the city (Duncan and Duncan, 1955: 211) is misleading since it does not include the concept of replacement of the relocated minority.

Some of these objections are obviously interrelated. We will deal with each briefly.

The ideal of evenness has been used predominately as illustrated by the fact that the zero point of most of the indices discussed by the Duncans (1955), including the Index of Dissimilarity itself, occurs when the minority is distributed evenly throughout the population. Our research has focused on an examination of the random fluctuations around this expectation as significant information. It is our argument that the opposite of segregation is more appropriately a randomness of distribution rather than an hypothetical "evenness." It might be natural, therefore, to construct an index which takes the value of zero when the distribution is random. However, due to the extensive use of D, we prefer to account for randomness by computing the expected value of D under the hypothesis that the distribution of nonwhites in any census tract is random. That is, one could compute a value for D which would be expected, given that the minority population were distributed randomly throughout the entire population. More explicitly, we assume that for a census unit with T_i residents, we randomly select *without replacement*, T_i individuals from the total population of the city. To compute the expected value of D, one needs the probability that the number of nonwhites (N_i) has some particular value (n_i). The assumptions of sampling without replacement and a population divided into two categories (i.e., white and nonwhite) specifies that this probability is given by the hypergeometric distribution (see equation (2) below). It is our contention that the comparison of the actual D for a city with its expected value contains more information than D itself and can alleviate some of the limitations of D. This comparison is most sensibly carried out in units of the standard deviation of D (cf. Bardwell, 1971) under the same assumptions used for the expected value of D. To compute the standard deviation of D, one needs the joint probability that in two census units with populations T_i and T_j, the numbers of nonwhites (N_i and N_j) in those units have some particular values (n_i and n_j).

In general, anytime that more than one census unit is involved in the placement of a given total number of individuals, interdependency exists. This interdependency is accounted for by the joint probability distribution in the calculation of the standard deviation.

The joint distribution was derived by conceptually selecting T_i plus T_j individuals for the two census units and then selecting T_i of these for the i^{th} unit. This joint probability is given by the product of two hypergeometric distributions as shown in equation (3) of the appendix.

We propose that the "standard score" of D,

$$Z = \frac{D-\mu D}{\sigma D} \qquad (1)$$

is more meaningful than D itself.

An additional motivation for building upon D is that even though evenness would not occur naturally, evenness is often a social goal. Examples of this are recent policies and attempts to distribute minority teachers evenly through a school system.

The other three objections or limitations raised above have been previously pointed out but not clearly demonstrated (Duncan and Duncan, 1955:216). Jahn et al. (1947:294) stated that "a satisfactory measure of ecological segregation should . . . not be distorted by . . . the proportion of Negroes." D is affected by the proportion of nonwhites in a particular city and is further affected by the number of households per unit of analysis (Taeuber and Taeuber, 1965:231–5). We demonstrate these effects in Table 1 which provides the expected value of the Index of Dissimilarity for computer-generated random placement with varying proportions of nonwhites and under the ideal conditions of many areal units with an equal number of households in each unit.

The expected values in Table 1 shows that the random fluctuations can make a significant difference in the expected value of D. The pattern is apparent—as q increases to .5, the expected value of D decreases. This relationship, of course, reverses as q exceeds .5. For example, a city with only 5% of the population nonwhite and a low density, such as 10 households per block, has an expected value of D of .63 even when the population has been residentially distributed without regard to race or any other demographic characteristic. An observed D value of .63 easily can be misinterpreted as evidence of a highly segregated city when, in fact, it corresponds to no segregation. In other words, this score is a result of particular values of q and N and not of any reality of segregation. It is therefore meaningless to use D to compare the relative segregation of this city to a city, which is 40% black and has one hundred households per block, which has an expected score of .08 when there is an equal lack of segregation. If we were simply to subtract out the effects of proportion and density (accomplished by observed D minus expected D), we would have a more realistic starting point than the conventional zero assumption. We have suggested that the magnitude of this difference is most sensibly carried out in units of the standard deviation. Thus the scores in Table 1 should be considered the zero points under the random assignment of the population to the census units. By employing the standard score corresponding to D for a particular city with its particular proportion of nonwhites and with its particular block structure, we obtain an

Table 1. Expected Value of the Index of Dissimilarity[a] with Random Placement of Nonwhites throughout a City

		(N) Number of Households in Each Areal Unit (e.g., Block)				
		10	25	50	100	1000
(q) Proportion of Minority in the City's Population	.01	.914	.786	.611	.370	.127
	.02	.833	.615	.372	.273	.090
	.05	.630	.369	.264	.180	.043
	.10	.387	.272	.185	.132	.042
	.20	.301	.196	.140	.099	.032
	.30	.266	.176	.122	.087	.028
	.40	.250	.161	.114	.081	.026
	.50	.246	.161	.112	.080	.025

[a] The range of Index of Dissimilarity may vary from 0.000 to 1.000. The values in this table were computed by use of the binomial approximation; see Appendix, equation (21).

improved measure of segregation for that city. Of course, the issue of the proper unit of analysis remains moot and our proposed analysis can be carried out for any choice for which the appropriate census data are available.

Despite the fact that the value of D is affected by the proportion of nonwhites and the block structure, and thus partially vitiates intercity comparisons, such comparisons have often been made in the literature (cf. Taeuber and Taeuber, 1965; Bahr and Gibbs, 1967). The Taeubers (1965:215) have stated, "Our purpose in computing segregation indexes is to permit intercity comparisons, and there must be some sense in which equal values mean equal degrees of segregation." The proposed standard score measure increases the validity of such comparisons.

We have suggested another objection which pertains to the interpretation of the index. D is interpreted as the proportion of nonwhites who would have to change their tract of residence to make the distribution of the minority even throughout the city (Duncan and Duncan, 1955:211). Such an interpretation is misleading in that D gives the proportion of the minority that would have to be moved without replacing them with whites. In particular, if the census tract were wholly populated by a minority, the entire tract would have to be evacuated in this interpretation. Actually, what is often desired is the proportion of minority population which would have to be *exchanged* while keeping the number of households per unit constant. It will be shown in equation (9) below that this "exchange proportion" is, in fact, given by (l-q)D where q is the proportion of the minority in the population.[1]

The Standard Score Index

Our proposal is to compute the expected value of D (μD) and the variance of D (σ^2D) under the hypothesis of randomness and to use the standard score

$$Z = \frac{D\text{-}\mu D}{\sigma D}$$

as an improved measure of segregation. Here D is the measure of deviation from evenness employed by the Duncans. The quantity μD may be interpreted as the value of D which would be achieved *if race had no effect* on the residential distribution. The value of Z gives the degree to which the actual distribution differs from randomness as measured in standard scores.

Since the exact distribution of Z is yet to be determined, no probabilistic interpretation of Z is yet possible. However, what we have provided is a necessary first step which does improve D in allowing, apparently for the first time, the comparison of cities with different nonwhite percentages. In addition, studies of one city over time have always been suspect due to the strong possibility that q has changed over the time period. Our improvement now allows us to deal with this.

If our basic idea for improving D is acceptable, the next step is the determination of the distribution of Z exactly or approximately or empirically. By empirically we mean a categorization of cities by their Z-scores. For example, the knowledge of which cities have Z-scores between 0 and 1, between 1 and 2, etc. would, by itself, be valuable information. In fact, we believe that such information is probably more important to sociologists than the exact statistical probabilities of occurrence. These calculations could be accomplished using the formulae in our mathematical appendix.

While the details of computing Z are left to the appendix, we now make precise what we mean by the "hypothesis of randomness." Suppose that a given city has k census units containing respectively T_1, T_2, . . ., T_k households and suppose that the city-wide proportion of nonwhites is q. We conceive of each unit of size T_1 as being a random sample from the total population, T. We assume the samples are to be selected without replacement, that is, we exclude the possibility that the same person occupies two or more residences. Therefore, the distribution of N_1, the number of nonwhites in unit_i, is *hypergeometric*: that is,

[1] Professor Taeuber has reminded us, in correspondence, that Duncan had pointed this out (see Taeuber and Taeuber, 1965: 30f.) and that Farley and Taeuber (1968:955) have used this formula, identifying it as the "replacement index."

$$\text{Prop}\ (N_i = n_i) = \frac{\binom{qT}{n_i}\binom{(1\text{-}q)T}{T_i\text{-}n_i}}{\binom{T}{T_i}}. \quad (2)$$

We conclude this section by making a general observation about indices. If one has a useful index and also a theory which explains some of the obvious and less relevant factors (e.g., random effects), one can always partial out these factors by constructing a Z-score analogous to that presented above.

The Exchange Proportion Problem

We now turn to a discussion of the interpretation of D as the proportion of minority which would have to change census tracts to bring about evenness. We asserted above that D actually gives this proportion on the assumption that the minority residents are *not* replaced by majority residents. We illustrate this point with two simple examples before passing to the general case. If in the i^{th} census tract there are N_i nonwhites and W_i whites and if in the city there are $N = \sum_i N_i$ nonwhites and $W = \sum_i W_i$ whites, then (see Duncan and Duncan, 1955)

$$D = \tfrac{1}{2} \sum_{i=1}^{k} \left| \frac{N_i}{N} - \frac{W_i}{W} \right|. \quad (3)$$

Here k denotes the total number of census tracts in the city.

For our first example, we consider a city with only two tracts for which $N_1 = 10$, $W_1 = 150$, and $N_2 = 30$, $W_2 = 150$. Clearly, to bring about evenness without replacement, we need merely to move 10 Ns from tract 2 to tract 1. This is a proportion $\frac{10}{40} = \tfrac{1}{4}$ of the minority. From the above formula for D we have:

$$D = \tfrac{1}{2}\left(\left|\frac{10}{40} - \frac{150}{300}\right| + \left|\frac{30}{40} - \frac{150}{300}\right|\right) = \tfrac{1}{4}$$

in agreement with our interpretation. Note that if tract 1 had only 160 $(N_1 + W_1)$ residential units, this would not be a practical way of achieving evenness.

For our second example, we again consider a city composed of two census tracts, but now $N_1 = 0$, $W_1 = 150$, and $N_2 = 30$, $W_2 = 0$ (i.e., a completely segregated city). Here, to bring about evenness without replacement, we must close down tract 2 completely and move all its Ns to tract 1. This is a proportion of $\frac{30}{30} = 1$ of the minority. From the formula for D, we have:

$$D = \tfrac{1}{2}\left(\left|\frac{0}{30} - \frac{150}{150}\right|\right) + \left(\left|\frac{30}{30} - \frac{0}{150}\right|\right) = 1$$

again in agreement with our interpretation.

To obtain the result, we introduce $T_i = N_i + W_i$, the total number of residents in tract i; and $T = N + W$, the total number of residents in the city; $q_i = \frac{N_i}{T_i}$, the proportion of nonwhites in tract i; and $q = \frac{N}{T}$, the city-wide proportion of nonwhites.

Consider a census tract i for which we have too high a proportion of the minority, i.e., $q_i > q$. We must remove enough nonwhites from this tract to bring this proportion down to q. Say we remove R_i: since we do not replace these people we require

$$q = \frac{N_i - R_i}{T_i - R_i}. \quad (4)$$

Solving the equation for R_i, we have

$$R_i = \frac{N_i - T_i q}{1 - q} = \frac{T_i(q_i - q)}{1 - q}. \quad (5)$$

Thus, from all the census tracts with too large a proportion of nonwhites, we remove

$$\sum_{q_i > q} R_i = \frac{1}{1-q} \sum_{q_i > q} T_i\,(q_i - q).$$

That is, a proportion

$$\frac{1}{N} \sum_{q_i > q} R_i = \frac{1}{(1-q)\,qT} \sum_{q_i > q} T_i\,(q_i - q) \quad (6)$$

is removed (recall $q = N/T$). We now claim that the quantity on the right hand side of equation (6) is equal to D. To establish this, we note that

$$\sum_{1=1}^{k} T_i (q_i - q) = \sum_{i=1}^{k} (N_i - T_i q) = N - Tq = 0;$$

hence, the proportion removed can be written as

$$\frac{1}{2(1-q)qT} \sum_{i=1}^{k} |T_i q_i - q| = \frac{1}{2(1-q)qT} \sum_{i=1}^{k} |N_i - T_i q| \qquad (7)$$

since

$$N_i - T_i q = N_i - (N_i + W_i) q = (1-q) N_i - qW_i,$$

the proportion removed is

$$\frac{1}{2} \sum_{i=1}^{k} \left| \frac{N_i}{qT} - \frac{W_i}{(1-q)T} \right| = \frac{1}{2} \sum_{i=1}^{k} \left| \frac{N_i}{N} - \frac{W_i}{W} \right| = D \qquad (8)$$

as we see by equation (3). Therefore, our interpretation of D as the proportion of the minority removed without replacement is established.

As stated earlier, what is usually desired is the proportion of the minority which would have to be *exchanged* with the majority to achieve evenness. Fortunately this quantity has a simple expression in terms of D as we shall now show. Since we now replace those moved, equation (4) becomes

$$q = \frac{N_i - R_i}{T_i}$$

and in place of (5) we have

$$R_i = N_i - qT_i = T_i (q_i - q).$$

Thus the proportion removed throughout the city is:

$$\frac{1}{N} \sum_{q_i > q} R_i = \frac{1}{qT} \sum_{q_i > q} T_i (q_i - q) = (1-q)D.$$

Here the last equality follows from (7, 8). Similarly, the proportion of the majority that must change tracts to bring about evenness is qD and the proportion of the total population that must move is:

$$(1-q) D \cdot \frac{N}{T} + qD \frac{W}{T} = 2q(1-q) D.$$

Therefore, once we have computed D, the proportion of the nonwhite, white, and the total population that must exchange places to achieve evenness are respectively given by

$$(1-q)D, \; qD, \; 2q(1-q). \qquad (9)$$

Conclusion

We feel that the objections to D which have been voiced previously by others or have been raised by us are not minor in their effect on the use of the index. Table 1 clearly demonstrates that the effects of q and N are not "loose" but do in fact produce a systematic deviation.

The corrections suggested here should go far toward allowing for meaningful intercity comparisons as well as providing a more practical interpretation of the index. The former is made possible through the use of the standard scores while the latter is aided by the use of the exchange proportion formulas (9).

It must be pointed out that the qualifications and modifications we have introduced have equal, if not greater, applicability in the use of D outside the confines of residential segregation. Multivariate analyses have used the index for several measures of differentiation or have included the proportion of nonwhites as an independent or intervening variable (cf. Jiobu and Marshall, 1971). Our findings strongly suggest that there is a high degree of confounding effect when q and D appear along with intercity comparisons in multivariate models.

Mathematical Appendix

In computing μD and $\sigma^2 D$, we assume that the T_i residents of the i^{th} census unit are a random sample selected without replacement from the total population of size T. Thus the probability distribution of N_i, the number of nonwhites in the i^{th} census unit, is given by the hypergeometric distribution,

$$P_i = \text{Prob}(N_i = n_i) = \frac{\binom{qT}{n_i}\binom{(1-q)T}{T-n_i}}{\binom{T}{T_i}}. \qquad (1)$$

Furthermore, if we define

$$s_{ij} = n_i + n_j \qquad (2)$$

the joint distribution of the nonwhite populations N_i and N_j on the i^{th} and j^{th} census units is given by

$$P_{ij} = \text{Prob}(N_i = n_i,\ N_j = n_j) = \qquad (3)$$

$$\frac{\binom{T_i}{n_i}\binom{T_j}{s_{ij} - n_i}}{\binom{T_i + T_j}{s_{ij}}} \cdot \frac{\binom{qT}{s_{ij}}\binom{(1-q)T}{T_i + T_j - s_{ij}}}{\binom{T}{T_i + T_j}}.$$

Since the mean and variance of a hypergeometric distribution are well known (Hays and Winkler, 1970), we have at once

$$\mu_i \equiv E[N_i] = qT_i \qquad (4)$$

and

$$\sigma_i^2 \equiv E[(N_i - \mu_i)^2] = T_i q(1-q)\ \frac{T - T_i}{T - 1}. \qquad (5)$$

From equations (7, 8) of the text we have

$$D = \frac{1}{2q(1-q)T} \sum_{i=1}^{k} |N_i - qT_i|. \qquad (6)$$

Thus from (4),

$$D = \frac{1}{B} \sum_{i=1}^{k} |N_i - \mu_i|, \qquad (7)$$

where

$$B = 2q(1-q)T. \qquad (8)$$

This is the form of D that we shall use for computing theoretical expressions for μD and $\sigma^2 D$.

We define the following expectations:

$$e_i = E[|N_i - \mu_i|] = \sum_{n_i=0}^{T_i} |n_i - \mu_i| P_i \qquad (9)$$

$$e_{ij} = E[|N_i - \mu_i|\ |N_j - \mu_j|] = \sum_{s_{ij}=0}^{T_i+T_j} \sum_{n_i=0}^{s_{ij}} |n_i - \mu_i|\ |(s_{ij} - n_i) - \mu_j|\ P_{ij}. \qquad (10)$$

Then, we have at once,

$$\mu D \equiv E(D) = \frac{1}{B} \sum_{i=1}^{k} e_i. \qquad (11)$$

Turning to the computation of $\sigma^2 D$, we have, from (7)

$$\sigma^2 D = \frac{1}{B^2} \{ \sum_{i=1}^{k} \text{Var}(|N_i - \mu_i|) + \sum_{i \neq j} \text{Cov}\ (|N_i - \mu|, |N_j - \mu_j|) \}. \qquad (12)$$

Now

$$\text{Var}(|N_i - \mu_i|) = E((N_i - \mu_i)^2) - E^2(|N_i - \mu_i|) = \sigma_i^2 - e_i^2 \qquad (13)$$

and

$$\text{Cov}(|N_i - \mu_i|, N_j - \mu_j|) = E(|N_i - \mu_i|\ |N_j - \mu_j|) - E(|N_i - \mu_i|)\ E(|N_j - \mu_j|) = e_{ij} - e_i e_j. \qquad (14)$$

Thus,

$$\sigma^2 D = \frac{1}{B^2} \{ \sum_{i=1}^{k} (\sigma_i^2 - e_i^2) + \sum_{i \neq j} (e_{ij} - e_i e_j) \} \qquad (15)$$

$$= \frac{1}{B^2} \{ \sum_{i=1}^{k} \sigma_i^2 - (\sum_{i=1}^{k} e_i)^2 + \sum_{i \neq j} e_{ij} \}.$$

Finally, using (11),

$$\sigma^2 D = \frac{1}{B^2}\ \{ \sum_{i=1}^{k} \sigma_i^2 + \sum_{i \neq j} e_{ij} \} - (\mu D)^2. \qquad (16)$$

Equations (11) and (16) of this appendix constitute theoretical expressions for μD and $\sigma^2 D$. In practice, however, various simplifications and approximations would be employed to simplify the calculations. We now briefly indicate some of the most obvious of these.

In general, since $E[(N_i - \mu_i)] = 0$, equation (9) can be replaced by

$$e_i = 2 \sum_{n_i=0}^{\text{INT}(\mu_i)} (\mu_i - n_i)\ P_i, \qquad (17)$$

where $\text{INT}(\mu_i)$ is the greatest integer less than or equal to μ_i.

We now consider some simplifications arising from approximating the hypergeometric distribution either by the binomial or by the normal distribution. We employ the notation,

$$B(n, q; x) = \binom{n}{x} q^x (1-q)^{n-x}, \quad (18)$$

for the binomial probabilities. In most demographic studies, the census unit populations T_i are all much smaller than the total population T. In this case, it is permissible to replace (1) by

$$P_i = \text{Prob}(N_i = n_i) \cong B(T_i, q; n_i). \quad (19)$$

By a result derived by Cramer (1946), Bardwell (1971) and others, we have in this approximation

$$e_i \cong 2\,\delta_i(1-q)\,B(T_i, q; \delta_i), \delta_i = 1 + \text{INT}(\mu_i), \quad (20)$$

and thus

$$\mu D \cong \frac{1}{qT} \sum_{i=1}^{k} \delta_i\, B(T_i, q; \delta_i), T_i << T. \quad (21)$$

If further $T_i\, q(1-q)$ is large for all the T_i, then the normal approximation to the binomial may be employed in (21) and we obtain

$$\mu D \cong \frac{1}{T\sqrt{2\pi q(1-q)}} \sum_{i=1}^{k} \sqrt{T_i}, \quad T_i << T,\ 1 << T_i\, q(1-q). \quad (22)$$

Similarly, a topic of further investigation is the consideration of the analogous simplifications of P_{ij} and their consequences for $\sigma^2 D$.

REFERENCES

Bahr, Howard M. and Jack P. Gibbs
1967 "Racial differentiation in American metropolitan areas." Social Forces 45: 521–32.

Bardwell, George
1971 Segregation—A Social Account. Denver: Colorado Civil Rights Commission.

Cowgill, Donald O. and Mary S. Cowgill
1951 "An index of segregation based on block statistics." American Sociological Review 16:825–31.

Cramer, Harald
1946 Mathematical Methods of Statistics. Princeton: Princeton University Press.

Duncan, O. D. and Beverly Duncan
1955 "A methodological analysis of segregation indices." American Sociological Review 20:210–7.

Farley, Reynolds and Karl E. Taeuber
1968 "Population trends and residential segregation since 1960." Science 159:953–6.

Hays, William L. and Robert L. Winkler
1970 Statitics, Vol. I. Holt-Rinehart-Winston.

Hornseth, Richard A.
1947 "A note on 'The measurement of ecological segregation' by Julius Jahn, Calvin F. Schmid and Clarence Schrag." American Sociological Review 12:603–4.

Jahn, Julius A.
1950 "The measurement of segregation: derivation on an index based on the criterion of reproducibility." American Sociological Review 15:100–4.

Jahn, Julius A., Calvin F. Schmidt and Clarence Schrag
1947 "The measurement of ecological segregation." American Sociological Review 12:293–303.
1948 "Rejoinder to Dr. Hornseth's note on 'The measurement of ecological segregation.'" American Sociological Re-Review 13:216–7.

Jiobu, Robert M. and Harvey H. Marshall, Jr.
1971 "Urban structure and the differentiation between blacks and whites." American Sociological Review 36:638–49.

Leasure, J. William and David H. Stern
1968 "A note on housing segregation indexes." Social Forces 46:406–7.

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

Williams, Josephine J.
1948 "Another commentary on so-called segregation indices." American Sociological Review 13:298–303.