Measures of Ordinal Segregation

sean f. reardon

520 Galvez Mall, #526
School of Education
Stanford University
Stanford, CA 94305-3084
650.736.8517 (office phone)
650.723.9931 (office fax)
sean.reardon@stanford.edu

Abstract

Purpose: To develop measures of segregation that are appropriate when either the groups or organizational units are defined by *ordered* categories. These methods allow the measurement of segregation among groups defined by ordered educational attainment categories or among ordered occupational categories, for example.

Approach: I define a set of desirable properties of such measures; develop a general approach to constructing such measures; derive three such measures; and show that these measures satisfy the required properties.

Originality: Traditional methods of measuring segregation focus on the measurement of segregation among groups defined by nominal categorical variables (e.g., race, gender) among organizational units also defined by nominal categorical units (e.g., schools, neighborhoods). Such methods are not appropriate to the measurement of occupational segregation, for example. The methods developed here are widely applicable and appropriate for such cases.

Traditional measures of segregation are designed to measure segregation between two groups (e.g., between black and white individuals, or males and females) among nominal categorical units (e.g., schools, neighborhoods) (James and Taeuber, 1985, Zoloth, 1976, Taeuber and Taeuber, 1976). More recently, traditional measures have been extended to allow the measurement of segregation among multiple unordered groups (e.g., multiple racial/ethnic categories) among nominal categorical units (Reardon and Firebaugh, 2002). However, neither the binary nor the multiple group measures are appropriate when either the groups or the units have an inherent ordering. For example, the study of gender segregation among occupations requires the measurement of segregation of two groups among ordered categorical units (occupations). Likewise, the study of residential segregation among groups defined by educational attainment requires the measurement of multiple ordered categorical groups (defined by ordered levels of educational attainment) among nominal categorical groups. In this paper, I develop measures of segregation appropriate to such cases.

The paper proceeds as follows. In the first section, I provide two motivating examples to illustrate the measurement problem. Section II briefly discusses existing approaches of measuring segregation among ordered groups, concluding that little useful work exists. Section III introduces the notation used in the paper. In section IV I describe a set of desirable properties of ordinal segregation measures. Section V derives a set of such measures and demonstrates their relation to existing binary segregation measures. Section VI shows that three of the proposed measures satisfy the criteria outlines in Section III. Section VII concludes with a discussion of the strengths and limitations of the measures.

Motivating Examples

Consider the following patterns of occupational location by gender. Under Scenario A (top panel of Table 1), of the 100 males and females in the labor market, males are disproportionately located in the highest occupational category, while females are disproportionately in the lowest. In Scenario B (bottom panel), the distributions in occupations 1 and 2 are switched, as are the distributions in occupations 3 and 4. As a result, men and women are somewhat less concentrated at one extreme of the occupational distribution.

Table 1 about here

A traditional (nominal) segregation measure—such as the dissimilarity index—would rank Scenarios A and B as identically segregated. In each case, there is one occupation with a 40/10 male/female ratio; one with a 30/20 ratio; one with a 20/30 ratio; and one with a 10/40 ratio. A good measure of occupational segregation, however, should describe A as more segregated than B, given the more dramatic segregation between males and females in A.

Likewise, consider the patterns of educational attainment by neighborhood described in Table 2. In Scenario A (top panel), those with the lowest level of educational attainment are concentrated in neighborhoods 1 and 2, where there are few individuals with the highest levels of education. In Scenario B, the distributions of the lowest two educational categories are switched with one another, as are the distributions of the highest two categories. Intuitively, educational attainment segregation is lower in B than A, because those with the lowest and highest educational levels are more likely to occupy the same neighborhoods.

Table 2 about here

As in Table 1, standard (nominal) segregation measures will rank Scenarios A and B identically segregated, because they will treat the ordered educational attainment categories as unordered and interchangeable. In both Tables 1 and 2, then, the need for measures of segregation that account for the ordered nature of one of the two categories involved is apparent.

Prior Methods of Measuring Ordinal Segregation

There is little prior methodological work developing measures of ordinal segregation. Most research on occupational segregation treats occupational categories as unordered (Charles and Grusky, 1995, Watts, 1992, Watts, 1997, Watts, 1998, Grusky and Pager, 1998, Hutchens, 1991, Silber, 1992, Siltanen, 1990). Some research on economic segregation relies on ordered income categories, but again, much of this work treats income as either binary or unordered (Telles, 1995, Fong, 2000). One exception is a recent paper that modifies traditional exposure-type segregation indices by introducing a "social difference coefficient"—a factor that accounts for the ordered nature of income categories (Meng et al., 2006). This paper, however, relies on somewhat arbitrary definitions of social differences (it assigns interval values to an ordinal scale) and derives exposure measures but not evenness measures of segregation.

A second type of prior work comes from measures of nominal-ordinal association. Reardon and Firebaugh argue that one way to conceptualize segregation measures is as measures of association between categorical variables (Reardon and Firebaugh, 2002). While they were concerned with segregation as association between two nominal categorical variables (e.g., racial groups and schools), measures of association between a nominal and ordinal category might be thought of as segregation measures. The generalized delta and alpha measures introduced by Agresti, for example, could perhaps be seen as measures of ordinal segregation (Agresti, 1981), though I am aware of no work that has made this link. Moreover, the properties of such measures are not clear.

The third strand of related work is the development of methods for measuring segregation along a continuously-measured variable, such as income. A considerable body of such work exists in the literature on residential household income segregation. For example, some work relies on ordered income categories and computes two-group segregation measures for many or all possible pairs of income categories, and then constructs some average or summary measure of these multiple pairwise indices (Massey and Eggers, 1990, Farley, 1977, Telles, 1995). This approach is rather *ad hoc* in nature, and depends heavily on the number of income categories used. Other measures compute a ratio of the between-neighborhood variation to overall variation, constructing measures of the proportion of variation in incomes that lies between neighborhoods (Davidoff, 2005, Wheeler, 2006, Wheeler and La Jeunesse, 2006, Jargowsky, 1996, Jargowsky, 1997, Ioannides, 2004, Ioannides and Seslen, 2002, Hardman and Ioannides, 2004, Watson, 2006). While there are many more examples of such measures, it is not clear how they inform the measurement of segregation along an ordinal variable, since they depend on the interval scale of income to compute between- and within-neighborhood variation.

In this paper, I develop a new approach to the measurement of ordinal segregation that is derived from the variation ratio approach described in Reardon and Firebaugh (2002) and similar to that used in many of the income segregation papers referred to above. This approach is based in the idea that segregation can be thought of as a measure of the extent to which variation within organizational units (unordered categories) is less than total variation in the population. By using well-defined measures of ordinal variation, the Reardon and Firebaugh approach is readily adapted to the case of ordinal segregation.

Notation

Let k be a variable denoting ordered categories $1, 2, \ldots, K$, and let m index unordered categories $1, 2, \ldots, M$. For example, k might index K ordered educational attainment categories (less than high school, high school diploma, bachelor's degree, graduate degree) and m might index neighborhoods. Conversely, m might index (unordered) racial groups and k might index (ordered) occupational status categories (manual labor, technical, professional). These two cases may seem conceptually quite different – in the first case, the ordered educational attainment category is a fixed characteristic of individuals (for the purposes of this example) and the unordered neighborhood category is changeable; while in the second, the opposite is true – but we shall see that they are mathematically identical.

Moreover, suppose we have a population of size T where each individual is cross-classified by m and k. Let $t_{m\cdot}$, $t_{\cdot k}$, and t_{mk} indicate the total population in category m, category k, and both, respectively. Within each unordered category m, denote the cumulative proportion of the population in m who are in ordered category k or below as c_{mk}; i.e.,

$$c_{mk} = \sum_{j=1}^{k} \frac{t_{mj}}{t_m}$$

[1]

Now we can write the distribution of individuals in category m as the K-tuple $\boldsymbol{c_m} = (c_{m1}, c_{m2}, \dots, c_{mK})$. (Note that $c_{mK} = 1$ for all m, so we need only the [K-1]-tuple $\boldsymbol{c_m} = \left(c_{m1}, c_{m2}, \dots, c_{m[K-1]}\right)$ to fully characterize the distribution in m). Likewise, we let c_k indicate the cumulative proportion of the population who are in category k or below, and write the distribution of individuals in the population as the [K-1]-tuple $\boldsymbol{c} = \left(c_1, c_2, \dots, c_{[K-1]}\right)$.

One additional piece of terminology will be useful in the following sections. We say that the distribution in m dominates the distribution in n over categories h to j (where $1 \leq h < j \leq K$) if $c_{mk} < c_{nk}$ for all $k \in (h, \dots, j)$. In other words, m dominates n over h to j if there is a greater proportion of the population of n than m at or below each category from h to j.

Desirable Properties of an Ordinal Segregation Measure

The goal of this paper is to define a class of ordinal segregation indices that measure what is generally termed the "evenness" dimension of segregation. That is, we wish to measure the extent to which ordered groups are evenly distributed across unordered categories. In the case of occupational segregation, this corresponds to measuring the extent to which ordered occupational categories are proportionately represented among unordered groups (typically gender or racial/ethnic groups). In the case of educational residential segregation, this corresponds to measuring the extent to which different educational attainment groups are evenly distributed among residential locations (e.g., neighborhoods).

In order to define a set of properties that a measure of ordinal segregation should satisfy, it is useful to consider several simple examples. Below we describe a set of desirable properties of an ordinal segregation measure, using examples of residential educational segregation or of racial occupational segregation in order to make the properties concrete.

Scale Interpretability. A segregation index is maximized *iff* within each unordered category m, all individuals occupy a single ordered category (if $c_{mk} \in \{0,1\}$ for all m and k). A segregation index is minimized *iff* within each unordered category m, the distribution of individuals is identical to that in the population (if $c_{mk} = c_k$, for all m and k).[1] The intuition here is that maximum segregation occurs when there is no variation in ordered category k within each unordered unit, and minimum segregation occurs when the distribution of k within each unit m mirrors that of the population as a whole, as illustrated in Scenarios A and B of Table 3.

It is important to note that this definition of maximum segregation does not depend on the overall amount of ordinal variation in the population, but only on the relative distribution among categories. This is a necessary requirement for defining evenness measures of segregation. In Scenario C of Table 3, for example, there is less variation in educational attainment than in Scenario A (everyone in C has either a HS diploma or a BA; no one has less than a HS diploma or more than a BA), yet by the definition here, Scenarios A and C are equally segregated, because in each case the unordered groups

[1] Note that in the case where the number of non-empty ordinal categories is greater than the number of groups, maximum segregation is not possible by this definition.

are as unevenly distributed among neighborhoods as possible (there is no neighborhood with any diversity of educational attainment). Conversely, to claim that Scenario A is more segregated than Scenario C would be to confound the extent of variation in educational attainment in the population with the unevenness of their distribution among neighborhoods. Given our interest in measuring the evenness dimension of segregation, this would be undesirable.

Table 3 about here

Unordered Category Equivalence. If the populations in two unordered categories m and n have the same distribution over k, and m and n are combined, segregation is unchanged. This is analogous to the standard organizational equivalence condition used in nominal measures of segregation (James and Taeuber, 1985, Reardon and Firebaugh, 2002). The intuition here is that two identically distributed unordered categories are not segregated from one another, so combining them does not alter overall segregation patterns. In Table 4, for example, Black and Hispanic individuals have the same distribution among occupations, so combining them in to a “minority” group should not alter conclusions about the levels of segregation.

Table 4 about here

Size invariance. If the number of members of each category k in each category m is multiplied by a positive constant, segregation is unchanged. The intuition here is simply that the total population counts do not matter, but only the distributions among categories.

Null group effects. If an ordered category j with no population (so that $t_{mj} = t_{j\cdot} = 0$ for all m) is added to the vector of possible ordered categories $k = 1, 2, \dots, K$, the following conditions should yield unambiguous results. First, if $j = 0$ (i.e, category j is lower than each of the K ordered categories) or $j = K + 1$ (i.e, category j is higher than each of the K ordered categories), then segregation is unchanged. Second, if $1 < j < K$ and $c_{mj} \in \{0,1\}$ for all m (in each unordered category m, all the population has $k < j$ or $k > j$), and there exist some unordered categories m and n where $c_{mj} = 0$ and $c_{nj} = 1$, then segregation will increase, unless segregation is already at its maximum possible value.

The intuition here is that the ordinal categories have some inherent meaning. Adding a null category at the extremes doesn't change the separation of groups, since it introduces no additional distance between any two individuals. But adding a null category that unambiguously increases the distance between higher and lower ordered categories (because for each unordered category m, all individuals are above or below the new ordered category), should increase segregation (unless it is already at its maximum).

Exchanges. If the distribution in category m dominates that in category n over categories j to k, and if an individual of category j is moved from n to m while an individual of category k is moved from m to n, then segregation is reduced. Likewise, if an individual in m is moved from category k to j, while an individual in n is moved from j to k, then segregation is reduced.[2]

The intuition here is that if we exchange individuals of different ordered categories in a way that makes the distributions of individuals in two unordered categories more similar to one another, segregation should be reduced. For example, if neighborhood m has a lower proportion of its population at or below each education level than neighborhood n, and if a highly-educated individual in neighborhood m swaps houses with a less educated individual in n, then segregation should be reduced—because we have raised the education levels in n and lowered them in m. Likewise, if there

[2] Note that we could specify transfer properties, but the exchange properties have the advantage of leaving the margins constant.

are more females than males at or below each occupation status level, and if a man in a high-status occupation changes jobs with a woman in a low status occupation, segregation should be reduced. This latter is illustrated in Table 4. In Scenario A, the male occupational distribution dominates the female distribution over categories 1-3 (there are a greater proportion of females than males at or below each occupational level from 1-3). If a male moves from occupation 3 to 2, while a female moves from 2 to 3, we expect segregation to be reduced, because it has reduced the disparity in cumulative proportions at or below income categories 2 and 3, while leaving the others unchanged.

Table 5 about here

Ordered Exchanges. If the distribution in unordered category m dominates that in category n over categories j to l (with $j < k < l$), and if an individual of category j is moved from n to m while an individual of category l is moved from m to n, then the resulting reduction in segregation will be greater than that resulting if an individual of category j is moved from n to m while an individual of category k is moved from m to n. Likewise, if an individual in m is moved from category l to j, while an individual in n is moved from j to l, then the resulting reduction in segregation will be greater than that resulting if an individual in m is moved from category k to j, while an individual in n is moved from j to k.

The intuition here is that the principal of exchanges ought to be sensitive to the ordering of the categories involved. Exchanges of individuals who are farther apart (in their ordered category levels) should change segregation more than exchanges of individuals who are closer. If a male doctor switches jobs with a female janitor, segregation should be reduced more than if a male doctor switches jobs with a female nurse. This is illustrated in the bottom panel of Table 4. Scenario C is obtained from A by having a category 3 male switch occupations with a category 1 female. By the principle of exchanges, segregation C is less than B, and B is less than A, so segregation in C should be less than in A.

Additive Unordered Group Decomposability. If *M* unordered categories are clustered into *N* clusters, then segregation is decomposable into a sum of independent within- and between-cluster components. The intuition here is that we would like to be able to decompose segregation into between- and within-category components, as described by Reardon and Firebaugh (2002).

Additive Ordered Group Decomposability. If K ordered categories are clustered into J supercategories through the combination of adjacent categories, then segregation is decomposable into a sum of independent within- and between-supercategory components. For example, we might like to be able to describe segregation among occupations as a sum of a component of segregation between professional and non-professional occupations, a component of segregation among non-professional occupations, and a component of segregation among professional occupations.

Measuring Segregation by an Ordinal Category

Reardon and Firebaugh (2002), Watson (2006), and Jargowsky and Kim (THIS VOLUME) suggest that one way of constructing a segregation measure is to think of it as a measure of the ratio of between-category inequality/variation to total inequality/variation. In this framework, segregation is the proportion of the total variation of some characteristic in a population that is due to differences in population composition of different categorical units (e.g., schools or census tracts). For an unordered categorical variable (such as race), population variation may be measured by diversity or entropy, but for an ordered (ordinal or continuous) variable, variation is typically measured using some index of the spread of the distribution (e.g., the variance or standard deviation, in the case of an interval-scaled variable). Following Reardon and Firebaugh (2002, Eq. 9, p. 45), if we have a suitable measure of ordinal variation v (a measure of the variation in ordinal variable k) we can define $\Lambda^{(v)}$—an ordinal segregation measure based on the ordinal variation measure v—as follows:

$$\Lambda^{(v)} = \sum_{m=1}^{M} \frac{t_m}{Tv}(v - v_m)$$

[2]

where m indexes unordered categories, t_m and v_m are the population count and ordered variation of k within category m, respectively, and T is the total population. Note that if v is the variance function (treating k as an interval-scale variable), then $\Lambda^{(v)}$ is equivalent to η^2 (or the R^2 from a regression of k on a set dummy variables for categories $m = 1, 2, \ldots, M$).

One way to construct a measure of ordinal segregation, then, is to define a suitable measure of ordinal variation (I define 'suitable' below), and then use it to construct a segregation measure as above.

Ordinal variation

Measuring the variation in a population of a quantity measured with an ordinal variable requires us to define what we mean by variation. For an ordinal variable k that can take on any of K ordered categories $1, 2, \ldots, K$, we define variation as having a maximum (which we can normalize to equal 1) when half the population has $k = 1$ and half has $k = K$. Variation is at a minimum (normalized to equal 0) when all observations have $k = c$ for some $c \in (1, 2, \ldots, K)$. Measuring ordinal variation then amounts to measuring how close the distribution of *x* is to these minimum and maximum variation states.

As described above, we can express the distribution of values of x in a sample as a [*K*-1]-tuple of cumulative proportions, $\boldsymbol{c} = \left(c_1, c_2, \ldots, c_{[K-1]}\right)$, where c_j is the cumulative proportion of the sample with values of k in category *j* or below (note that $c_K = 1$ by definition, so is not needed to characterize the distribution of k). A distribution of k has maximum variation when $\boldsymbol{c} = \boldsymbol{c_0} = (½, ½, ½, \ldots, ½)$, corresponding to the case where half the population has the lowest possible value and half has the highest possible value of k. There are K possible distributions of k such that there is no variation in k, corresponding to the K cases of the pattern $\boldsymbol{c} = (0,0, \dots 0,0,1,1, \dots 1,1)$ (where all observations have the same value of $k = c$, so that $c_j = 0$ for $j < k$ and $c_j = 1$ for all $j \geq k$).

Blair and Lacy (1996) suggest that it is helpful to think of $\boldsymbol{c}$ as a point in $[K-1]$-space, which leads to the insight that variation can be measured as an inverse function of the distance from $\boldsymbol{c}$ to $\boldsymbol{c_0}$, the point of maximum variation (it is easier to define variation in terms of the distance from the single point of maximum variation rather than from one of the K points of zero variation). Alternately, we can think of $\boldsymbol{c}$ as describing a cumulative density function of an ordinal variable, where variation is measured as a function of the closeness of $\boldsymbol{c}$ to the line $c_k = ½$. This suggests a general form of a variation measure as:

$$v = \frac{1}{K-1} \sum_{j=1}^{K-1} f(c_j)$$

[3]

where $f(c)$ is a continuous function defined on the interval $[0,1]$ such that $f(c)$ is increasing for $c \in \left(0, \frac{1}{2}\right)$ and decreasing for $c \in \left(\frac{1}{2}, 1\right)$, maximized on the interval $[0,1]$ at $f(½) = 1$, and minimized on the interval $[0,1]$ at $f(0) = f(1) = 0$. Four such possible functions f are defined below:[3]

$$f_1(c) = -[c\log_2 c + (1-c)\log_2(1-c)]$$

$$f_2(c) = 4c(1-c)$$

$$f_3(c) = 2\sqrt{c(1-c)}$$

$$f_4(c) = 1 - |2c - 1|$$

[4]

Substituting these functions into [3] above yields four potential measures of ordinal variation,

[3] Note that we define $0 \cdot \log_2(0) = \lim_{x \to 0} x \log_2 x = 0$.

which are denoted v_1, v_2, v_3, and v_4, respectively.[4] Each of these ordinal variation measures equals its maximum value, 1, if and only if the distribution of k is such that $k = 1$ for half the observations and $k = K$ for the other half (corresponding to the cumulative proportion vector $\boldsymbol{c} = (½, ½, ½, \dots, ½, 1)$), and each equals its minimum value of 0 if and only if all observations have the same value of k (corresponding to a cumulative proportion vector $\boldsymbol{c} = (0,0, \dots 0,1, \dots ,1,1)$).

Ordinal segregation

We can substitute v_1, v_2, v_3, and v_4 into Equation [2] above to produce measures of ordinal segregation Λ_1, Λ_2, Λ_3, and Λ_4:

$$\Lambda_1 = \sum_{m=1}^{M} \frac{t_m}{Tv_1}(v_1 - v_{1m})$$

$$\Lambda_2 = \sum_{m=1}^{M} \frac{t_m}{Tv_2}(v_2 - v_{2m})$$

$$\Lambda_3 = \sum_{m=1}^{M} \frac{t_m}{Tv_3}(v_3 - v_{3m})$$

$$\Lambda_4 = \sum_{m=1}^{M} \frac{t_m}{Tv_4}(v_4 - v_{4m})$$

[5]

The index Λ_1 is an ordinal generalization of the information theory index H, so we will denote Λ_1 as H^O, and call it the *ordinal information theory index*. In the case where $K = 2$, H^O is identical to H, the

[4] The second of these, v_2, is the index of ordinal variation (IOV), which has been derived in a number of alternate, but equivalent forms (Berry and Mielke, 1992a, Berry and Mielke, 1992b, Blair and Lacy, 1996, Kvålseth, 1995a, Kvålseth, 1995b). The other three indices of ordinal variation are new, to my knowledge. Note that, in the case where $K = 2$, several of these variation measures collapse to familiar measures of binary variation. When $K = 2$, $v_1 = -[c\log_2 c + (1-c)\log_2(1-c)]$, which is the entropy (Theil, 1972) of a two-group population when one group makes up c percent of the population. Likewise, when $K = 2$, $v_2 = 4c(1-c)$, which is twice the Simpson interaction index diversity measure for a two-group population when one group makes up c percent of the population.

conventional two-group information theory index (Theil, 1972, Theil and Finezza, 1971, Zoloth, 1976). Similarly, Λ_2 is an ordinal generalization of the variation ratio index R (which goes by many names in the literature—see, e.g., James and Taeuber, 1985, Reardon and Firebaugh, 2002), and is identical to that index when $K = 2$, so we will denote Λ_2 as R^O, and call it the *ordinal variation ratio index*. In addition, Λ_3 is an ordinal generalization of Hutchen's square root index (Hutchens, 2001, Hutchens, 2004), and is identical to that index when $K = 2$, so we will denote Λ_3 as S^O, and call it the *ordinal square root index*. Finally, Λ_4 might be termed the *ordinal absolute difference index*, denoted D^O. Unlike the other three indices, however, D^O does not correspond to any standard binary index when $K = 2$; in particular, it is not equivalent to the dissimilarity index (James and Taeuber, 1985, Taeuber and Taeuber, 1976), despite its apparently similar form. Each of these four indices is interpreted as the average difference between overall and within-unit ordinal variation, expressed as a ratio of the overall ordinal variation of the population. They differ only in the functional forms of the ordinal variation measures each employs.

<u>Alternate formulations of the ordinal segregation measures</u>

The ordinal segregation measures can be written in another form that illustrates their relation to existing measures. First, we require some additional notation. Let $\nu_j = f(c_j)$ and $\nu_{mj} = f(c_{mj})$; these can be thought of as measures of the (ordinal) variation in the population and in category m, respectively, in the case where the population is divided into only two categories—those with $k \leq j$ and those with $k > j$. Note that Equation [3] implies $\sum_{i=1}^{K-1} \nu_j = v(K-1)$. In addition, let Λ^j indicate the segregation (as defined in [2]) when the population is similarly divided into two categories at j. With some substitution and algebraic manipulation, Equation [2] can be written as

$$\Lambda = \sum_{m=1}^{M} \frac{t_m}{Tv}(v - v_m)$$

$$= \sum_{m=1}^{M} \frac{t_m}{Tv(K-1)} \sum_{j=1}^{K-1} [f(c_j) - f(c_{mj})]$$

$$= \sum_{j=1}^{K-1} \frac{v_j}{v(K-1)} \sum_{m=1}^{M} \frac{t_m}{Tv_j} (v_j - v_{mj})$$

$$= \sum_{j=1}^{K-1} w_j \Lambda^j$$

[6]

where $w_j = \frac{v_j}{\sum_{i=1}^{K-1} v_i}$ and $\sum_{j=1}^{K-1} w_j = 1$. The key insight to be gained from [6] is that a measure of ordinal segregation defined by Equation [2] (as a function of the difference between the ordinal variation of the population and the average ordinal variation within each category *m*) can be rewritten as the weighted average of *K*-1 binary (non-ordinal) segregation indices. The weights are proportional to $v_j = f(c_j)$, the variation in k when collapsed into two categories corresponding to those above and below category j. Because $f(c_j)$ is defined to have its maximum at $c_j = \frac{1}{2}$, Equation [6] weights segregation defined by categories that most evenly split the population.

For three of the four ordinal measures defined in [4], the pairwise indices Λ^j in Equation [6] are familiar indices: $\Lambda_1^j = H^j$ (the two-group Theil information theory index); $\Lambda_2^j = V^j$ (the two-group variance ratio index); and $\Lambda_3^j = O^j$ (the two-group Hutchens square root index). This means that the properties of Λ in these cases can be derived easily from known properties of the corresponding Λ^j's. The index defined by Λ_4^j, however, corresponds to no existing pairwise index. Moreover, as we shall see, it fails to satisfy the scale interpretability and principle of exchanges properties desired in a segregation index.

Properties of the ordinal segregation measures

Scale interpretation. Each of the four indices we describe are bounded between 0 and 1, obtaining their maximum when segregation is complete and their minimum when each unordered category has an identical distribution of members of the ordered categories. For three of indices (except Λ_4), this follows most obviously from Equation [6] and the fact that each of their corresponding binary indices are bounded between 0 and 1 (Hutchens, 2001, Hutchens, 2004, James and Taeuber, 1985, Reardon and Firebaugh, 2002). For Λ_4, the minimum value of segregation can be obtained even when segregation is not complete (that is, when members are not evenly distributed among categories). To see this, consider the example in Table 6. The value of Λ_4 in this example is 0, despite the fact that the distribution of males and females among occupations is not equal. Thus, Λ_4 does not meet the scale interpretability criterion.

Table 6 about here

Unordered category equivalence. Each index of the form defined in Equation [2] satisfies this property. This follows readily from the definition in [2] and the fact that v depends only on cumulative proportions and not cell totals. Let r be the unordered category formed by combining unordered categories p and q, where $c_{pk} = c_{qk}$ for all $k \in \{1,2,\ldots,K\}$. From Equation [1], we then have $c_{rk} = c_{pk} = c_{qk}$ for all $k \in \{1,2,\ldots,K\}$, and thus $v_r = v_p = v_q$. Now

$$\Lambda = \sum_{m \neq p,q} \frac{t_m}{Tv}(v - v_m) + \frac{t_p}{Tv}(v - v_p) + \frac{t_q}{Tv}(v - v_q)$$

$$= \sum_{m \neq p,q} \frac{t_m}{Tv}(v - v_m) + \frac{t_p + t_q}{Tv}(v - v_r)$$

$$= \sum_{m \neq p,q} \frac{t_m}{Tv}(v - v_m) + \frac{t_r}{Tv}(v - v_r)$$

Size invariance. Each index of the form defined in Equation [2] satisfies this property. This follows readily from the fact that each c_k and c_{mk} is unchanged if the number of members of each category k in each category m is multiplied by a constant. As a result, v and each v_m are unchanged as well. It follows that Equation [2] is unchanged under scalar multiplication of all cell sizes.

Null group effects. Recall from [3] that ordinal variation is the average

$$v = \frac{1}{K-1}\sum_{j=1}^{K-1} f(c_j)$$

If we add a null category i below category 1 ($i = 0$) or above K ($i = K + 1$), the cumulative proportion in category *i* will be 0 (if $i = 0$) or 1 (if $i = K + 1$) for the population as a whole and for each group *m.* As a result, both v and v_m are multiplied by a factor of $\frac{K-1}{K}$ (because $f(c_i) = f(c_{mi}) = 0$); these factors will cancel in Equation [2], leaving Λ unchanged.

If we add a null category i such that $1 < i < K$, $0 < c_i < 1$, and $c_{mi} \in \{0,1\}$ for all m, each v_m will be multiplied by a factor of $\frac{K-1}{K}$ as well. The overall ordinal variation in the population, v, will be multiplied by a factor greater than $\frac{K-1}{K}$, because $f(c_i) > 0$. As a result, Λ will increase. If, however, $0 < c_{mj} < 1$ for some m, then, in general, Λ may increase, decrease, or remain unchanged.

Exchanges. To investigate the exchange property, it is simplest to use the form of Λ in Equation [6].

$$\Lambda = \sum_{j=1}^{K-1} w_j \Lambda^j$$

Under an exchange, the marginal distributions of individuals among the categories does not change; that is, the c_k's and the t_m's do not change (and because the c_k's are unchanged, so are the v_j's and v). As a result, Λ can change only if the Λ^j's change.

Recall that Λ^j measures the segregation of those with $x > j$ from those with $x \leq j$. In an exchange involving individuals of category j and k (where $j < k$), there will be no change in the cumulative proportions c_{mi}, where $i < j$ or $i > k$. As a result, there will be no change in Λ^i. Now consider categories i such that $j \leq i \leq k$. If the distribution in category m dominates that in category n from j to k, then $c_{mi} < c_{ni}$ for all these i. Exchanging an individual of category j from n with an individual of category k from m will increase c_{mi} and decrease c_{ni} for $j \leq i \leq k$. If Λ^i is a binary segregation index satisfying the principle of exchanges, then Λ^i will decrease for each $j \leq i \leq k$ in such a case. As a result, Λ will decrease as well.

Three of the Λ^j's described correspond to binary segregation indices satisfying the principle of exchanges. The Information Theory Index, Variance Ratio Index, and Square Root Index each satisfy the principle of exchanges (James and Taeuber, 1985, Hutchens, 2001, Hutchens, 2004, Reardon and Firebaugh, 2002). The fourth index does not satisfy the principle of exchanges except when the exchange involves transfers from categories whose cumulative proportion is above .5 to those whose cumulative proportion is less than .50, or vice versa. In other cases, exchanges will leave the index unchanged. To see this, consider the example in Table 7.

Table 7 about here

In this example, Scenario B is obtained from Scenario A by the exchange of a male in category 1 with a female in category 2 (a male moves from occupation category 1 to category 2, while a female moves from category 2 to category 1). Intuitively, this should increase segregation, because this exchange exacerbates the disproportionate representation of females in the lowest category. However, in both cases, $\Lambda_4 = 0$.[5] Thus, Λ_4 does not meet the exchange criterion.

Ordered Exchanges. Following the logic above, an exchange between ordered categories j and l, where $j < k < l$, will reduce each Λ^i for $i \in \{j, \dots, l\}$. An exchange between categories j and k will reduce only the Λ^i's where $i \in \{j, \dots, k\}$. As a result, an index which can be written as in Equation [6] and which satisfies the principle of exchanges will also satisfy the principle of ordered exchanges.

Additive Unordered Group Decomposability. If the binary index Λ^j is additively decomposable in the sense described by Reardon and Firebaugh (2002), then so is Λ. Let the M unordered categories be combined into some smaller number of supercategories, indexed by g. Then for an additively decomposable index, we can write:

$$\begin{aligned}
\Lambda &= \sum_{j=1}^{K-1} w_j \Lambda^j \\
&= \sum_{j=1}^{K-1} w_j \left[\Lambda_G^j + \sum_g \frac{t_g v_g}{Tv} \Lambda_g^j \right] \\
&= \sum_{j=1}^{K-1} w_j \Lambda_G^j + \sum_g \sum_{j=1}^{K-1} w_j \frac{t_g v_g}{Tv} \Lambda_g^j \\
&= \Lambda_G + \sum_g \frac{t_g v_g}{Tv} \sum_{j=1}^{K-1} w_j \Lambda_g^j \\
&= \Lambda_G + \sum_g \frac{t_g v_g}{Tv} \Lambda_g
\end{aligned}$$

[8]

Thus, Λ can be written as a sum of a component due to segregation among the supercategories and a

[5] The failure of Λ_4 to meet the scale interpretability and exchange criteria is a direct result of the fact that the variation function v_4 is not a concave down function of the c_j.

weighted sum of the segregation within each supercategory.

Additive Ordered Group Decomposability. None of the four indices described here satisfy the ordered group decomposability criterion. While this is a desirable analytic property, it is not clear whether any index can satisfy it.

Discussion and Conclusion

The approach I have described above yields four potential measures of segregation along an ordered dimension. Of these, three of the measures satisfy the set of criteria described at the outset. The characteristics of each measure are described in Table 8 below. Each of the measures can be interpreted in two ways: 1) as a measure of relative ordinal variation (a measure of the difference in the ordinal variation of the population to the average ordinal variation within each unordered category); and 2) as a weighted average of the binary segregation between those above and below each threshold of the ordered variable. The first interpretation links the measures to the analysis of within- and between-group variation; the measures are analogous to η^2-type statistics, denoting the proportion of variation in a population that lies between, rather than within, organizational units. The second interpretation links the measures to traditional binary segregation measures. It shows that we can compute ordinal segregation by computing binary segregation measures between the groups of individuals above and below each threshold of an ordered categorical variable, and then computing a weighted average of these binary measures over all thresholds. In each case, the weights are greatest for thresholds that divide the ordered variable such that half the population is above and half below the threshold.

Table 8 here

The weights w_j in Equation [6] have intuitive interpretations in some cases. In the case of the ordinal information theory index, the weight w_j assigned to H^j in [6] can be interpreted as a the expected information content contained in H^j about a randomly chosen individual from the population, where information is defined as $\log_2 c_j$, where c_j is the proportion of the population in or below the j^{th} ordered category (Theil, 1972, Pielou, 1977). If c_j is near 0 or 1, then H^j contains little information about the segregation experienced by an individual, since it distinguishes among individuals only at one extreme of the income distribution. Conversely, if c_j is near 0.5, then H^j contains maximal information, since it distinguishes at the median of the population.

In the case of the ordinal relative variation index, the weight w_j likewise has an appealing interpretation. For a given threshold j, the probability that two randomly-selected individuals from the population will have incomes on opposite sides of threshold j is $2c_j(1 - c_j)$, which is proportional to $f(c_j)$. Since the segregation level R^j describes the extent of segregation between individuals on either side of the j^{th} ordinal threshold, we can interpret Equation [6] in this case as a weighted average of the segregation between individuals on either side of the j^{th} ordinal threshold, where the segregation level at each threshold is weighted by how informative segregation measured at that threshold is for a randomly chosen pair of individuals in the population.

In the case of the ordinal square root index, the weights have no clear intuitive interpretation (they are the square root of the weights in the ordinal relative variation index). Nonetheless, because each of the variation functions in Equation [4] have their maximum at $c_j = \frac{1}{2}$, and their minima at $c_j = 0$ and $c_j = 1$, each of the ordinal segregation measures weight binary segregation between groups defined by the median of the ordinal distribution most heavily, and segregation between the extreme groups and the remainder least heavily. Intuitively, this makes sense, since a segregation level computed between those in a small extreme ordinal category and all others, for example, tells us very little about the segregation between two randomly chosen individuals, while segregation between those above and below the median income tells us more about overall income segregation.

Although the weights in some cases have intuitive interpretations, it is important to note that the class of ordinal segregation measures described here were not derived originally as a weighted sum of $K-1$ pairwise segregation indices and thus do not rely on the interpretation of the weights for their derivation. Rather, the weights derived in Equation [6] are simply a consequence of the definition of ordinal segregation as a function of the ratio of within-unit to overall ordinal variation. If the weights are less intuitive than one might like, they can nonetheless be understood as a useful as providing a means of computing the ordinal segregation measures from easily computed and familiar pairwise indices. One could perhaps imagine an alternate class of segregation measures defined explicitly as a weighted average of pairwise measures, where the weights were based on some preferred value function rather than derived from the approach here. Weighting each of the pairwise measures equally, for example, might be appealing under some circumstances.

In some cases, one might prefer a segregation measure that is not normalized by the variance in the population. Some recent papers have argued that measures of absolute, rather than relative, deviation, such as the mutual information index (Theil, 1971), are preferable to normalized indices because of their superior decomposition properties (Frankel and Volij, 2008, Mora and Ruiz-Castillo, 2003).[6] In fact, Equation [6] can be easily rewritten to suggest that the ordinal segregation measures here can be understood as normalized unweighted averages of pairwise segregation measures that are not normalized by population variation. Rearranging the terms in the third line of Equation [6] yields

$$\Lambda = \frac{1}{v}\sum_{j=1}^{K-1}\frac{v^j\Lambda^j}{(K-1)}$$

[6] The choice between measures like the mutual information theory index M (which is not normalized by the population variation) and the information theory index H (which is) lies more in how one believes indices should compare segregation between populations that differ in their overall variation. The decomposition differences are, in my opinion, a minor issue compared to this issue. A fuller discussion of the comparison is beyond the scope of this paper.

$$= \frac{1}{v} \sum_{j=1}^{K-1} \frac{\Lambda^{*j}}{(K-1)}$$

[9]

where $\Lambda^{*j} = v^j \Lambda^j$ is a measure of pairwise segregation that is not normalized by dividing it by v (see Equation [2]). The mutual information theory index M is such an index. Because $M^j = v_1 H^j$ by definition (see Theil, 1971), Equation [9], in fact, can be readily used to show that the ordinal information theory index H^O is easily transformed into an alternate index one might call the ordinal mutual information theory index, M^O:

$$H^O = \frac{1}{v_1} \sum_{j=1}^{K-1} \frac{M^j}{(K-1)}$$

$$= \frac{1}{v_1} M^O$$

[10]

where M^j is the mutual information theory index when the population is divided into two categories at j; and where M^O is the (unweighted) average of M^j over the $K-1$ ordered thresholds. Thus, the approach described in this paper easily produces a set of measures that are not based on normalized measures. These measures would not satisfy the scale interpretability and null group effects properties described above (though advocates of the mutual information theory index typically discard scale interpretability in any case). A full investigation of the properties of this set of measures, however, is beyond the scope of the present paper.

In this paper, I have described an approach to measuring segregation among groups defined by ordered categories. The approach described here combines the appealing features of defining segregation as a ratio of variation within categories to variation in a population and those of averaging binary segregation measures among pairs of categories defined by an ordered variable. Three measures of ordinal segregation derived from this approach—the ordinal information theory index (H^O), the ordinal variation ratio index (R^O), and the ordinal square root index (S^O)—each satisfy the set of desired characteristics for such measures defined here. In addition to the characteristics described above, each of the indices described here can easily be adapted to incorporate spatial proximity, using the methods described by Reardon and O'Sullivan (2004). This may be useful, for example, if one wished to measure the spatial residential segregation among groups defined by ordered categories of educational attainment.

Acknowledgements

This paper benefitted from helpful conversation with Paul Jargowsky, Robert Hutchens, and Glenn Firebaugh, as well as the thoughtful comments of an anonymous and insightful reviewer. Support was provided by the William T. Grant Foundation and the National Science Foundation (SES-0520400). I claim responsibility for all errors.

References

AGRESTI, A. (1981) Measures of Nominal-Ordinal Association. *Journal of the American Statistical Association,* 76**,** 524-529.

BERRY, K. J. & MIELKE, P. W., JR. (1992a) Assessment of variation in ordinal data. *Perceptual and Motor Skills,* 74**,** 63-66.

BERRY, K. J. & MIELKE, P. W., JR. (1992b) Indices of ordinal variation. *Perceptual and Motor Skills,* 74**,** 576-578.

BLAIR, J. & LACY, M. G. (1996) Measures of variation for ordinal data as functions of the cumulative distribution. *Perceptual and Motor Skills,* 82**,** 411-418.

CHARLES, M. & GRUSKY, D. B. (1995) Models for describing the underlying structure of sex segregation. *American Journal of Sociology,* 100**,** 931-971.

DAVIDOFF, T. (2005) Income sorting: Measurement and decomposition. Berkeley, CA, UC-Berkeley.

FARLEY, R. (1977) Residential segregation in urbanized areas of the United States in 1970: An analysis of social class and racial differences. *Demography,* 14**,** 497-518.

FONG, E., & SHIBUYA, KUMIKO (2000) The spatial separation of the poor in Canadian cities. *Demography,* 37**,** 449-459.

FRANKEL, D. M. & VOLIJ, O. (2008) Measuring school segregation.

GRUSKY, D. B. & PAGER, D. (1998) Models for analyzing occupational segregation by race.

HARDMAN, A. & IOANNIDES, Y. M. (2004) Neighbors' income distribution: Economic segregation and mixing in U.S. urban neighborhoods. *Journal of Housing Economics,* 13**,** 368-382.

HUTCHENS, R. M. (1991) Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations. *Mathematical Social Sciences,* 21**,** 31-51.

HUTCHENS, R. M. (2001) Numerical measures of segregation: Desirable properties and their implications. *Mathematical Social Sciences,* 42**,** 13-29.

HUTCHENS, R. M. (2004) One measure of segregation. *International Economic Review,* 45**,** 555-578.

IOANNIDES, Y. M. (2004) Neighborhood income distributions. *Journal of Urban Economics,* 56**,** 435-457.

IOANNIDES, Y. M. & SESLEN, T. N. (2002) Neighborhood wealth distributions. *Economics Letters,* 76**,** 357-367.

JAMES, D. R. & TAEUBER, K. E. (1985) Measures of segregation. *Sociological Methodology,* 14**,** 1-32.

JARGOWSKY, P. A. (1996) Take the money and run: Economic segregation in U.S. metropolitan areas. *American Sociological Review,* 61**,** 984-998.

JARGOWSKY, P. A. (1997) *Poverty and place: Ghettos, barrios, and the American city,* New York, Russell Sage Foundation.

KVÅLSETH (1995a) Coefficients of variation for nominal and ordinal categorical data. *Perceptual and Motor Skills,* 80**,** 843-847.

KVÅLSETH (1995b) Comment on the coefficient of ordinal variation. *Perceptual and Motor Skills,* 81**,** 621-622.

MASSEY, D. S. & EGGERS, M. L. (1990) The ecology of inequality: Minorities and the concentration of poverty, 1970-1980. *American Journal of Sociology,* 95**,** 1153-1188.

MENG, G., HALL, G. B. & ROBERTS, S. (2006) Multi-group segregation indices for measuring ordinal classes. *Computers, Environment and Urban Systems,* 30**,** 275-299.

MORA, R. & RUIZ-CASTILLO, J. (2003) Additively Decomposable Segregation Indexes. The Case of Gender Segregation by Occupations and Human Capital Levels in Spain. *Journal of Economic Inequality,* 1**,** 147-179.

PIELOU, E. C. (1977) *Mathematical ecology,* New York, John Wiley & Sons.

REARDON, S. F. & FIREBAUGH, G. (2002) Measures of multi-group segregation. *Sociological Methodology,* 32**,** 33-67.

REARDON, S. F. & O'SULLIVAN, D. (2004) Measures of spatial segregation. *Sociological Methodology,* 34**,** 121-162.

SILBER, J. (1992) Occupational segregation indices in the multidimensional case: A note. *The Economic Record,* 68**,** 276-277.

SILTANEN, J. (1990) Social change and the measurement of occupational segregation by sex: An assessment of the sex ratio index. *Work, Employment and Society,* 4**,** 1-29.

TAEUBER, K. E. & TAEUBER, A. F. (1976) A practitioner's perspective on the index of dissimilarity. *American Sociological Review,* 41**,** 884-889.

TELLES, E. E. (1995) Structural sources of socioeconomic segregation in Brazilian metropolitan areas. *American Journal of Sociology,* 100**,** 1199-1223.

THEIL, H. (1971) *Principles of Econometrics,* New York, Wiley & Sons.

THEIL, H. (1972) *Statistical decomposition analysis,* Amsterdam, North-Holland Publishing Company.

THEIL, H. & FINEZZA, A. J. (1971) A note on the measurement of racial integration of schools by means of informational concepts. *Journal of Mathematical Sociology,* 1**,** 187-194.

WATSON, T. (2006) Metropolitan growth, inequality, and neighborhood segregation by income. Williamstown, MA, Williams College.

WATTS, M. (1992) How should occupational sex segregation be measured? *Work, Employment and Society,* 6**,** 475-487.

WATTS, M. (1997) Multidimensional indexes of occupational segregation: A critical assessment. *Evaluation Review,* 21**,** 461-482.

WATTS, M. (1998) Occupational Gender Segregation: Index Measurement and Econometric Modeling. *Demography,* 35**,** 489-496.

WHEELER, C. H. (2006) Urban decentralization and income inequality: Is sprawl associated with rising income segregation across neighborhoods? St. Louis, MO, Federal Reserve Bank of St. Louis.

WHEELER, C. H. & LA JEUNESSE, E. A. (2006) Neighborhood income inequality. St. Louis, MO, Federal Reserve Bank of St. Louis.

ZOLOTH, B. S. (1976) Alternative measures of school segregation. *Land Economics,* 52**,** 278-298.

(Note: Need to insert the Jargowsky & Kim reference from THIS VOLUME:

JARGOWSKY, P. A. & KIM, J. (THIS VOLUME) The information theory of segregation: uniting segregation and inequality in a common framework. *Research on Economic Inequality*.

Table 1: Stylized Occupational Distribution Patterns

	Occupational Level				
	1	2	3	4	Total
			Scenario A		
Male	10	20	30	40	100
Female	40	30	20	10	100
Total	50	50	50	50	200
			Scenario B		
Male	20	10	40	30	100
Female	30	40	10	20	100
Total	50	50	50	50	200

Table 2: Stylized Educational Neighborhood Distribution Patterns

	Neighborhood				
	1	2	3	4	Total
		Scenario A			
Less than HS	50	40	10	0	100
HS Diploma	40	30	20	10	100
BA	10	20	30	40	100
BA+	0	10	40	50	100
Total	100	100	100	100	400
		Scenario B			
Less than HS	40	30	20	10	100
HS Diploma	50	40	10	0	100
BA	0	10	40	50	100
BA+	10	20	30	40	100
Total	100	100	100	100	400

Table 3: Maximum and Minimum Segregation Patterns

	Neighborhood				
	1	2	3	4	Total
		Scenario A			
Less than HS	0	100	0	0	100
HS Diploma	100	0	0	0	100
BA	0	0	0	100	100
BA+	0	0	100	0	100
Total	100	100	100	100	400
		Scenario B			
Less than HS	25	25	25	25	100
HS Diploma	25	25	25	25	100
BA	25	25	25	25	100
BA+	25	25	25	25	100
Total	100	100	100	100	400
		Scenario C			
Less than HS	0	0	0	0	0
HS Diploma	100	100	0	0	200
BA	0	0	100	100	200
BA+	0	0	0	0	0
Total	100	100	100	100	400

Table 4: Unordered Category Equivalence

	Occupational Level				
	1	2	3	4	Total
		Scenario A			
White	10	20	30	40	100
Black	20	20	10	0	50
Hispanic	10	10	5	0	25
Total	40	50	45	40	175
		Scenario B			
White	10	20	30	40	100
Minority	30	30	15	0	75
Total	40	50	45	40	175

Table 5: Exchange Properties of Segregation

	Occupational Level				
	1	2	3	4	Total
		Scenario A			
Male	10	20	30	40	100
Female	40	30	20	10	100
Total	50	50	50	50	200
		Scenario B			
Male	10	21	29	40	100
Female	40	29	21	10	100
Total	50	50	50	50	200
		Scenario C			
Male	11	20	29	40	100
Female	39	30	21	10	100
Total	50	50	50	50	200

Table 6: Scale Interpretability Example

	Occupational Level				
	1	2	3	4	Total
	Scenario A				
Male	10	40	25	25	100
Female	40	10	25	25	100
Total	50	50	50	50	200

Table 7: Principle of Exchanges Example

	Occupational Level				
	1	2	3	4	Total
		Scenario A			
Male	10	40	25	25	100
Female	40	10	25	25	100
Total	50	50	50	50	200
		Scenario B			
Male	9	41	25	25	100
Female	41	9	25	25	100
Total	50	50	50	50	200

Table 8: Measures of Ordinal Segregation

Ordinal Index	Symbol	Variation Formula	Two-Group Analog	Weight
Ordinal Information Theory Index	H^O	$v_1 = \frac{-1}{K-1}\sum_{j=1}^{K-1}\left[c_j \log_2 c_j + (1-c_j)\log_2(1-c_j)\right]$	H	$v_j = -\left[c_j \log_2 c_j + (1-c_j)\log_2(1-c_j)\right]$
Ordinal Variation Ratio Index	R^O	$v_2 = \frac{4}{K-1}\sum_{j=1}^{K-1} c_j(1-c_j)$	R	$v_j = 4c_j(1-c_j)$
Ordinal Square Root Index	S^O	$v_3 = \frac{2}{K-1}\sum_{j=1}^{K-1}\sqrt{c_j(1-c_j)}$	S	$v_j = 2\sqrt{c_j(1-c_j)}$
Ordinal Absolute Difference Index	D^O	$v_4 = 1 - \frac{1}{K-1}\sum_{j=1}^{K-1}\lvert 2c_j - 1\rvert$	none	n/a

Note: H is the Theil Information Theory Index (Theil, 1972, Theil and Finezza, 1971); R is the Relative Variation Index (which goes by many names in the literature—see, e.g., James and Taeuber, 1985, Reardon and Firebaugh, 2002); S is Hutchens' Square Root Index (Hutchens, 2001, Hutchens, 2004).