INTERNATIONAL ECONOMIC REVIEW
Vol. 45, No. 2, May 2004

ONE MEASURE OF SEGREGATION*

BY ROBERT HUTCHENS[1]

School of Industrial and Labor Relations, Cornell University, U.S.A.

This article considers the problem of deriving a numerical measure of segregation, i.e., a measure of inequality in the distribution of people across groups. It proposes a list of eight desirable properties for a good numerical measure of segregation. These properties yield a class of segregation indexes that are related to generalized entropy indexes of income inequality. Moreover, one and only one index—termed the square root index—satisfies seven of the properties.

1. INTRODUCTION

Concern about unequal access to jobs, schools, and neighborhoods has resulted in a large body of empirical work that examines inequality in the distribution of people across groups. For example, the literature on occupational segregation of men and women is partly a response to concerns about unequal access to jobs. Since the seminal work of Duncan and Duncan (1955), numerical measures have played a central role in such studies. This article contributes to the methodological side of that literature by proposing a list of desirable properties—or criteria—for what constitutes a "good" numerical measure of segregation. It then proves that one and only one measure satisfies those criteria.

The main innovation in this article—the innovation that makes it possible to prove the existence of a unique measure—is the introduction of two related properties: an aggregative property and an additive decomposability property. Specifically, the article is built on concepts of aggregation and additive decomposability presented in Anthony Shorrocks' analysis of income inequality (Shorrocks 1980, 1984). It essentially applies these concepts to measures of segregation. In building on Shorrocks' article, this article also helps to forge links between the literature on measuring income inequality and the literature on measuring segregation.

In accordance with the aggregative property, measured segregation for an aggregate should be a continuous and increasing function of segregation within subaggregates. For example, suppose we split the occupational structure into two subaggregates: blue collar occupations and white collar occupations. An

* Manuscript received October 2002; revised March 2003.

[1] I am grateful to Gary Fields for helpful comments and suggestions. Special thanks to Alexei Egorov and, at an earlier stage, Eduardo Zambrano, for help with the proofs. Special thanks also to Kim Weeden for graciously letting me use her data. What errors remain are, of course, mine alone. Please address correspondence to: Robert Hutchens, School of Industrial and Labor Relations, Cornell University, 263 Ives Hall, Ithaca, NY 14853-3901. Phone: 607-255-2749. Fax: 607-255-4496. E-mail: *rmh2@cornell.edu*.

aggregative property requires that the measure of segregation for all occupations be a function of measured segregation within the blue collar and the white collar subaggregates. As developed below, an additive decomposable property requires that this function be additive.

Not only does an aggregative property place restrictions on the functional form of an inequality measure, but it opens the way "to decompose the overall inequality level into the inequality contributions associated with each of the subgroups, or, alternatively, to aggregate upwards from the subgroup values to derive the composite figure" (Shorrocks, 1984, p. 1369). As such, an aggregative property is a desirable property for a measure of segregation. A further advantage is that the property provides the key to a proof that links a small set of properties to a single measure of segregation.

This article first introduces eight properties for a measure of segregation and then establishes that seven of these properties are sufficiently restrictive to yield a single index, termed the square root index. Alternatively stated, a measure of segregation satisfies the seven properties if and only if it is a square root index. The aricle goes on to illustrate the use of the index by decomposing changes in occupational segregation in the United States between 1980, 1990, and 2000.[2]

2. DESIRABLE PROPERTIES OF A MEASURE OF SEGREGATION

Consider a society where two types of people are distributed over T groups. Let x_{ij} be the number of persons of type i in group j ($i = 1, 2; j = 1, \ldots, T$), and let x be the matrix of these x_{ij}. The total number of people in group j is then $x_{1j} + x_{2j}$. As an example, type 1 could be women and type 2 could be men, and group j could be one of T occupations. Denote the total number of type i people over all groups as $N_i(x) = \sum_j x_{ij}$, $i = 1, 2$. To insure that the problem is meaningful, assume that T is greater than one, that the x_{ij} are nonnegative real numbers,[3] and that the numbers of type 1 and type 2 people ($N_1(x)$ and $N_2(x)$) are both positive. For example, for three occupations,

$$\text{(1)} \qquad x = \begin{array}{c} \\ \\ \text{type 1} \\ \text{type 2} \end{array} \begin{array}{c} \text{Occupation} \\ \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \end{array} \\ \begin{bmatrix} 3 & 1 & 4 \\ 8 & 4 & 4 \end{bmatrix} \end{array}$$

Thus, if type 1 people are women the distribution in (1) has three women and eight men in the first occupation, and the complete distribution contains a total of eight women and 16 men ($N_1 = 3 + 1 + 4 = 8$ and $N_2 = 8 + 4 + 4 = 16$). To

[2] Other authors have decomposed total segregation into components associated with population subgroups. For example, Flückiger and Silber (1999), Kakwani (1994), and Watts (1998). This article is different in that it applies the insights of Shorrocks (1984) to the segregation problem, and it presents a set of axioms that yield a unique measure.

[3] An x matrix with noninteger values is certainly possible. Part-time workers could be treated as fractional workers. Irrational numbers are also theoretically admissible (e.g., a full-time worker is counted as "1" and a part-time worker as the square-root of 0.3). When the domain is specified in terms of real numbers, it is possible to exploit results in the income inequality literature that utilize the same domain.

simplify matters, when possible the $N_i(x)$ are written N_i, $i = 1, 2$. Denoting the vector space of all $2 \times T$ real matrices with nonnegative elements by $\mathbf{R}_+^{2\times T}$, the domain of x shall be $D = \cup_{T=2}^{\infty} D_T$ where

$$D_T = \left\{x \in \mathbf{R}_+^{2\times T} : N_i(x) > 0, i = 1, 2\right\}.$$

The basic goal of this article is to develop a numerical measure that indicates whether the matrix x is more segregated than another matrix y. Let $O(x)$ be a numerical measure of segregation. More precisely, $O(x)$ is a continuous function defined on D such that $O(x) \geq O(y)$ iff x has as much segregation as y for each x, $y \in D$. This article thus seeks to characterize $O(x)$.

Segregation curves provide a useful starting point for this problem. Segregation curves, which were introduced in Duncan and Duncan (1955), are similar to the Lorenz curves used to assess income inequality. A segregation curve plots the cumulative fraction of type 1 people (on the vertical axis) and type 2 people (on the horizontal axis), both fractions being ranked from low to high values of x_{1j}/x_{2j}. Figure 1 illustrates for the 1980, 1990, and 2000 occupational distributions of men and women across occupations. (Section 6 discusses the data underlying these curves.) We say that one distribution is more segregated than another if the curve for the former distribution lies everywhere below the latter. Thus, Figure 1 indicates that the 1980 distribution is more segregated than the 2000 distribution.

A segregation curve indicates complete integration (or no segregation) when it lies along the diagonal line OA. In this case the ratio of women to men is the same in all occupations. An example of complete integration would be, $\left[\begin{smallmatrix} 4 & 2 & 2 \\ 8 & 4 & 4 \end{smallmatrix}\right]$. Note that the ratio of women to men in each occupation is 1 to 2. Introducing notation, denote a distribution with no segregation by the $2 \times T$ matrix $\Omega(T, N_1, N_2)$,

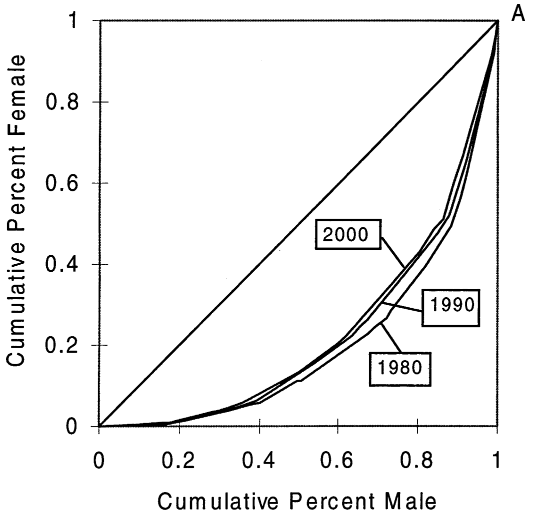

FIGURE 1

SEGREGATION CURVES FOR 1980, 1990, AND 2000—DERIVED FROM TABLE 1 MALE AND FEMALE OCCUPATIONAL DISTRIBUTIONS where the ω_{ij}, the row i and column j elements of $\Omega(T, N_1, N_2)$, satisfy the property $\omega_{1j}/N_1 = \omega_{2j}/N_2$, $j = 1, T$.

A segregation curve indicates complete segregation when it lies along the two axes, i.e., it takes the form of the right angle formed by the vertical line through A and the horizontal axis. In this case all the men are in one set of occupations and all the women are in another set; no occupation contains both men and women. For a distribution with 8 women and 16 men, an example of complete segregation would be, $\left[\begin{smallmatrix} 8 & 0 & 0 \\ 0 & 9 & 7 \end{smallmatrix}\right]$.

Like Lorenz curves, segregation curves have both advantages and disadvantages. Their main disadvantage is that they provide an incomplete or partial ordering of distributions from most segregated to least segregated. If the curves intersect, segregation curves are silent about which distribution is more segregated. Therein lies the reason for pursuing numerical measures: Unlike segregation curves, numerical measures produce complete rankings.

The main advantage to segregation curves is that they provide a simple way to rank distributions, and that ranking rests on particularly unrestrictive value judgments. Hutchens (1991) shows that if the segregation curve for distribution y lies at some point below and at no point above the segregation curve for x, then every member of a well-defined class of measures will indicate that y is more segregated than x. That class of measures is termed the RIMFO class, where RIMFO is short for *relative inequality measure for occupations*. A member of the RIMFO class must exhibit the first four properties below.[4] Although these properties are not sufficiently restrictive to yield a complete ordering, they form the basis for numerical measures that do yield complete orderings.

The first property is scale invariance. By this property, if N_1 and/or N_2 are multiplied by a positive scalar and the share of both types of people in the T occupations remains the same, then segregation does not change. Thus follows:

> P1. Scale Invariance: Let $\Phi = \left[\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right]$ where α and β are positive scalars. If $y = \Phi x$, then $O(x) = O(y)$.

To illustrate, suppose that over a decade the number of women (N_1) in the labor force doubles, whereas the number of men (N_2) remains the same. Suppose, in addition, that the new occupational distribution of women duplicates the old in the sense that an occupation that previously contained no women continues to contain no women, and an occupation that previously contained 30% of all women continues to do so. Thus, despite the increase in the *level* of women in the labor force, the share of women in an occupation does not change (where the share is calculated as the number of women in the occupation divided by the number of women in the labor force). According to the scale invariance property, such a change in the number of women in the labor force does not change the level of segregation.

[4] Hutchens (1991) indicates that any index satisfying P1–P3 will yield the same ranking as nonintersecting segregation curves. In fact, an index must satisfy P1–P4. Theorem 1 of that paper presents a standard result linking the inequality ranking of Lorenz curves to every relative inequality measures for income (RIMFI). It then recasts the theorem in a notation of groups that is almost the same as that used here. That change in notation implicitly assumes that the value of a RIMFI is invariant to the change, and thereby implies that a RIMFI must also satisfy P4.

The second property is symmetry in groups. By this property if the people in group i trade places with those in group k then segregation is unaffected. Thus follows:

P2. Symmetry in Groups: If $(j_1, \ldots, j_T)$ is any permutation of $1, \ldots, T$,

$$x = \begin{bmatrix} x_{11}, x_{12}, \ldots, x_{1T} \\ x_{21}, x_{22}, \ldots, x_{2T} \end{bmatrix} \text{ and } y = \begin{bmatrix} x_{1j_1}, x_{1j_2}, \ldots, x_{1j_T} \\ x_{2j_1}, x_{2j_2}, \ldots, x_{2j_T} \end{bmatrix}, \text{ then } O(x) = O(y)$$

In other words, segregation does not depend on whether a group is labeled 3, 6, or 22. What is important is the number of type 1 and type 2 people in the group.

The third property is movement between groups. This property is similar to the principle of transfers in the income inequality literature. Whereas the income inequality literature utilizes a regressive transfer, here I use a "disequalizing movement," which is defined as follows:

A Disequalizing Movement: The distribution y is obtained from x by a disequalizing movement of type 1 people if, for i and j, (a) $x_{2i} = x_{2j} = y_{2i} = y_{2j} > 0$, (b) $x_{1i}/x_{2i} < x_{1j}/x_{2j}$, (c) $y_{1i} = x_{1i} - d$, and $y_{1j} = x_{1j} + d$ for $0 < d \leq x_{1i}$, and (d) $x_{hk} = y_{hk}$, $h = 1, 2$; $k \neq i, j$.

This then leads to the third property.

P3. Movement Between Groups: If y is obtained from x by a disequalizing movement, then $O(y) > O(x)$.

To illustrate, let type 1 people be women and type 2 people be men, and suppose that the initial distribution is as in (1) above. Note that since occupation 2 contains one woman and four men and occupation 3 contains four women and four men, $x_{12}/x_{22} = 1/4 < 4/4 = x_{13}/x_{23}$. According to the third property, if one woman in occupation 2 moves to occupation 3 (so that occupation 2 contains no women and four men and occupation 3 contains five women and four men), then segregation has increased. The movement not only increased the percent male in the more "male" occupation, but also increased the percent female in the more "female" occupation.

The fourth property is termed insensitivity to proportional divisions, and is concerned with splitting a large group into subgroups. If the subgroups have the same ratio of type 1 to type 2 people as the original large group, then segregation should not be affected. For example, if one large group with 100 men and 100 women is split into five subgroups, each containing 20 men and 20 women, then measured segregation should not change. When new wine is poured into old bottles, it remains new wine.

More formally,

Proportional Division of a Group: The distribution y is obtained from x by a proportional division of a group if

$$\begin{aligned} y_{ij} &= x_{ij}, \qquad i = 1, 2;\ j = 1, \ldots, T-1 \\ y_{ij} &= x_{iT}/M, \quad i = 1, 2;\ j = T, \ldots, T+M-1 \end{aligned}$$

where M is a positive integer.

The fourth property simply states that when a group is divided into subgroups through a proportional division, segregation should not be altered. Thus follows:

P4. Insensitivity to Proportional Divisions: If y is obtained from x by a proportional division of a group, then $O(x) = O(y)$.

As noted above, P1–P4 are only sufficient to justify a partial ranking of distributions. All measures that satisfy P1–P4 will rank y as more segregated than x so long as the segregation curve for y lies at some point below and at no point above the segregation curve for x. If the segregation curves for x and y intersect, then measures that satisfy P1–P4 can produce different conclusions about which distribution is more equal. Additional properties are needed to obtain a complete ranking. An aggregative property serves that purpose.

3. AGGREGATIVE MEASURES

In analyzing measures of income inequality Shorrocks (1984) introduced an aggregative property. This section first summarizes key elements of Shorrocks (1984), and then extends that analysis to the problem of measuring segregation.

Shorrocks defines an aggregative measure for incomes in a population of n individuals. Denote the income of the ith individual by v_i, and let $v = (v_1, v_2, \ldots, v_n)$ represent incomes for the population. Divide this population into two mutually exclusive and exhaustive subgroups, containing n_1 and n_2 people. Let $v^1 = (v_1^1, v_2^1, \ldots, v_{n1}^1)$ and $v^2 = (v_1^2, v_2^2, \ldots, v_{n2}^2)$ be vectors of incomes for people in groups 1 and 2, and let μ^1 and μ^2 denote mean incomes for the two groups. A continuous income inequality measure $I(v)$ is *aggregative* if and only if there exists an "aggregator" function A such that

$$I(v) = I(v^1; v^2) = A\big(I(v^1), \mu^1, n_1, I(v^2), \mu^2, n_2\big)$$

where A is continuous and strictly increasing in $I(v^1)$ and $I(v^2)$.

In the class of aggregative income inequality measures, within-group and between-group inequality have precise meaning. Within-group inequality is $I(v^1)$ for group 1 and $I(v^2)$ for group 2. Between-group inequality is that level of inequality that occurs when there is no within-group inequality for either group 1 or group 2, i.e., all elements of v^1 and v^2 are set at their respective group means. Thus, between-group inequality equals $A(I(\mu^1 e^{n_1}), \mu^1, n_1, I(\mu^2 e^{n_2}), \mu^2, n_2)$, where e^{n_g} is a vector of n_g ones.

In practical applications analysts often prefer measures that are not simply aggregative but also *additively decomposable*. An additive decomposable measure of income inequality takes the form

$$I(v) = \sum_g w_g I(v^g) + B$$

where w_g is the weight attached to group g ($g = 1, \ldots, G$), B is a between-group term, and

$$w_g = w_g(\mu^1, \ldots, \mu^G, n_1, \ldots, n_G) > 0$$
$$B = B(\mu^1, \ldots, \mu^G, n_1, \ldots, n_G)$$

All additive decomposable measures are aggregative with A taking an additive form.[5]

Because they permit transparent decompositions of total inequality into "within" and "between" components, additive decomposable measures have enjoyed wide use. For example, Schwarze (1996) uses an additive decomposable measure (a Theil index) to decompose German income inequality into components that measure within-East Germany inequality, within-West Germany inequality, and between East and West inequality.[6] He then examines how Germany's reunification affected the three components.

Not all measures of income inequality are aggregative. Perhaps the most prominent example of a measure that is not aggregative is the Gini coefficient.[7] Thus, if $I(v)$, $I(v^1)$, and $I(v^2)$ are specified as Gini measures of income inequality, there does not exist an aggregator function A that is continuous and strictly increasing in $I(v^1)$ and $I(v^2)$ for all feasible v^1 and v^2.

What income inequality measures are aggregative? Shorrocks (1984) examined measures of income inequality that are both aggregative and relative,[8] and concluded that a small family of income inequality measures satisfies these criteria: the generalized entropy measures and their positive transformations. See Shorrocks' (1984) Theorem 5. This paper uses Foster's (1985) restatement of Shorrocks' result. (Rather than refer to Foster's restatement of Shorrocks' Theorem 5, this article simply refers to Shorrocks' Theorem 5.) Denoting the generalized entropy measures as $I_c(v)$, and defining the domain as the positive real numbers D_+,[9] we have the following:

SHORROCKS' THEOREM 5. *Let I: $D_+ \to \mathbf{R}$ be continuous and let $I(e^n) = 0$ for each $n \geq 1$, where e^n is a vector of n ones. Then I is an aggregative relative inequality measure iff there exists a continuous, strictly increasing function F:$[0, \infty) \to \mathbf{R}$, with $F(0) = 0$, such that $F(I(v)) = I_c(v)$ (all v element D_+) for some c, where*

$$I_c(v) = \begin{cases} \frac{1}{n}\frac{1}{c(c-1)}\sum_{i=1}^n \left(\left(\frac{v_i}{\bar{v}}\right)^c - 1\right), & c \neq 0, 1 \\ \frac{1}{n}\sum_{i=1}^n \frac{v_i}{\bar{v}}\ln\left(\frac{v_i}{\bar{v}}\right), & c = 1 \\ \frac{1}{n}\sum_{i=1}^n \ln\left(\frac{v_i}{\bar{v}}\right), & c = 0 \end{cases}$$

and $\bar{v} = \frac{1}{n}\sum_{i=1}^n v_i$.

[5] Foster (1985, p. 60).

[6] Although Schwarze (1996) provides a useful illustration, there exist many earlier applications of these methods. See for example, Mookherjee and Shorrocks (1982) and Anand (1983).

[7] See Pyatt (1976) and Mookherjee and Shorrocks (1982).

[8] As discussed in Foster (1985) a relative measure of income inequality exhibits the properties of homogeneity, symmetry, the Pigou–Dalton transfers principle, and the population principle. These relative measures are also termed relative inequality measures for income (RIMFI).

[9] More precisely, $D_+ := \cup_{n=1}^{\infty} D_{n+}$ where $D_{n+} = \{v \varepsilon \mathbf{R}^n : v_i > 0 \text{ for all } i\}$.

Suitably modified, this theorem is applicable to aggregative measures of segregation.[10] As before let x represent a $2 \times T$ matrix of type 1 and type 2 people over T occupations. Partition this matrix into two mutually exclusive and exhaustive subaggregates containing T^1 and T^2 occupations such that $x = [x^1, x^2]$ where x^1 is a $2 \times T^1$ matrix that denotes the distribution of people across the T^1 occupations, and x^2 is a similar matrix for the remaining $T^2 = T - T^1$ occupations. Given this, define an aggregative measure of segregation in a manner analogous to Shorrocks (1984).[11]

An Aggregative Measure: A measure of segregation is aggregative if there exists an "aggregator" function A such that,

$$\begin{aligned} O(x) &= O(x^1, x^2) \\ &= A\big(O(x^1), N_1(x^1)\big/N_2(x^1), N_2(x^1), O(x^2), N_1(x^2)\big/N_2(x^2), N_2(x^2)\big) \end{aligned}$$

where A is continuous and strictly increasing in $O(x^1)$ and $O(x^2)$.[12]

It is useful to state this as a property. Thus follows:

P5. Aggregation: $O(x)$ is aggregative.

Given this, one obtains a theorem for segregation that parallels Shorrocks' Theorem 5. Using "aggregative RIMFO" to indicate a measure of segregation that satisfies properties 1–5, we have the following:

[10] Shorrocks (1988) introduces a property of subgroup consistency, and uses this to provide an alternative justification for the generalized entropy measures. While the argument in this article could have been built from Shorrocks (1988), since the results in Shorrocks (1984) are used in the Shorrocks (1988) proof, it seemed sensible to similarly build this article from Shorrocks (1984).

[11] In fact, the definition of aggregative measure of segregation simply uses an alternative notation to restate the definition of an aggregative measure of income inequality. To see this consider a population of n people with incomes described by the $n \times 1$ vector v. In a manner similar to Hutchens (1991) classify this population into T homogeneous groups such that the members of each group have identical incomes. (If K people in v have an income of \$1000, then those K people would constitute one of the T groups.) Let x_{1j} be the total amount of income received by people in group j, and let x_{2j} be the number of people in group j. Then the vector v can be restated as the matrix x. When the definition of aggregative measure of income inequality is rewritten in terms of this alternative notation, one obtains the text's expressions for an aggregative measure of segregation. The same logic applies to the additive decomposable measure of segregation in Section 4.

[12] Neither of the most widely used measures of segregation—the dissimilarity index or the Gini—are aggregative. The literature on income inequality establishes that the Gini measure of income inequality is not aggregative, and that argument applies to the Gini measure of occupational inequality. See Foster (1985, p. 64). With regard to the dissimilarity index, let $x_1 = \left[\begin{smallmatrix} 8 & 6 & 14 \\ 1 & 1 & 1 \end{smallmatrix}\right]$ and $x_2 = \left[\begin{smallmatrix} 6 & 8 \\ 1 & 1 \end{smallmatrix}\right]$. Writing the dissimilarity index, $O_d(x) = \frac{1}{2}\sum_{i=1}^{T} |x_{1i}/N_1 - x_{2i}/N_2|$, $O_d(x_1) = 0.1667$ and $O_d(x_2) = 0.0714$, and overall $O_d(x_1, x_2) = 0.1333$. Suppose one type 1 person in x_1 moves from the second group to the first, yielding distribution x_1^*, thus, $x_1^* = \left[\begin{smallmatrix} 9 & 5 & 14 \\ 1 & 1 & 1 \end{smallmatrix}\right]$. This movement does not alter the dissimilarity index for x_1, i.e., $O_d(x_1^*) = O_d(x_1) = 0.1667$. If $O_d(x)$ were aggregative, we would observe $O_d(x_1^*; x_2) = O_d(x_1; x_2)$, since $O_d(x_1^*) = O_d(x_1)$. In fact $O_d(x_1^*; x_2) = 0.1476 > 0.1333 = O_d(x_1, x_2)$. Then $O_d(x)$ is not aggregative.

THEOREM 1. *Let* $O: D \to \mathbf{R}$ *be a continuous RIMFO with* $\Omega(T, N_1, N_2) = 0$ *for each* $T > 1$, *and all* $N_1 > 0$, and $N_2 > 0$. *Then O is an aggregative RIMFO if and only if there exists a continuous, strictly increasing function* $F: [0, \infty) \to \mathbf{R}$, *with* $F(0) = 0$, *such that* $F(O(x)) = O_c(x)$ (*all* $x \in D$), *where*,

$$O_c(x) = -\sum_{j=1}^{T} s_{2j}\left[(s_{1j}/s_{2j})^c - 1\right] \qquad 0 < c < 1$$

with $s_{1j} = x_{1j}/N_1$ *and* $s_{2j} = x_{2j}/N_2$.

PROOF. See the Appendix.

Thus, within the set of continuous measures that satisfy properties 1–4, there is a small family of measures that is aggregative. This family contains $O_c(x)$ and its positive transformations. To remain consistent with the income inequality literature, $O_c(x)$ shall be referred to as a generalized entropy measure of segregation.[13]

These generalized entropy measures of segregation should be quite useful. They are easy to compute, and they permit decompositions of total segregation into segregation within and between components. Before moving to applications, however, there remains the issue of deriving a unique index.

4. ONE INDEX

Three additional properties are required to derive a unique index. The first property is additive decomposability. An additive decomposable measure of segregation can be defined in a manner analogous to the additive decomposable measures of income inequality in Shorrocks (1984). Thus follows:

An Additive Decomposable Measure of Segregation: A measure of segregation is additive decomposable if it takes the form,

$$O(x) = \sum_g w_g O(x^g) + B$$

where w_g is the weight attached to group g ($g = 1, \ldots, G$), B is the between-group term, and

$$\begin{aligned} w_g &= w_g\big(N_1(x^1), \ldots, N_1(x^G), N_2(x^1), \ldots, N_2(x^G)\big) > 0 \\ B &= B\big(N_1(x^1), \ldots, N_1(x^G), N_2(x^1), \ldots, N_2(x^G)\big) \\ &= A\big(\Omega\big(T^1, N_1(x^1), N_2(x^1)\big), N_1(x^1)\big/N_2(x^1), N_2(x^1), \ldots, \\ &\qquad \Omega\big(T^G, N_1(x^G), N_2(x^G)\big), N_1(x^G)\big/N_2(x^G), N_2(x^G)\big) \end{aligned}$$

[13] This index should not be confused with the entropy measure analyzed in Theil (1972). Let p_i represent the percent of nonwhite students in school i, let $(1 - p_i)$ represent the percent of white students in school i, and let w_i represent the fraction of all students in the school district (white and nonwhite), who are in school i. Then the entropy measure for "n" schools in Theil (1972, p. 19) is, $\sum_{j=1}^{n} w_j[p_j \log \frac{1}{p_j} + (1 - p_j) \log \frac{1}{(1-p_j)}]$. This measure is neither a member of the above generalized entropy class, nor is it a RIMFO. In particular, it does not satisfy the homogeneity property (Property 1).

Since this property is a restricted version of P5 (all additive decomposable measures are aggregative, but not all aggregative measures are additive decomposable), denote the new property as P5′,

P5′. Additive Decomposability: $O(x)$ is additive decomposable.

Additive decomposable measures of segregation are quite useful. They can be used to partition occupational segregation in a manner similar to Schwarze's decomposition of income inequality in Germany. Schwarze used an additive decomposable measure of income inequality to decompose total income inequality into an East German component, a West German component, and a between-sector component. Using an additive decomposable measure of segregation, one can similarly decompose total segregation into a blue-collar component, a white-collar component, and a between-sector component, and assess how the different components changed through time. This is illustrated in Section 6.

For purposes of deriving a unique index, the advantage of additive decomposability is that it places restrictions on the monotonic increasing function $F(\cdot)$ in Theorem 1. If the measure of segregation is additive decomposable, then $F(\cdot)$ must be linear. This can be summarized in a result that is not only linked to Theorem 1, but for which there is a parallel result in the income inequality literature.[14]

Corollary to Theorem 1. *$O(x)$ is an additive decomposable RIMFO if and only if $O(x)$ is a positive multiple of $O_c(x)$ $0 < c < 1, x \in D$.*

Proof. See the Appendix.

The next property, denoted P6, places restrictions on the value of c. The value of c determines the effect of type 1 versus type 2 people on measured segregation. For example, consider the distributions x^* and y^*,

$$x^* = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} \qquad y^* = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$$

The only difference between these distributions is that the first group in x^* contains two type 1 people and one type 2 person, whereas the first group in y^* contains one type 1 person and two type 2 people. A generalized entropy index with $c = 0.5$ indicates that x^* and y^* have the same level of segregation ($O_c(x^*) = O_c(y^*)$ $= 0.0576$). A generalized entropy index with $c \neq 0.5$ indicates that x^* and y^* have different levels of segregation. For example, if c is 0.2, the generalized entropy index for x^* and y^* are 0.0570 and 0.0583, respectively. Because the generalized entropy index with $c = 0.5$ is insensitive to who is labeled type 1 or type 2, it receives special attention in what follows. Since it can be written as a product of square roots,

[14] See Shorrocks (1980, Theorem 5).

$$O(x) = O_{0.5}(x) = -\sum_{j=1}^{T} s_{2j}\left[(s_{1j}/s_{2j})^{0.5} - 1\right] = 1 - \sum_{j=1}^{T} \sqrt{(s_{2j})(s_{1j})}$$

Hutchens (2001) terms it the square root index.[15]

There may be reasons for an analyst to choose a value of c other than 0.5. Suppose that type 1 people are women and type 2 people are men. One could argue that the x^* distribution above constitutes a less egregious form of segregation than the y^* distribution, perhaps noting that male-dominated occupations like that in y^* tend to be associated with past, often legally sanctioned, barriers to entry (e.g., medical colleges that excluded women). By this argument, one would want $c < 0.5$. Obviously, the choice of c requires a value judgment.

Is there a way to determine whether one distribution is less segregated than another for any feasible value of c? Yes, plot segregation curves. If one segregation curve always lies above the other, then we know that the distribution associated with the higher curve is less segregated for any measure of segregation that satisfies properties 1–4 (i.e., any RIMFO). Since a generalized entropy measure satisfies properties 1–4, we know that, regardless of the chosen value of c, any generalized entropy measure will indicate that the distribution associated with the higher curve is the more equal distribution.

Suppose, however, that one is willing to make the value judgment that the index should be insensitive to whether men or women are labeled type 1 or 2. Not only is that a reasonable value judgment, but also there is ample precedent for it in the literature.[16] The idea can be stated as a formal property.

P6. Symmetry in Types: Let $\Gamma = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. If $y = \Gamma x$, then $O(x) = O(y)$.

A generalized entropy index with $c = 0.5$ obviously satisfies this property.

The seventh and final property limits the range of the index. This property simply states that the index should lie between zero and one. Thus follows:

P7. Range: $0 \leq O(x) \leq 1$.

The two most widely used measures of segregation—the dissimilarity index and the Gini index—satisfy this property. They lie on the unit interval, with zero representing complete integration and one representing complete segregation.

Properties P5′, P6, and P7 in combination with P1–P4, are sufficient to yield a unique index. Specifically, a numerical measure of segregation satisfies P1–P4, P5′–P7 if and only if it is a square root index. More formally we have the following:

[15] Hutchens (2001) introduces the square root index, and conjectures (without proof) that there may exist conditions under which the square root index is unique. This article presents both a set of conditions and a proof. There remains a good question about whether arguments underlying this index can be extended to a measure of inequality for more than two types of people, for example an index of the form,

$$1 - \sum_{j=1}^{T} \sqrt[K]{(s_{1j})(s_{2j})\cdots(s_{Kj})},$$

where K is the number of types of people (e.g., ethnic groups). That, however, will have to be a topic for future work.

[16] See Kakwani (1994), Chakravarty and Silber (1994), and Hutchens (2001).

THEOREM 2. *Let O: $D \to \mathbf{R}$ be a continuous RIMFO with a maximum value of* 1 *and a minimum value of* $\Omega(T, N_1, N_2) = 0$ *for each* $T > 1$, *and all* $N_1 > 0$, *and* $N_2 > 0$. *Then O is an additive decomposable RIMFO that satisfies symmetry of types* (P6) *if and only if O is a square root index, i.e.,*

$$O(x) = O_{0.5}(x) = -\sum_{j=1}^{T} s_{2j}\left[(s_{1j}/s_{2j})^{0.5} - 1\right] = 1 - \sum_{j=1}^{T} \sqrt{(s_{2j})(s_{1j})}$$

PROOF. See the Appendix.

Of course, a unique index must necessarily rest on strong—and thereby debatable—assumptions about the nature of inequality. Since the square root index is a member of the generalized entropy class, the critical literature on that class of measures applies.[17] In using a square root index, one should perhaps adopt the approach often taken in using a Theil index of income inequality. Although it yields useful decompositions, other indexes that do not satisfy additive decomposability (e.g., the Gini coefficient) can also be useful.

5. LINKS TO MEASURES OF INCOME INEQUALITY

It is reasonable to ask whether the argument underlying Theorem 2 also applies to measures of income inequality. The square root index is a generalized entropy index of segregation with the parameter c set to 0.5. From Shorrock's Theorem 5 we know that the analogous generalized entropy index of income inequality with $c = 0.5$ takes the form,[18]

$$I_{0.5}(v) = -\frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{v_i}{\bar{v}}\right)^{0.5} - 1\right) = 1 - \sum_{i=1}^{n}\sqrt{1/n}\sqrt{v_i/n\bar{v}} = 1 - \frac{1}{n\sqrt{\bar{v}}}\sum_{i=1}^{n}\sqrt{v_i} \tag{2}$$

Can the properties underlying Shorrocks' Theorem 5 be augmented with properties analogous to P5′, P6, and P7 so as to demonstrate that this index is the only index that satisfies the properties? The answer is "yes," but that at least one of the seven properties is likely to be controversial. In particular, the symmetry property (P6) is less compelling for a measure of income inequality than it is for a measure of segregation.

[17] For example, see the excellent discussion in Chapter A.5 of the Appendix by Foster and Sen in Sen (1997).

[18] This is a slightly altered version of the index in Shorrock's Theorem 5. That index was multiplied by 4.

The point can be made by considering how "equivalent" movements of type 1 and type 2 people alter a distribution of people that initially has no segregation.[19] Define a movement of δ_1 type 1 people from group i to j as "equivalent" to a movement of δ_2 type 2 people from group i to j if $\delta_2 = \delta_1(N_2/N_1)$. For example, let x be a distribution with two groups of men and women. Consistent with the definition of a distribution with no segregation, let the first group have four women and four men and the second group six women and six men. Thus follows:

$$x = \begin{bmatrix} 4 & 6 \\ 4 & 6 \end{bmatrix}$$

Now, suppose one woman is moved from group 1 to group 2 yielding the distribution,

$$y^{o} = \begin{bmatrix} 3 & 7 \\ 4 & 6 \end{bmatrix}$$

An equivalent movement of men from the initial x distribution yields

$$y^{oo} = \begin{bmatrix} 4 & 6 \\ 3 & 7 \end{bmatrix}$$

A numerical measure of segregation that exhibits the symmetry property (P6) will obviously indicate that y^{o} and y^{oo} exhibit the same level of segregation. More generally, given an initial distribution with no segregation and a numerical measure of segregation that satisfies P1–P7, equivalent movements have identical effects on segregation.

The same idea extends to measures of income inequality and equivalent movements of income and people. To see this, consider a population of 10 people with an aggregate income of \$10. In a manner analogous to the previous example, let the population be divided into two groups; the first contains four people who divide a total income of \$4, and the second contains six people who divide a total income of \$6. Assume these groups divide income equally such that each person

[19] Another way to make the point is to note that any vector of incomes can be converted to "groups" notation using the method described in footnote 11. Thus, P6 can be applied to any pair of income vectors. The problem is that although P6 is plausible in the segregation problem, it is controversial in the income inequality problem. For example, application of P6 to the income inequality problem implies that the income vectors $v_x = (2,1,1)$ and $v_y = (0.5,0.5,1,1)$ have the same level of income inequality. To see this, use the method in footnote 11 to convert v_x and v_y into "groups" notation; the resulting matrixes are x^* and y^* in the text. Although it is plausible to claim that x^* and y^* imply the same degree of segregation, the claim is less compelling for the income vectors v_x and v_y.

Yet another way to make the point is to note that two income vectors that satisfy P6 will have Lorenz curves that are images of each other. Specifically, if two vectors v_x and v_y are symmetric in the sense of P6, then take the Lorenz curve for v_x, draw a diagonal from the Northwest corner to the Southeast corner of the square, and rotate the Lorenz curve for v_x around that diagonal. The resulting Lorenz curve is that for v_y.

Note that even if one imposed P5′, P6, and P7 on the income inequality problem, the resulting index is not a measure of segregation since it is defined over the smaller domain D_+.

in a group receives that group's average income of $1. Since everyone receives the same income, there is no inequality in this initial distribution. Letting the income distribution over n people be denoted as $v = (v_1, v_2, \ldots, v_n)$, we initially have

$$v = (1, 1, 1, 1;\ 1, 1, 1, 1, 1, 1)$$

where the semicolon separates the first group from the second.

Given this, consider how "equivalent" movements of people and income alter a distribution of income that initially has no inequality. Let the movement of δ_1 dollars of income from group i to j be defined as "equivalent" to a movement of δ_2 people from group i to j if $\delta_2 = \delta_1(Y/n)$, where n is the total number of people in the population and Y is the total amount of income in the population. In a manner analogous to the above treatment of segregation, suppose $1 is moved from group 1 to group 2. As a result, the first group has four people sharing $3 (implying a per capita income of 3/4), and the second group has six people sharing $7 (implying a per capita income of 7/6). Letting w^{o} denote the new distribution, we have the following:

$$w^{\mathrm{o}} = (3/4, 3/4, 3/4, 3/4;\ 7/6, 7/6, 7/6, 7/6, 7/6, 7/6)$$

An equivalent movement of people from the initial distribution would involve movement of one person from group 1 to group 2 and no movement of income. In this case the first group would have three people sharing $4 (implying a per capita income of 4/3), and the second group would have seven people sharing $6 (implying a per capita income of 6/7). Call this new distribution w^{oo}, where

$$w^{\mathrm{oo}} = (4/3, 4/3, 4/3;\ 6/7, 6/7, 6/7, 6/7, 6/7, 6/7, 6/7)$$

Do w^{o} and w^{oo} have the same level of income inequality? Looking at them the answer is not at all obvious. If, however, one is willing to claim a property analogous to P6 and assert that a measure of income inequality should be symmetric in its treatment of equivalent movements of people and income, then the answer is "yes." Indeed, some numerical measures of income inequality exhibit that property. In particular, the square root index in Equation (2), an Atkinson index with $\varepsilon = 0.5$, and a Gini coefficient indicate that w^{o} and w^{oo} have the same level of income inequality. But other commonly used measures of income inequality do not exhibit the property. Included are Theil 1, Theil 2, the coefficient of variation, and an Atkinson index with $\varepsilon \neq 0.5$.[20] Thus, it is controversial to claim that the income distributions w^{o} and w^{oo} have the same level of inequality.

And therein lies the point. Because measures of segregation are closely linked to measures of income inequality, one can use an argument like that underlying Theorem 2 to show that the square root index of income inequality in Equation (2) is the only index that satisfies properties analogous to P1–P7. However, that argument is likely to be less compelling for a measure of income inequality than for a measure of segregation. This is because of P6. Specifically, it is not difficult to

[20] These measures of income inequality are fully described in Sen (1997).

TABLE 1
OCCUPATIONAL DISTRIBUTION BY GENDER, 1980, 1990, AND 2000

Occupation	1980 Percent of All Women	1980 Percent of All Men	1990 Percent of All Women	1990 Percent of All Men	2000 Percent of All Women	2000 Percent of All Men
Management	4.9	9.4	7.1	9.6	7.2	10.3
Farmers and farm managers	0.3	2.0	0.3	1.4	0.2	0.9
Business and financial operations	2.5	2.7	4.4	3.1	5.1	3.7
Computer and mathematical	0.4	0.8	0.9	1.4	1.6	3.1
Architects, engineers, and scientists	1.2	5.1	1.6	5.4	1.5	4.4
Community, social services, and legal	1.5	1.7	2.1	2.0	2.9	2.1
Education, training, and library	7.3	2.6	7.4	2.6	8.3	2.7
Arts, design, entertainment, sports, and media	1.4	1.4	1.8	1.7	1.9	1.8
Healthcare practitioner and technical	5.4	1.6	6.3	1.9	7.1	2.3
Healthcare support	3.7	0.4	3.5	0.4	3.8	0.4
Protective service	0.4	2.3	0.6	2.6	0.8	2.8
Food preparation, building, and grounds maintenance	9.8	6.0	8.6	6.6	9.1	7.9
Personal care and service	4.2	0.7	4.4	0.7	4.7	1.1
Sales and related	11.3	8.8	12.7	11.0	11.8	10.7
Office and administrative support	30.0	6.6	26.3	6.5	24.9	7.3
Farming, forestry, and fishing	0.7	2.3	0.6	2.5	0.3	1.1
Construction trades	0.3	9.7	0.4	9.7	0.4	10.1
Extraction	0.0	0.5	0.0	0.2	0.0	0.2
Installation, maintenance, and repair	0.3	6.5	0.3	6.2	0.4	6.8
Production	11.5	17.1	8.2	13.3	5.9	10.6
Transportation and material moving	2.7	11.8	2.4	11.0	2.2	9.8
Total	100.0	100.0	100.0	100.0	100.0	100.0
Indexes						
Square Root	0.166		0.147		0.138	
Gini	0.596		0.559		0.539	

convince people that y^{o} and y^{oo} have the same level of segregation. It *is* difficult to convince people that w^{o} and w^{oo} have the same level of income inequality.

6. AN APPLICATION

The ideas in this article can be illustrated with data on occupational segregation by gender. Table 1 presents data on the distribution of men and women across 21 occupations in 1980, 1990, and 2000. The data set was put together by Kim Weeden and utilized in Weeden (forthcoming).[21] Over the two decades women increased their share of employment in traditional male occupations (e.g., management)

[21] The 2000 Census changed the occupation classification scheme. At the time of writing, there did not exist a cross-walk permitting classification of the 1980 and 1990 occupation distributions in terms of the 2000 taxonomy. Professor Weeden managed to match "the detailed occupation titles from earlier schemes to a list of the detailed titles constituting each of the 21 aggregate occupations." She notes

and decreased their share in traditional female occupations (e.g., administrative support). As displayed in Figure 1, segregation curves derived from these data indicate reduced occupational segregation between 1980 and 2000.[22] Numerical measures of segregation that are in the RIMFO class will, in consequence, also register reduced occupational segregation. Since the square root index is a member of that class, it behaves accordingly, decreasing from 0.166 to 0.138.

Some argue that the observed trend toward decreased segregation in all occupations masks a trend toward increased segregation in occupational subgroups. For example, in their analysis of occupational segregation in Europe, Rubery et al. (1996) argue that whereas higher level occupations (e.g., management) are becoming less segregated, occupational segregation is either increasing or remaining constant in intermediate and low-level occupations. Is there evidence for that in the U.S. data? The methods introduced in this article can answer that question.

Table 2 rearranges the occupations in Table 1 into three sectors labeled high, intermediate, and low. The sectors are formed on the basis of 2001 average annual earnings.[23] For example, annual earnings in the intermediate sector range from \$27,600 in "production occupations" to \$39,130 in "education, training, and library occupations." Average earnings for the high sector lie above this range and average earnings for the low sector lie below it. Note that in 1980 the ratio of the female share to the male share (denoted as "q" in the table) was below one in the high and intermediate groups (0.76 and 0.66 respectively) and above one in the other low group (1.72). Note also that by 2000 the ratios for the low and high sector moved closer to 1, whereas that for the intermediate sector moved away from 1.

One way to analyze the data in Table 2 is to examine each sector's contribution to the overall index. Like any additive index, the square root index can be rewritten as the sum of its parts. Thus follows:

$$O_c(x^*) = \sum_{j\in \text{High}} \left[s_{2j} - \sqrt{s_{1j}s_{2j}}\right] + \sum_{j\in \text{Int}} \left[s_{2j} - \sqrt{s_{1j}s_{2j}}\right] + \sum_{j\in \text{Low}} \left[s_{2j} - \sqrt{s_{1j}s_{2j}}\right]$$

where the first sum is taken over the seven high occupations, the second over the seven intermediate occupations, and the third over the seven low occupations. The top panel in Table 3 presents the three sums. Clearly, most of the measured segregation arises in the "intermediate" sector. That is in part because the sector is large, and in part because the seven occupations within the intermediate sector have quite different ratios of females to males. Note also that between 1980 and 2000 the high sector's contribution declined dramatically, whereas the other two

that "given the unavoidable guesswork this matching process entailed, results based on these data should be interpreted cautiously." (Weeden, forthcoming, p. 147). Thus, although these data are ideal for illustrating applications of the square root index, results based on the data should not be viewed as the final word on trends in occupational segregation.

[22] The 1980 and 2000 segregation curves do not intersect. The 1990 segregation curve intersects with the 2000 curve.

[23] U.S. Bureau of Labor Statistics, Office of Employment and Unemployment Statistics, *The 2001 Occupational Employment Survey*. http://www.bls.gov/oes/2001/oes_nat.htm

TABLE 2

REARRANGING THE OCCUPATIONS IN TABLE 1 INTO SUBGROUPS WITH HIGH, LOW, AND INTERMEDIATE EARNINGS

	1980			1990			2000		
Occupation	(1) Percent of All Women	(2) Percent of All Men	(3) q	(4) Percent of All Women	(5) Percent of All Men	(6) q	(7) Percent of All Women	(8) Percent of All Men	(9) q
High									
Management	0.049	0.094	0.52	0.071	0.096	0.73	0.072	0.103	0.70
Business and financial operations	0.025	0.027	0.91	0.044	0.031	1.41	0.051	0.037	1.37
Computer and mathematical	0.004	0.008	0.54	0.009	0.014	0.61	0.016	0.031	0.51
Architects, engineers and scientists	0.012	0.051	0.24	0.016	0.054	0.30	0.015	0.044	0.34
Community, social services and legal	0.015	0.017	0.86	0.021	0.020	1.05	0.029	0.021	1.38
Arts, design, entertainment, sports and media	0.014	0.014	0.99	0.018	0.017	1.08	0.019	0.018	1.07
Healthcare practitioner and technical	0.054	0.016	3.31	0.063	0.019	3.24	0.071	0.023	3.13
Subtotal for high	0.173	0.228	0.76	0.241	0.252	0.96	0.273	0.276	0.99
Intermediate									
Education, training, and library	0.073	0.026	2.82	0.074	0.026	2.89	0.083	0.027	3.12
Protective service	0.004	0.023	0.18	0.006	0.026	0.22	0.008	0.028	0.29
Sales and related	0.113	0.088	1.28	0.127	0.110	1.15	0.118	0.107	1.11
Construction trades	0.003	0.097	0.03	0.004	0.097	0.04	0.004	0.101	0.04
Extraction	0.000	0.005	0.04	0.000	0.002	0.03	0.000	0.002	0.05
Installation, maintenance, and repair	0.003	0.065	0.05	0.003	0.062	0.05	0.004	0.068	0.06
Production	0.115	0.171	0.67	0.082	0.133	0.62	0.059	0.106	0.55
Subtotal for intermediate	0.313	0.474	0.66	0.296	0.457	0.65	0.276	0.438	0.63
Low									
Farmers and farm managers	0.003	0.020	0.15	0.003	0.014	0.20	0.002	0.009	0.19
Healthcare support	0.037	0.004	10.04	0.035	0.004	8.16	0.038	0.004	8.79
Food preparation, building and grounds maintenance	0.098	0.060	1.63	0.086	0.066	1.31	0.091	0.079	1.15
Personal care and service	0.042	0.007	5.76	0.044	0.007	6.00	0.047	0.011	4.14
Office and administrative support	0.300	0.066	4.55	0.263	0.065	4.03	0.249	0.073	3.39
Farming, forestry, and fishing	0.007	0.023	0.31	0.006	0.025	0.25	0.003	0.011	0.31
Transportation and material moving	0.027	0.118	0.23	0.024	0.110	0.22	0.022	0.098	0.22
Subtotal for low	0.514	0.298	1.72	0.463	0.292	1.59	0.451	0.286	1.58
Grand total	1.000	1.000	1.00	1.000	1.000	1.00	1.000	1.000	1.00

NOTES: q in column 3 is the ratio of the number in column 1 to that in column 2. Similarly, q in column 6 (9) is the ratio of the number in column 4 (7) to that in column 5 (8).

TABLE 3

AN ANALAYSIS OF SEGREGATION WHEN THE OCCUPATIONAL DISTRIBUTION IS DIVIDED INTO HIGH, INTERMEDIATE, AND LOW WAGE SECTORS: SECTORAL CONTRIBUTIONS AND DECOMPOSITION OF TOTAL SEGREGATION INTO WITHIN AND BETWEEN COMPONENTS

Sectoral Contribution to Total Segregation				
Year	All (Total)	High	Intermediate	Low
1980	0.166	0.043	0.148	−0.025
1990	0.147	0.019	0.144	−0.016
2000	0.138	0.016	0.148	−0.026
Square Root Indexes Computed Over the Detailed Occupations Within Each Sector				
Year	All (Total)	High	Intermediate	Low
1980	0.166	0.071	0.153	0.174
1990	0.147	0.055	0.150	0.163
2000	0.138	0.051	0.167	0.132
Decomposition				
Year	All (Total)	Within	Between	Between Percent
1980	0.166	0.141	0.025	0.150
1990	0.147	0.129	0.019	0.126
2000	0.138	0.119	0.019	0.135

sectors hardly changed their contribution. Obviously, the declining contribution of the high sector is the main reason for the decline in total measured segregation over the decade.

But what about the question of whether the overall trend toward decreased segregation masks a trend toward increased segregation in the low and intermediate occupational subgroups? That question can be answered by computing square root indexes for the three subgroups and then decomposing total segregation into within and between components. To that end, the middle panel in Table 3 presents square root indexes that are computed over the detailed occupations within each of the sectors. Note that the indexes for the high and low sectors decline between 1980 and 2000, whereas that for the intermediate sector declines between 1980 and 1990 and then increases. There is then some evidence of divergent trends, although it is not of the form described by Rubery, Fagan, and Maier in their European analysis. In these data occupational segregation is declining in both high-income and low-income occupations. The divergence occurs in the intermediate occupations, it is small, and it only occurs after 1990.

The lower panel of Table 3 summarizes. The entries in the column labeled "within" are weighted averages of the within-sector square root indexes in the middle panel.[24] The column labeled "between" indicates the value that the square

[24] As established in the proof of the corollary to Theorem 1, the square root index can be written, $O_{0.5}(x) = O_{0.5}(x^B) + \sum_{i=1}^{3} w_i O_{0.5}(x^i)$, where $w_i = \{N_1(x^i)/N_1(x)\}^{0.5}\{N_2(x^i)/N_2(x)\}^{0.5}$. For 2000, $O_{0.5}(x) = 0.138 = 0.019 + 0.051w_1 + 0.167w_2 + 0.132w_3$, where $w_1 = 0.275$, $w_2 = 0.348$, and $w_3 = 0.359$. Thus, the between component equals 0.019 and the within component equals $0.051w_1 + 0.167w_2 + 0.132w_3$. For 1990 the corresponding weights are $w_1 = 0.247$, $w_2 = 0.368$, and $w_3 = 0.367$, and for 1980 they are $w_1 = 0.199$, $w_2 = 0.385$, and $w_3 = 0.391$.

root index would take if men and women within each sector were redistributed across that sector's occupations such that the within-sector measure was zero.[25] Note that the 1980 sum of the "within" and "between" components equals the total, as does the 1990 and 2000 sum. The table indicates that the "within" component fell over the two decades whereas the "between" component fell between 1980 and 1990 and then remained constant between 1990 and 2000. The main message from this lower panel is similar to the message for the middle panel: In terms of occupational segregation by gender, the decade of the 1990s differed from the 1980s. In the 1980s occupational segregation declined within and between sectors. In the 1990s, the dominant trend was toward declining segregation, but the story is more complex. In particular, there was no decline in between-sector segregation.

Finally, since total measured segregation is the sum of the "between" and "within" components, one can determine what percent of total segregation takes the form of between-sector segregation. As indicated in the final column of the lower panel of Table 3, this percentage decreases from 15.0% in 1980 to 13.5% in 2000, with the 1990 figure diverging slightly from the trend. Since the three sectors represent different levels of earnings, this decrease suggests that the link between occupational segregation and earnings became weaker through time. Alternatively stated, occupational segregation is increasingly a matter of gender differences within occupational subgroups that have similar earnings and less a matter of gender differences between occupations that have disparate earnings.

7. CONCLUSION

This article considers the problem of analyzing inequality in the distribution of people across groups. It establishes that there is one and only one index—the square root index—that satisfies seven reasonable properties of a good segregation index. The key innovation in the article is to utilize concepts of aggregation and additive decomposability that were initially introduced in the analysis of income inequality. Although there have been previous axiomatic derivations of segregation indexes, this is the first proof of a unique index. The article concludes with an illustration: a decomposition of changes in occupational segregation by gender in the United States between 1980, 1990, and 2000.

The ideas in this article are part of a broader effort at bringing together the literature on measuring income inequality and the literature on measuring segregation. Although the two literatures have largely developed independently, they deal with conceptually similar problems. Similar methods, axioms, and theorems can be used to address both problems.

APPENDIX

A.1. *Proof of Theorem 1.*

Sufficiency: If $O(\mathrm{x})$ is a continuous aggregative RIMFO defined over all $x \in D$, and $O(\Omega(T, N_1, N_2)) = 0$ for each $T > 1$, and all $N_1 > 0$, and $N_2 > 0$, then there

[25] See the proof of the corollary to Theorem 1 for the formula for B.

exists a continuous, strictly increasing function F: $[0, \infty) \to \mathbf{R}$, with $F(0) = 0$, such that $F(O(x)) = O_c(x)$ (all $x \in D$) for some c where $0 < c < 1$.

PROOF.

Step 1. Some auxiliary results. Let $x \in D$. For any $i = 1, 2$ and $j = 1, 2, \ldots, T$ take a sequence $x_{ij}^k = [10^k x_{ij} + 1]/10^k$, $k = 1, 2, 3\ldots$ and denote x^k a matrix formed by replacing each element x_{ij} in x by x_{ij}^k. Obviously, $x^k \in D^1 \equiv \cup_{T=1}^{\infty} D_T^1$, where D_T^1 is the set of all $2 \times T$ matrices with positive real elements in its first and positive rational elements in its second row. Moreover, $x^k \to x$ as $k \to \infty$. Let y^k be a vector obtained by replacing each column j ($j = 1, 2, \ldots, T$) in matrix x by $10^k x_{2j}^k$ elements x_{1j}^k / x_{2j}^k. Thus, $y_i^k = x_{11}^k / x_{21}^k$, $i = 1, 10^k x_{21}^k, \ldots, y_i^k = x_{1T}^k / x_{2T}^k$, $i = 10^k \sum_{l=1}^{T-1} x_{2l}^k + 1, 10^k \sum_{l=1}^{T} x_{2l}^k$. Note that $y^k \in D_+$, so Foster's restatement of Shorrocks' Theorem 5 applies. Restating this theorem in terms of x^k instead of y^k we get that if I: $D^1 \to R$ is a continuous function and $I(\Omega(T, N_1, N_2)) = 0$ for each $T > 1$, and all $N_1 > 0$, and $N_2 > 0$, then I is an aggregative relative inequality measure iff there exists a continuous, strictly increasing function $G: [0, \infty) \to R$, with $G(0) = 0$, such that $G(I(z)) = I_c(z)$ (all $z \in D^1$) for some c, where

$$I_c(z) = \begin{cases} \frac{1}{c(c-1)} \sum_{j=1}^{T} s_{2j}(z)((s_{1j}(z)/s_{2j}(z))^c - 1), & c \neq 0, 1 \\ \sum_{j=1}^{T} s_{1j}(z) \ln(s_{1j}(z)/s_{2j}(z)), & c = 1 \\ \sum_{j=1}^{T} s_{2j}(z) \ln(s_{2j}(z)/s_{1j}(z)), & c = 0 \end{cases}$$

with $s_{1j}(z) = z_{1j}/N_1(z)$ and $s_{2j}(z) = z_{2j}/N_2(z)$. For any $x \in D$ define $I_c(x) = \lim_{k\to\infty} I_c(x^k)$ utilizing the above notation. Note that if $0 < c < 1$ then $I_c\,(x)$ is finite, but for the other values of c it is infinite when at least for one $j = 1, 2, \ldots, T$, $s_{1j}(x) = 0$ or $s_{2j}(x) = 0$.

Step 2. By assumption $O(x)$ is a continuous aggregative RIMFO defined over all $x \in D$. O is also a relative inequality measure for incomes (RIMFI) (footnote 4 discusses the RIMFI class), since (i) the domain of the RIMFI class is a subset of the domain of the RIMFO class ($D^1 \subset D$) and (ii) the properties of a RIMFO imply the properties of a RIMFI (cf. Hutchens 1991, Lemma 2). Reformulating the aggregative property for an income inequality measure similarly to what is done in step 1 for Shorrocks' Theorem 5 (see also the argument in footnote 11), and taking into account that the domain of the RIMFI class is a subset of the domain of the RIMFO class, we conclude that O is also an aggregative RIMFI.

Step 3. From step 2 it follows that we can take $I(\cdot) \equiv O(\cdot)$ in the reformulated Shorrocks' Theorem 5. Since O and F^{-1} are continuous on their respective domains, we conclude that $I_c\,(x)$ is also continuous for any $x \in D$, from which follows that $0 < c < 1$. Defining $F(\cdot) \equiv -c(c-1)G(\cdot)$ finishes the proof of sufficiency. ■

Necessity: If $O(x)$ is a continuous RIMFO defined over all $x \in D$, and $O(\Omega(T, N_1, N_2)) = 0$ for each $T > 1$, and all $N_1 > 0$ and $N_2 > 0$, and there exists a continuous, strictly increasing function F: $[0, \infty) \to \mathrm{R}$, with $F(0) = 0$, such that $F(O(\mathrm{x})) = O_c(x)$ (all $x \in D$) for some $0 < c < 1$, then O is aggregative.

PROOF. Let the distribution of type 1 and type 2 people in the set of T occupations be partitioned into two mutually exclusive and exhaustive subsets containing T^1 and T^2 occupations such that $x = (x^1, x^2)$. We need to show that there exists an "aggregator" function A that is continuous and strictly increasing in $O_c(x^1)$ and $O_c(x^2)$ with

$$\begin{aligned} O_c(x) &= O_c(x^1, x^2) \\ &= A\big(O_c(x^1), N_1(x^1)/N_2(x^1), N_2(x^1); O_c(x^2), N_1(x^2)/N_2(x^2), N_2(x^2)\big) \end{aligned}$$

Consider the function Z

$$\begin{aligned} Z = {} & O_c(x^1)\big(N_1(x^1)/N_1(x)\big)^c\big(N_2(x^1)/N_2(x)\big)^{1-c} \\ & + O_c(x^2)\big(N_1(x^2)/N_1(x)\big)^c\big(N_2(x^2)/N_2(x)\big)^{1-c} \\ & + 1 - \big(N_1(x^1)/N_1(x)\big)^c\big(N_2(x^1)/N_2(x)\big)^{1-c} \\ & - \big(N_1(x^2)/N_1(x)\big)^c\big(N_2(x^2)/N_2(x)\big)^{1-c} \end{aligned}$$

Z is continuous and strictly increasing in $O_c(x^1)$ and $O_c(x^2)$, and $O_c(x) = O_c(x^1{,}x^2) = Z$. Moreover, Z can be written as a function of the form $A(O_c(x^1)$, $N_1(x^1)/N_2(x^1)$, $N_2(x^1)$; $O_c(x^2)$, $N_1(x^2)/N_2(x^2)$, $N_2(x^2))$. Thus O_c is aggregative. Finally, since F is continuous and strictly increasing, there exists an inverse function F^{-1} that is also continuous and strictly increasing. Then $O(x) = F^{-1}[O_c(x)] = F^{-1}[A(O_c(x^1), N_1(x^1)/N_2(x^1), N_2(x^1); O_c(x^2), N_1(x^2)/N_2(x^2), N_2(x^2))] = F^{-1}[A(F(O(x^1)), N_1(x^1)/N_2(x^1), N_2(x^1); F(O(x^2)), N_1(x^2)/N_2(x^2), N_2(x^2))]$. So, O is aggregative. This finishes the proof of necessity and Theorem 1 as a whole. ■

A.2. *Proof of Corollary to Theorem 1.* It is easiest to begin with the following lemma.

LEMMA A.1. *O_c is additive decomposable.*

PROOF. The proof of Theorem 1 establishes that O_c is aggregative. From that proof note that

$$\begin{aligned} O_c(x) &= O_c(x^1, x^2) = Z = \gamma_1 O_c(x^1) + \gamma_2 O_c(x^2) + \varphi, \quad \text{where,} \\ \gamma_i &= (N_1(x^i)/N_1(x))^c(N_2(x^i)/N_2(x))^{1-c}, \quad i = 1, 2 \\ \varphi &= 1 - (N_1(x^1)/N_1(x))^c(N_2(x^1)/N_2(x))^{1-c} \\ &\quad - (N_1(x^2)/N_1(x))^c(N_2(x^2)/N_2(x))^{1-c} \end{aligned}$$

Then O_c is additive decomposable. ■

Given Lemma A.1, a proof of the corollary to Theorem 1 proceeds as follows.

Sufficiency: If O is an additive decomposable RIMFO, then $O(x)$ is a positive multiple of $O_c(x)$, $0 < c < 1$, $x \in D$.

PROOF. Assume that x takes the form of $N_1(x)$ type 1 people, and $N_2(x)$ type 2 people distributed over $2K$ occupations. Partition the occupations into K mutually exclusive and exhaustive subsets, each containing two occupations such that $x = (x^1, x^2, \ldots, x^K)$, where x^i denotes the distribution of people across the two occupations in subset i. In addition, let $N_1(x^i) = N_1(x)/K = N_{1o}$; $N_2(x^i) = N_2(x)/K = N_{2o}$ for all $i = 1, 2, \ldots, K$.

If $O(x)$ is additive decomposable, then Theorem 1 and Lemma A.1 imply

$$F(O(x)) = F\left(\sum_i w_i O(x^i) + \mathrm{B}\right) = O_c(x) = \sum_i \gamma_i O_c(x^i) + \varphi$$

Given that $N_1(x^i) = N_{1o}$; $N_2(x^i) = N_{2o}$ for all $i = 1, 2, \ldots, K$, it follows that $\varphi = 0$ and $\gamma = 1/K$. Moreover, if $O(x^i) = O(\Omega(2, N_{1o}, N_{2o})) = 0$ then $O(x) = O(\Omega(2K, N_1(\mathrm{x}), N_2(\mathrm{x})) = 0$, implying $B = 0$. Given $B = 0$ and $w_i = w_o = w(N_{1o}, N_{2o})$, $i = 1, K$, then $O(x) = w_o \Sigma_i O(x^i)$. Since this must be valid for $O(x^i) = O(x)$, $i = 1, K$, it follows that $w_o = 1/K$. Then $F(O(x)) = F(\Sigma_\mathrm{i} O(x^i)/K) = O_c(x) = \Sigma_i \mathrm{O}_c(x^i)/K = \Sigma_i F(O(x^i))/K$.

Reallocate the people in each subset $i = K - J + 1, K - J + 2, \ldots, K$ across its two occupations such that $N_1(x^i)$ and $N_2(x^i)$ are fixed at N_{1o} and N_{2o} and there is no segregation, i.e., reallocate the people such that $O(x^i) = O(\Omega(2, N_{1o}, N_{2o})) = 0$. Since $O_c((\Omega(2, N_{1o}, N_{2o})) = 0$, and since this reallocation does not alter $w_i = \gamma_i = 1/K$ for all $i = 1, 2, \ldots, K$, it follows that

$$F\left(\sum_{i=1}^{K-J} O(x^i) \Big/ K\right) = \sum_{i=1}^{K-J} F(O(x^i)) \Big/ K \quad \text{for all } i = 1, 2, \ldots, K$$

Note that this is true for any $x \in D$, subject only to $T = 2K$. So, some of the x^i may be identical, and thus $F(\lambda O(x)) = \lambda F(O(x))$ for any positive rational λ. Since F is continuous and the set of rational numbers is everywhere dense in the set of real numbers, we can conclude that F is homogeneous of degree one. Taking into account that F is a function of one real variable, it then follows that F is a linear function of its argument. Moreover, given that F is strictly increasing, it follows that $F(O(x)) = bO(x)$ for some positive constant b. Thus, $O(x)$ is a positive multiple of $O_c(x)$, $0 < c < 1$, $x \in D$. ■

Necessity: If O is a positive multiple of O_c than it is an additive decomposable RIMFO.

PROOF. The proof of necessity in Theorem 1 and Lemma A.1 establish that O_c is an additive decomposable RIMFO. Then it follows trivially that if O is positive multiple of O_c, then O is an additive decomposable RIMFO. ■

To prove Theorem 2, it is easiest to begin with two lemmas, labeled A.2 and A.3.

LEMMA A.2. *A generalized entropy index of occupational segregation of the form,*

$$O_c(x) = -\sum_{j=1}^{T} s_{2j}\left[(s_{1j}/s_{2j})^c - 1\right] \quad 0 < c < 1$$

satisfies P6 (symmetry in types) if and only if $c = 0.5$.

PROOF. Let $\Gamma = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. By P6, for any $x \in D$, $O(x) = O(\Gamma(x))$. Since $O_c(x) = F(O(x))$, we get

$$\begin{aligned} 0 &= O_c(x) - O_c(\Gamma(x)) \\ &= \sum_{j=1}^{T} \sqrt{s_{1j}(x)s_{2j}(x)}\left[(s_{1j}(x)/s_{2j}(x))^{c-1/2} - (s_{2j}(x)/s_{1j}(x))^{c-1/2}\right] \end{aligned}$$

This is true for any $x \in D$, which means that the expression in brackets is equal to zero for any $x \in D$, and thus $c = 1/2$. This ends the proof of Lemma A.2. ■

The next lemma follows immediately from Theorem 1 and Lemma A.2.

LEMMA A.3. *Let* $O: D \to \mathbf{R}$ *be a continuous RIMFO with* $O(\Omega(T, N_1, N_2)) = 0$ *for each* $T > 1$, *and all* $N_1 > 0$, *and* $N_2 > 0$. *Then O is an aggregative RIMFO that satisfies P6 iff there exists a continuous, strictly increasing function* $F: [0, \infty) \to \mathbf{R}$, *with* $F(0) = 0$, *such that* $F(O(x)) = O_c(x)$ *(all* $x \in D$*) for* $c = 0.5$, *i.e.*,

$$F(O(x)) = O_{0.5}(x) = -\sum_{j=1}^{T} s_{2j}\left[(s_{1j}/s_{2j})^{0.5} - 1\right] = 1 - \sum_{j}^{T} \sqrt{s_{1j}s_{2j}}$$

A.3. *Proof of Theorem 2.*

Sufficiency: Let $O: D \to \mathbf{R}$ be a continuous additive decomposable RIMFO with a maximum value of 1 and a minimum value of $O(\Omega(T, N_1, N_2)) = 0$ for each $T > 1$, and all $N_1 > 0$, and $N_2 > 0$. In addition, let O satisfy P6. Then O is a square root index.

PROOF. From Lemma A.3, if O is an aggregative RIMFO that satisfies P6, then $F(O(\cdot)) = O_{0.5}(\cdot)$, where $O_{0.5}$ is a square root index. If O is additive decomposable, than by the corollary to Theorem 1, O is a positive multiple of $O_{0.5}$. Then $O = \alpha O_{0.5}$, where α is a positive number. Since the maximum value of the square root index is 1, and since from P7, the maximum value of O is 1, α must equal 1. Then $O = O_{0.5}$. ■

Necessity: If O is a square root index, then O is an additive decomposable RIMFO that satisfies P6.

PROOF. From Lemma A.3, the square root index is an aggregative RIMFO that satisfies P6. From Lemma A.1 it follows that it is additive decomposable. ■

REFERENCES

ANAND, S., *Inequality and Poverty in Malaysia: Measurement and Decomposition* (London: Oxford University Press, 1983).

CHAKRAVARTY, S. R., AND J. SILBER, "Employment Segregation Indices: An Axiomatic Characterization," in W. Eichorn, ed., *Models and Measurement of Welfare and Inequality* (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag, 1994), pp. 912–20.

DUNCAN, O. D., AND B. DUNCAN, "A Methodological Analysis of Segregation Indices," *American Sociological Review* 20 (1955): 210–17.

FLÜCKIGER, Y., AND J. SILBER, *The Measurement of Segregation in the Labor Force* (Heidelberg, New York: Physica-Verlag, 1999).

FOSTER, J., "Inequality Measurement," in H. Peyton Young, ed., *Fair Allocation* (Providence: American Mathematical Society, 1985), pp. 31–68.

HUTCHENS, R. M., "Segregation Curves, Lorenz Curves, and Inequality in the Distribution of People across Occupations," *Mathematical Social Sciences* 21 (February 1991), 31–51.

——, "Numerical Measures of Segregation: Desirable Properties and Their Implications," *Mathematical Social Sciences* 42 (July 2001), 13–29.

KAKWANI, N. C., "Segregation by Sex: Measurement and Hypothesis Testing," in S. Neuman and J. Silber, eds., *Inequality in Labor Markets: The Economics of Labor Market Segregation and Discrimination, Research on Economic Inequality* (Greenwich, CT and London, England: JAI Press, 1994), pp. 1–26.

MOOKHERJEE, D., AND A. SHORROCKS, "A Decomposition Analysis of the Trend in U.K. Income Inequality," *The Economic Journal* 92 (December 1982), 886–902.

PYATT, G., "On the Interpretation and Disaggregation of Gini Coefficients," *The Economic Journal* 86 (June 1976), 243–55.

RUBERY, J., C. FAGAN, and F. MAIER, "Occupational Segregation, Discrimination, and Equal Opportunity," in G. Schmid, J. O'Reilly, and K. Schomann, eds., *International Handbook of Labour Market Policy and Evaluation* (Cheltenham: Edward Elger, 1996).

SCHWARZE, J., "How Income Inequality Changed in Germany Following Reunification: An Empirical Analysis Using Decomposable Inequality Measures," *Review of Income and Wealth* (March 1996), 1–12.

SEN, A., *On Economic Inequality*, Expanded Edition (Oxford: Clarendon Press, 1997).

SHORROCKS, A. F., "The Class of Additively Decomposable Inequality Measures," *Econometrica* (April 1980), 613–25.

——, "Inequality Decompositions by Population Subgroups," *Econometrica* 48 (November 1984), 1369–86.

——, "Aggregation Issues in Inequality Measurement" in W. Eichorn, ed., *Measurement in Economics* (New York: Physica-Verlag, 1988), 429–52.

THEIL, H., *Statistical Decomposition Analysis* (Amsterdam: North-Holland, 1972).

WATTS, M., "Occupational Gender Segregation: Index Measurement and Econometric Modeling," *Demography* 35 (November 1998), 489–96.

WEEDEN, K. A., "Profiles of Change: Sex Segregation in the United States, 1910–1990," in M. Charles and D. B. Grusky, eds., *Occupational Ghettos: The Worldwide Segregation of Men and Women* (Stanford, CA: Stanford University Press), forthcoming.