Measuring Segregation*

David M. Frankel — Iowa State University

Oscar Volij — Iowa State University

May 19, 2005

Abstract

This paper gives a rigorous foundation for the statement "city A is more segregated than city B." We define segregation to be the tendency of groups (e.g., ethnic groups, religious groups) to be distributed differently across locations (neighborhoods, schools, occupations, etc.). We propose a set of basic properties that any reasonable segregation index should have. We then show that a unique segregation index satisfies these basic properties. This index can sometimes be sensitive to the introduction of small new demographic groups. We also study an alternative index that is less sensitive to such changes. It is the unique index that satisfies a slightly different set of basic properties. Our analysis applies to any finite number of demographic groups.

Journal of Economic Literature Classification Numbers: C43, C81, D63.

Keywords: Segregation, segregation indices.

---

*Email addresses: dfrankel@econ.iastate.edu; volij@iastate.edu. We thank seminar participants at University of Haifa, University of Iowa, Rice University, the 2004 Summer Meetings of the Econometric Society, and the Murat Sertel Memorial Conference on Economic Theory.

1 Introduction

Segregation is a pervasive social issue. The segregation of men and women into different occupations helps explain the gender gap in earnings.[1] The continued racial segregation of schools appears to contribute to low educational achievement among minorities.[2] Residential segregation between blacks and whites has been blamed for black poverty, high black mortality, and increases in prejudice among whites.[3] In other contexts, segregation is viewed more positively. The formation of homogeneous living areas has been discussed as a solution to highly polarized conflicts in the Middle East, Yugoslavia, and elsewhere.

But what exactly is segregation? An influential paper by James and Taeuber [11] defines it as "the differences in the distribution of social groups, such as blacks and whites, among units of social organization such as schools." Massey and Denton [15] discuss five different dimensions of segregation. The first, evenness, agrees with James and Taeuber's definition:

> Groups may live apart from one another and be "segregated" in a variety of ways. Minority members may be distributed so that they are over-represented in some areas and underrepresented in others, varying on the characteristic of evenness.

Their other dimensions are isolation from the majority group, concentration in a small area, centralization in the urban core, and clustering in a contiguous enclave.

This paper will focus on Massey and Denton's first dimension of segregation: the tendency of demographic groups to have different distributions across neighborhoods.

---

[1] See Cotter et al [3], Lewis [13], and Macpherson and Hirsh [14].

[2] See Meldrum and Eaton [16], Orfield [18], and Schiller [21].

[3] See Cutler and Glaeser [5], Collins and Williams [2], and Kinder and Mendelberg [12], respectively.

Formally, we define a segregation ordering as a total order on cities: a ranking from most segregated to least segregated. We propose a set of basic properties that any reasonable segregation ordering should have. We then prove that there is a unique segregation ordering with these properties (Theorems 1, 2).

This ordering is represented by a simple index: one minus the sum, over all neighborhoods, of the geometric means of the percentages of each group who live in the neighborhood. For example, suppose 40% of blacks and 10% of whites live in neighborhood A while 60% of blacks and 90% of whites live in neighborhood B. The index is $1 - \sqrt{.4 * .1} - \sqrt{.6 * .9} = 0.065$.

This index has an important property. It is invariant to changes in the size of one group that preserve that group's distribution across locations. For instance, if the number of blacks in every neighborhood of a city is doubled, the index is unchanged. This property makes the index suitable for studying changes over periods in which the growth rates of the demographic groups differed from one another.

The index can sometimes be sensitive to the introduction of small new demographic groups. We also study an alternative index that is less sensitive to such changes (Theorem 3). It is the unique index that satisfies a slightly different set of basic properties. This alternative index is not always invariant to changes in the size of one group even if they preserve the group's distribution across locations. But it may be more suitable to studying changes over periods in which the demographic schema was redefined. One example is the introduction of tiny, mixed-race groups in the 2000 U.S. Census.

The papers that are most closely related to this paper are Hutchens [8, 9]. These papers study the measurement of segregation in the case of two demographic groups. Hutchens [8] proves that a family of indices satisfy a set of basic properties. Hutchens [9] strengthens one axiom and obtains a unique segregation index, which is the index produced by our first axiomatization (Theorem 1) in the case of two demographic groups.

While we assume properties of the underlying segregation ordering, Hutchens follows the inequality literature (e.g., Shorrocks [22, 23]) by imposing restrictions directly on the segregation index. Some of these restrictions are cardinal in nature and do not have a natural translation into properties of segregation orderings. For instance, it is not clear why a segregation ordering should be represented by an index that is additively separable over neighborhoods. Making an analogy with decision theory, while one may or may not accept the independence axiom of expected utility, it has a clear interpretation in terms of preferences. But the restriction that the utility representation be linear in probabilities has no such interpretation.

Another related paper is Echenique and Fryer [7]. Echenique and Fryer use data on individuals' social networks to measure the strength of an individual's isolation from members of other demographic groups. The resulting "spectral segregation index" is most closely related to Massey and Denton's second dimension of segregation: isolation from the majority group. Each group has a spectral segregation index that measures the extent to which that group is isolated from the other groups.

The rest of the paper is organized as follows. After setting up some basic notation in Section 2, we introduce the notion of a segregation ordering in Section 3 and provide some known examples of segregation indices that represent various orderings. Section 4 proposes some properties that a satisfactory segregation ordering should satisfy, and Section 5 shows that a unique ordering satisfies them all. The alternative segregation index in which small groups carry little weight is axiomatized in Section 5.3.

2 Notation

Definition 1 *A city consists of*

- *A nonempty and finite set of demographic groups* G
- *A nonempty and finite set of neighborhoods* N

- *For each group $g \in G$ and for each neigborhood $n \in N$, a nonnegative number T_g^n (the number of members of group g that reside in neighborhood n) such that for all $g \in G$, the number of members of group g is positive, i.e., $\sum_{n \in N} T_g^n > 0$.*

When the set of groups is understood, we will sometimes use the following simpler notation to discuss examples. The expression $\langle (1,2), (3,1) \rangle$ denotes a city with two groups (e.g., blacks and whites) and two neighborhoods. The first neighborhood, $(1,2)$, contains one black and two whites; the second, $(3,1)$, contains three blacks and one white. The order of the neighborhoods does not matter; e.g., $\langle (1,2), (3,4) \rangle$ can also be written $\langle (3,4), (1,2) \rangle$.

The following notation will be useful:

$$\begin{aligned}
T_g &= \sum_{n \in N} T_g^n\text{: the number of members of group } g \text{ in the city} \\
T^n &= \sum_{g \in G} T_g^n\text{: the total population of neighborhood } n \\
T &= \sum_{g \in G} T_g\text{: the total population of the city} \\
P_g &= \frac{T_g}{T}\text{: the proportion of city residents who are in group } g \\
p_g^n &= \frac{T_g^n}{T^n} \text{ (for } T^n > 0\text{): the proportion of residents of } n \text{ who are in } g \\
t_g^n &= \frac{T_g^n}{T_g} \text{ (for } T_g > 0\text{): the proportion of members of } g \text{ who live in } n
\end{aligned}$$

For any city $X = \left\langle (T_g^n)_{g \in G} \right\rangle_{n \in N}$, we will denote the set of neighborhoods of X by $N(X)$ and the set of demographic groups of X by $G(X)$.

The *group distribution of a city* X is the vector $(P_g)_{g \in G}$ of proportions of the city's residents who are in each group. The *group distribution of a nonempty neighborhood* n is the vector $\left(p_g^n\right)_{g \in G}$ of proportions of the neighborhood's residents who are in each group.

For any city X and any nonnegative constant c, cX denotes the city that results from multiplying the number of members of each group in each neighborhood of X by c: $cX = \left(cT_g^n\right)_{g \in G, n \in N}$. For example, if $X = \langle (1,2), (3,4) \rangle$, then $2X = \langle (2,4), (6,8) \rangle$.

For any city X and any vector of positive scalars $\overrightarrow{\alpha} = (\alpha_g)_{g \in G(X)}$, $\overrightarrow{\alpha} * X$ denotes the city in which the number of members of group g in neighborhood n is $\alpha_g T_g^n$. For example, if $X = \langle (1,2), (3,4) \rangle$, and $\overrightarrow{\alpha} = (1,2)$, then $\overrightarrow{\alpha} * X = \langle (1,4), (3,8) \rangle$.

For any two cities X and Y, $X \uplus Y$ denotes a city whose neighborhoods consist of the neighborhoods of X and Y. For instance, if $X = \langle (1,2), (3,4) \rangle$, and $Y = \langle (1,2) \rangle$ where the two cities contain the same two groups (e.g., blacks and whites), then $X \uplus Y = \langle (1,2), (1,2), (3,4) \rangle$.[4]

Neighborhood n in city X is *representative* if the demographic distributions of n and X are the same: if $p_g^n = P_g$ for all $g \in G$. A neighborhood that is not representative of the city is said to be *unrepresentative*.

3 Segregation orderings, and their measures

We will sometimes restrict attention to some special classes of cities. For example, the class of cities with exactly $K \geq 1$ demographic groups will be denoted $\mathcal{C}_K$. A *segregation ordering* $\succcurlyeq$ on a class of cities is a complete and transitive binary relation on that set of cities. We interpret $X \succcurlyeq Y$ to mean "city X is at least as segregated as city Y." The relations $\sim$ and $\succ$ are derived from $\succcurlyeq$ in the usual way.[5]

A related concept is the segregation index. A *segregation index* is a function that assigns a nonnegative number to each city in a class. Any segregation index induces a segregation ordering: given a segregation index S, the *associated segregation order* is defined by $X \succcurlyeq Y \Leftrightarrow S(X) \geq S(Y)$.

We depart from the literature by imposing axioms not on the segregation index but on the underlying segregation ordering. These approaches are not equivalent. First, a segregation ordering may be represented by more than one index: if an index

[4]Note that $G(X \uplus Y) = G(X) \cup G(Y)$. In the example, $G(X) = G(Y)$, but this need not be the case.

[5]That is $X \sim Y$ if both $X \succcurlyeq Y$ and $Y \succcurlyeq X$; $X \succ Y$ if $X \succcurlyeq Y$ but not $Y \succcurlyeq X$.

represents a given ordering, then so does any increasing function of the index. Second, there are segregation orderings that are not captured by any index. Examples are easy to construct: given two segregation indices S_1 and S_2, consider the ordering $\succcurlyeq$ defined as follows: $X \succcurlyeq Y$ if either (a) $S_1(X) > S_1(Y)$ or (b) $S_1(X) = S_1(Y)$ and $S_2(X) \geq S_2(Y)$. That is, S_1 is applied first and S_2 is used to break ties. In general, there will not be a single segregation index that represents this ordering.

3.1 Examples of segregation indices

We now discuss several examples of segregation indices.

Multigroup Atkinson Index This index is derived from the measure of income inequality of the same name (Atkinson [1]). It is given by

$$A(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G(X)} t_g^n \right)^{\frac{1}{|G(X)|}} \tag{1}$$

When X contains exactly two nonempty groups, this index gives the same segregation ordering as the usual Atkinson index with parameter 1/2 (Massey and Denton [15, p. 286]).[6]

Weighted Multigroup Atkinson Index This is a weighted version of (1).

$$A^W(X) = 1 - \sum_{n \in N(X)} \prod_{g \in G(X)} \left(t_g^n \right)^{P_g(X)} \tag{2}$$

$A(X)$ equals one minus the sum of the *unweighted* geometric averages of the group proportions in each neighborhood. In contrast, $A^W(X)$ equals one minus the sum of the *weighted* geometric averages of the group proportions in each neighborhood, where the weight assigned to a group equals the proportion of city residents who

[6] One can show that the usual Atkinson index with parameter 1/2 equals $1 - (1 - A(X))^2$, which is an increasing transformation of A.

belong to that group. While $A(X)$ treats each group the same regardless of its size, $A^W(X)$ gives more weight to larger groups.

Multigroup Dissimilarity Index This is a multigroup version of the Index of Dissimilarity:[7]

$$D(X) = \sum_{n \in N(X)} f(t^n) \text{ where } f(t^n) = \frac{1}{|G(X)|} \sum_{g \in G(X)} \left| t_g^n - \frac{\sum_{g' \in G(X)} t_{g'}^n}{|G(X)|} \right| \quad (3)$$

In the case of two groups, this measures the proportion of either group who would have to change neighborhoods in order to attain complete integration: for every neighborhood to be representative of the city. The Index of Dissimilarity was first discussed by Jahn et al [10]. It was used by Cutler, Glaeser, and Vigdor [6] to measure the evolution of segregation in American cities.

Entropy This index, adapted from the information theory literature, was proposed by Theil and Finizza [26]. In general, the entropy of a random variable is a measure of how much one typically learns when one observes the value of the variable. The entropy of a city is defined as one minus the ratio of the within-neighborhood entropy to the total entropy:

$$H(X) = 1 - \frac{H_{\text{within}}}{H_{\text{total}}} \quad (4)$$

where

$$H_{\text{within}} = \sum_{n \in N(X)} \frac{T^n}{T} H_n \quad (5)$$

$$H_n = \sum_{g \in G(X)} p_g^n \log_2 \left(\frac{1}{p_g^n} \right) \quad (6)$$

$$H_{\text{total}} = \sum_{g \in G(X)} P_g \log_2 \left(\frac{1}{P_g} \right) \quad (7)$$

[7] Equation (3) is just one possible way to generalize the index of Dissimilarity to multiple groups; see Morgan [17] and Sakoda [20] for an alternative.

The total entropy in the city, H_{total}, is the average information conveyed when one is told the race of a randomly chosen city resident. The within-neighborhood entropy, H_{within}, is the average information conveyed when one is told the race of random city resident, if one initially knows the resident's neighborhood. This is just the weighted average of the entropy H_n of each neighborhood n. If there is complete segregation, an agent's neighborhood perfectly signals her race, so H_{within} is zero; the entropy index equals one. If there is complete integration, an agent's neighborhood conveys no information about her race, so H_{within} equals H_{total}: the entropy index equals zero.

4 Axioms

We first propose a number of properties that, in our view, a segregation ordering should satisfy. Later we characterize the family of indices that satisfy all these properties.

Denote the class of all cities by $\mathcal{C}^A = \bigcup_{K=1}^{\infty} \mathcal{C}_K$. Sometimes we will refer to a transformation $\tau : \mathcal{C}^A \to \mathcal{C}^A$ of cities. We will say that a class of cities is closed under the transformation τ if for all cities X in the class, the transformed city, $\tau(X)$ is also in the class. Let $\mathcal{C}$ be a class of cities that is closed under the transformation τ, and let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$. We will say that the segregation of a city is unaffected by a transformation τ if the application of τ does not affect the city's location in the segregation ordering: if $X \sim \tau(X)$ for all cities $X \in \mathcal{C}$. Many of our axioms state that a segregation ordering on some class $\mathcal{C}$ is invariant to a given type of transformation. Unless otherwise specified, we will always assume the class $\mathcal{C}$ to be closed under this transformation.

We state our axioms with respect to a general class of cities $\mathcal{C}$. Later we will apply them to specific classes. The first axiom states that the segregation of a city is unaffected by relabeling the demographic groups (e.g., by calling whites "blacks" and blacks "whites"). For any permutation $\sigma : G(X) \to G(X)$, let $\sigma(X)$ be the city in which the number of members of group g in each neighborhood $n \in \sigma(X)$ equals the number of members of group $\sigma(g)$ who reside in the corresponding neighborhood n of X. For example, if X has two groups, b and w, then the permutation $\sigma(b) = w$, $\sigma(w) = b$ swaps the labels of the two races. In this case, σ applied to the city $X = \langle(1,2),(3,1)\rangle$ yields $\sigma(X) = \langle(2,1),(1,3)\rangle$.

Group Symmetry (GS) The segregation in a city is unaffected by relabeling the demographic groups: for any city $X \in \mathcal{C}$ and any permutation $\sigma : G(X) \to G(X)$, $X \sim \sigma(X)$.

The next axiom states that the size of a group does not affect the segregation of a city. For example, the city $\langle(1,2),(3,2)\rangle$ is as segregated as the city $\langle(1,1),(3,1)\rangle$. This is one of the five requirements that Jahn *et al* [10] say a satisfactory measure of segregation should satisfy.[8]

Scale Invariance (SI) The segregation in a city is unaffected if the number of agents of a given demographic group is multiplied by the same nonzero factor in all neighborhoods: for any city $X \in \mathcal{C}$ and any positive scalars $\overrightarrow{\alpha} = (\alpha_g)_{g \in G(X)}$, $X \sim \overrightarrow{\alpha} * X$.

---

[8] Jahn *et al* [10] write:

> A satisfactory measure of ecological segregation should (1) be expressed a single quantitative value so as to facilitate such statistical procedures as comparison, classification, and correlation; (2) be relatively easy to compute; (3) not be distorted by the size of the total population, the proportion of Negroes, or the area of a city; (4) be generally applicable to all cities; and (5) differentiate degrees of segregation in such a way that the distribution of intmediate scores cover most of the possible range between the extremes of 0 and 100.

Property (3) is Scale Invariance.

For instance, suppose we are interested in studying the changing pattern of segregation between whites, blacks, and Asians in a given city over time. Suppose the black population grew faster than the populations of the other groups, but each group retained its original distribution across neighborhoods. Since the tendency of each group to live apart from the other groups has not changed, one may desire a segregation measure that remains constant. Scale Invariance guarantees that this will be the case. In other applications, such as when the number and definition of the groups is changing over time, Scale Invariance might be less desirable. We consider an alternative axiom in section 5.3.

The next axiom states that if a neighborhood is divided into two neighborhoods, then segregation is unaffected if either the two new neighborhoods have the same group distribution or one of the new neighborhoods is empty (has zero residents); otherwise, segregation weakly rises. All indices described in the previous section satisfy this axiom.

Neighborhood Division Property (NDP) Let $X \in \mathcal{C}$ be a city and let n be a neighborhood of X. Let X' be the city that results from dividing n into two neighborhoods, n_1 and n_2. If either (a) at least one of n_1 and n_2 is empty or (b) n_1 and n_2 have the same group distributions (i.e., $\left(p_g^{n_1}\right)_{g\in G} = \left(p_g^{n_2}\right)_{g\in G}$), then $X' \sim X$. Otherwise, $X' \succcurlyeq X$.

The central idea of NDP is that if the partition of a city into neighborhoods becomes finer, one may detect more segregation, but not less.

NDP is related to two properties that are discussed by James and Taeuber [11] and subsequent authors. The first is organizational equivalence: if a neighborhood is divided into two neighborhoods that have the same group distribution, this does not change the city's level of segregation. For instance, the first neighborhood in the city $\langle (3,6), (2,1) \rangle$ is 1/3 black and 2/3 white. In the city $\langle (1,2), (2,4), (2,1) \rangle$, this neighborhood has been divided into two neighborhoods that also are 1/3 black, 2/3 white. Organizational equivalence states that these two cities are equally segregated.

The second property is the transfer principle. When there are two demographic groups, the transfer principle states that if a black (white) person moves from one neighborhood to another neighborhood in which the proportion of blacks (whites) is higher, then segregation in the city rises. For example, consider the city $\langle(5,5),(0,0)\rangle$. It is perfectly integrated since all residents live in a single neighborhood. Suppose that blacks then move, one by one, to the second neighborhood. The end result, $\langle(0,5),(5,0)\rangle$, is clearly more segregated than the initial city. The transfer principle implies that segregation rises along the way as well: each black who moves causes an increase in segregation.

When there are two demographic groups, NDP follows from organizational equivalence and the transfer principle.[9] On the other hand, it is not clear how to generalize the transfer principle to more than two groups.[10] For instance, consider a city with blacks, whites, and Asians. Suppose a black moves to a neighborhood that has higher proportions of both blacks and Asians. Since there are more blacks in the destination neighborhood, one might argue (using the transfer principle) that segregation has gone up. On the other hand, blacks are now more integrated with Asians, so

---

[9] To see this, let X be a city and let n be a neighborhood of X. Let X' be the city that results from dividing n into two neighborhoods, n_1 and n_2. Organizational equivalence directly implies $X' \sim X$ if n_1 and n_2 have the same demographic distributions. If they don't, assume without loss of generality that the proportion black is higher in n_1 than in n_2: $p_{n_1} > p_{n_2}$. Neighborhood n in city X can be written $(B_{n_1} + B_{n_2}, W_{n_1} + W_{n_2})$. Split this neighborhood into two neighborhoods with identical percents black: $n_1' = \left([B_{n_1} + B_{n_2}] \frac{W_{n_1}}{W_{n_1}+W_{n_2}}, W_{n_1}\right)$ and $n_2' = \left([B_{n_1} + B_{n_2}] \frac{W_{n_2}}{W_{n_1}+W_{n_2}}, W_{n_2}\right)$. Let the resulting city be X''. By organizational equivalence, $X \sim X''$. Since the proportion black is higher in n_1 than in n_2, neighborhood n_1' must have a proportion black that lies between p_{n_1} and p_{n_2}. Since the number of whites is the same in n_1' as in n_1, the number of blacks must be lower in n_1' than in n_1. Now move blacks from n_2' to n_1' until n_2' and n_1' have the same number of blacks as n_2 and n_1, respectively. (Note that the number of whites is also the same.) The city that results is X'. By the transfer principle, this operation strictly raises segregation: $X' \succ X'' \sim X$, so by transitivity, $X' \succ X$.

[10] One attempt to do so appears in Reardon and Firebaugh [19].

perhaps segregation has fallen. This issue does not arise with only two groups since if the proportion of blacks in the destination neighborhood is higher, the proportion of whites must be lower.

The next axiom states that if two cities have the same group distributions and the same total populations, then adjoining a given set of neighborhoods to each of them does not affect which of the two cities is more segregated. Intuitively, since they have the same size and group distributions, the degree of segregation *between* each city and the new set of neighborhoods is the same. Since the degree of segregation *within* the new set of neighborhoods is the same regardless of which city these neighborhoods are added to, which combined city is more segregated should be governed by the degree of segregation within each original city.

Independence (IND) Let $X, Y \in \mathcal{C}$ be two cities with the same set G of nonempty groups. Suppose X and Y have the same group distributions and the same total populations. Then for all cities $Z \in \mathcal{C}^A$ such that $G(Z) \subset G$,

$$X \succcurlyeq Y \text{ if and only if } X \uplus Z \succcurlyeq Y \uplus Z.$$

To see why the conditions on X and Y are needed, consider first an example in which X and Y have very different group distributions but equal sizes:

$$\begin{aligned} X &= \langle (1, 100) \rangle \\ Y &= \langle (100, 1) \rangle \\ Z &= \langle (100, 1) \rangle \\ X \uplus Z &= \langle (1, 100), (100, 1) \rangle \\ Y \uplus Z &= \langle (100, 1), (100, 1) \rangle \end{aligned}$$

Since Y is just X with the groups relabeled, they are equally segregated. Now consider adjoining Z to each of them. Each neighborhood in $Y \uplus Z$ is representative of the city, while in $X \uplus Z$ they are very unrepresentative: $X \uplus Z$ seems to be more segregated than $Y \uplus Z$. For this reason, we restrict to cities X and Y that have the same group distribution.

To see why X and Y must also be of the same size, consider the following example, in which X and Y have the same group distributions but very different sizes:

$$\begin{aligned} X &= \langle (90000, 10000), (10000, 90000) \rangle \\ Y &= \langle (1, 0), (0, 1) \rangle \\ Z &= \langle (100, 100) \rangle \end{aligned}$$

In this example, Y is completely segregated; Z is completely integrated; and X lies inbetween. Thus, $Y \succ X \succ Z$. In addition, since X is much larger than Z, $X \uplus Z$ should be about as segregated as X. Since Z is much larger than Y, $Y \uplus Z$ should be about as segregated as Z. Together, these imply that $X \uplus Z$ should be more segregated than $Y \uplus Z$: adding Z should reverse the segregation order. To permit segregation orderings with this reasonable property, the axiom of Independence restricts to cities X and Y that have the same size.

The final axiom is used to rule out the trivial ordering, in which all cities are equally segregated.

Nontriviality Suppose the class $\mathcal{C}$ contains some cities with exactly K nonempty groups, where $K \geq 2$. Then there exist cities $X, Y \in \mathcal{C}$, each with exactly K nonempty groups, such that $X \succ Y$.

5 Results

5.1 Fixed number of groups

In this section we consider the class $\mathcal{C}_K$ of cities that contain exactly K nonempty groups.[11] Fix a city $X \in \mathcal{C}_K$. For any neighborhood n in X, let $t^n(X) =$

[11] More precisely, $\mathcal{C}_K = \{X \in \mathcal{C} : |G(X)| = K\}$.

$\left(t_g^n(X)\right)_{g\in G(X)}$ denote the vector of the proportions of each group that live in n.[12] For example, in the city $X = \langle (3,4), (7,6) \rangle$, which is in $\mathcal{C}_2$, 30% of blacks and 40% of whites live in the first neighborhood, so $t^1(X) = (0.3, 0.4)$. We will omit the argument "X" when the city is clear from the context.

Recall that $\mathcal{C}^A$ is the set of all cities. Our first result is that the five axioms of section 4 are jointly satisfied by a unique segregation ordering, which is represented by A.

Theorem 1 *A segregation ordering on $\mathcal{C}_K$ satisfies GS, SI, NDP, IND, and N if and only if it is represented by A on $\mathcal{C}_K$.*

While A represents this unique ordering, any increasing transformation of A also represents it. However, A is the unique index that satisfies an intuitive property. Let $\underline{X}^K = \langle (1, ..., 1) \rangle$ be the completely integrated city with a single neighborhood that contains one member of each of the K groups. Let $\overline{X}^K = \langle (1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1) \rangle$ be the completely segregated city with exactly K residents, each from a different group and living in its own neighborhood. Note that $\underline{X}^1 = \langle (1) \rangle$ is well defined but $\overline{X}^K$ is only defined for $K \geq 2$. For any $K \geq 2$, consider the city $\alpha\overline{X}^K \uplus (1-\alpha)\underline{X}^K$. A proportion α of this city is completely segregated while $1-\alpha$ is completely integrated. We will refer to this city as the *α-segregated city in $\mathcal{C}_K$*.

We propose the following principle for segregation indices S on $\mathcal{C}_K$:

Cardinalization Principle on $\mathcal{C}_K$ There are two cases.

1. $K = 1$: S assigns a value of zero to the city $\underline{X}^1 = \langle (1) \rangle$, which contains a single person living in a single neighborhood.
2. $K > 1$: For any $\alpha \in [0, 1]$,

$$S\left(\alpha\overline{X}^K \uplus (1-\alpha)\underline{X}^K\right) = \alpha \tag{8}$$

[12] Note that the length of t^n is the number of nonempty groups in X, not the number of nonempty groups in n.

Proposition 1 *For all $K \geq 1$, the restriction of A to $\mathcal{C}_K$ satisfies the Cardinalization Principle on $\mathcal{C}_K$ and has an induced segregation ordering that satisfies GS, SI, NDP, IND, and N. Moreover, A is the unique segregation index on $\mathcal{C}^A$ with this property.*

5.2 Variable Number of Groups

Weak Group Division Property (WGDP)

Thus far we have considered how cities with a common number of groups should be ranked. In this section, we consider segregation orderings that allow comparisons among cities with different numbers of groups. This is often needed for longitudinal comparisons, since data-collecting agencies often change their racial schemas. It is also useful for international comparisons, since the number of racial groups may differ across countries.

In order to obtain a unique segregation ordering on the set of all cities, we introduce a weak assumption on how the ordering ranks certain pairs of cities with different numbers of groups. This axiom states that if each group g is divided into a common number of equal-sized groups, each with the same distribution across neighborhoods as the original distribution of g, segregation does not change. For example, consider the two-neighborhood city $\langle (2,4), (6,8) \rangle$ with two demographic groups. If we divide each group into two equal sized groups which are identically distributed acros neigborhoods, we get the two-neighborhood, four-group city $\langle (1,1,2,2), (3,3,4,4) \rangle$. According to the following axiom, these two cities are equally segregated.

Weak Group Division Property (WGDP) Let $X \in \mathcal{C}$ be a city in which the set of groups is G. Let X' be the result of partitioning each group $g \in G$ into $M \geq 2$ equal sized groups, g_1 through g_M, such that the M groups have the same distribution across neighborhoods as the original group: $\left(t^n_{g_m}\right)_{n \in N} = \left(t^n_g\right)_{n \in N}$.[13]

[13] Note that X' has the same set N of neighborhoods as X and for each neighborhood $n \in N$, $T^n_{g_m} = T^n_g / M$ for all $m \in \{1, \cdots, M\}$.

If $X' \in \mathcal{C}$, then $X' \sim X$.

With the addition of this axiom, we obtain a unique segregation ordering on the set of all cities:

Theorem 2 *A segregation ordering on $\mathcal{C}^A$ satisfies GS, SI, NDP, IND, N, and WGDP if and only if it is represented by the index A.*

Proof. By Theorem 1, for any K, the segregation ordering on $\mathcal{C}_K$ that is represented by the index A is the only ordering on $\mathcal{C}_K$ that satisfies the axioms GS, SI, NDP, IND, and N. We first show that the segregation ordering on $\mathcal{C}^A$ that is represented by the index A satisfies WGDP. Recall that $A(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g=1,\ldots,K} t_g^n \right)^{1/K}$ where K is the number of nonempty groups in X. Let X' be the result of partitioning each group g into M equal-sized groups, g_1 through g_m. Suppose that for all g, the M subgroups g_1 through g_M have the same distribution across neighborhoods as g itself: $t_{g_m}^n = t_g^n$ for all subgroups m and neighborhoods n. Then

$$
\begin{aligned}
A(X') &= 1 - \sum_{n \in N(X)} \left(\prod_{\substack{g=1,\ldots,K \\ m=1,\ldots,M}} t_{g_m}^n \right)^{\frac{1}{KM}} \\
&= 1 - \sum_{n \in N(X)} \left[\prod_{g=1,\ldots,K} \left(\prod_{m=1,\ldots,M} t_{g_m}^n \right)^{\frac{1}{M}} \right]^{\frac{1}{K}} \\
&= 1 - \sum_{n \in N(X)} \left[\prod_{g=1,\ldots,K} t_g^n \right]^{\frac{1}{K}} = A(X)
\end{aligned}
$$

Thus, the ordering induced by A on the set of all cities satisfies WGDP.

We now show that this is the only ordering on $\mathcal{C}^A$ that satisfies GS, SI, NDP, WGDP, IND, and N. Consider any two cities X and Y. Let K_X (K_Y) be the number of nonempty groups in X (Y). Let X' be the result of replacing each group in X by K_Y equal sized groups that have the same distribution across neighborhoods.

Let Y' be the result of replacing each group in Y by K_X equal sized groups that have the same distribution across neighborhoods. Note that X' and Y' have the same number of groups, $K_X K_Y$. By WGDP, $X \sim X'$ and $Y \sim Y'$. Thus, $X \succcurlyeq Y$ if and only if $X' \succcurlyeq Y'$. But by Theorem 1, $X' \succcurlyeq Y'$ if and only if $A(X') \geq A(Y')$. Since the ordering induced by A satisfies WGDP, $A(X) = A(X')$ and $A(Y) = A(Y')$. Hence, $X \succcurlyeq Y$ if and only if $A(X) \geq A(Y)$. Q.E.D. ■

Section 5.1 shows that A is the only index that satisfies the five axioms of Theorem 1 as well as the Cardinalization Principle when the number of groups is fixed. Theorem 2 shows that the ordering represented by A is the only one that satisfies these five axioms plus WGDP when the number of groups is variable. Any strictly increasing transformation of A also represents this unique ordering. However, for any $K \geq 2$, any such transformation must violate the Cardinalization Principle for K groups. Why? By inspection, A assigns to the most and least segregated cities the values of one and zero, respectively, regardless of the number of groups. This implies that the unique ordering $\succcurlyeq$ of Theorem 2 must rank all completely integrated cities as equally segregated:

$$\underline{X}^K \sim \underline{X}^{K'} \text{ for all } K, K' \geq 1 \tag{9}$$

Likewise, it must rank all completely segregated cities as equally segregated:[14]

$$\overline{X}^K \sim \overline{X}^{K'} \text{ for all } K, K' \geq 2 \tag{10}$$

Let A' be any increasing transformation of A and let X be a city for which $A(X) \neq$

[14] The first property is a direct implication of WGDP (using also Scale Invariance) but the second is not. To see why, note first that $\langle (1) \rangle$ is as segregated as $\left\langle \left(\underbrace{\frac{1}{K}, ..., \frac{1}{K}}_{K \text{ times}} \right) \right\rangle$ by WGDP and the latter city is as segregated as $\underline{X}^K$ by SI. So all completely integrated cities are equally segregated. The property that all completely segregated cities are equally segregated is not a direct consequence of WGDP. WGDP implies e.g., that $\overline{X}^2 = \langle (1,0), (0,1) \rangle$ is as segregated as $\langle (1,1,0,0), (0,0,1,1) \rangle$ (using Scale Invariance as well) but the latter city could, in principle, be less segregated than $\overline{X}^4$.

$A'(X)$. If X has $K' \geq 2$ groups, let $\alpha \in [0,1]$ be such that $X \sim \alpha \overline{X}^{K'} \uplus (1-\alpha)\underline{X}^{K'}$. By (9) and (10), $X \sim \alpha \overline{X}^{K} \uplus (1-\alpha)\underline{X}^{K}$ as well, for any $K \geq 2$. But $A'(X)$ cannot equal α since $A(X) = \alpha$ and $A(X) \neq A'(X)$. Hence, A' violates the Cardinalization Principle for each fixed number $K \geq 2$ of groups.

Is the Cardinalization Principle a reasonable requirement for a segregation index when K is variable? It is if the underlying ordering satisfies (9) and (10). These conditions imply that $\alpha \overline{X}^{K} \uplus (1-\alpha)\underline{X}^{K}$ is as segregated as $\alpha \overline{X}^{K'} \uplus (1-\alpha)\underline{X}^{K'}$ for any $K, K' \geq 2$. Thus, they must have the same segregation index. Because of the form these cities take, it is natural to require that this index be α.

5.3 Axiom Set II: WSI and GDP

The previous sections considered the Weak Group Division Property, which required that segregation be invariant to simultaneous division of existing groups into equisized and identically distributed subgroups. This axiom is weak because it applies only when

1. all the existing groups are subdivided into the same number of subgroups,
2. the subgroups of a given group have equal size, and
3. the subgroups of a given group are identically distributed across neighborhoods.

Condition (3) is essential: if the subgroups have different distributions across neighborhoods, then the city is more segregated after the subdivision, so segregation should rise.[15] However, conditions (1) and (2) do not contribute to the axiom's plausibility.

[15] Consider a city with 100 whites, who are split 20-80 between two neighborhoods: $\langle (20), (80) \rangle$. This city is completely integrated. Now suppose we modify the ethnic schema to separate the 20% of whites who are hispanic from the 80% who are not. Suppose, moreover, that the hispanics live entirely in the first neighborhood while the nonhispanics inhabit the second. By subdividing the whites, we have discovered considerable segregation. A reasonable segregation ordering will reflect this fact by ranking the new city as more segregated than the old one.

The following axiom strengthens WGDP by dropping these conditions.

Group Division Property (GDP) Let $X \in \mathcal{C}$ be a city in which the set of groups is G. Let X' be the result of partitioning some group $g \in G$ into two groups, g_1 and g_2, such that both groups have the same distribution across neighborhoods: $\left(t_{g_1}^n\right)_{n \in N} = \left(t_{g_2}^n\right)_{n \in N}$.[16] If $X' \in \mathcal{C}$, then $X' \sim X$.

This axiom states that segregation is invariant to the subdivision of an existing group into two identically distributed subgroups. For instance, if whites are divided into white female and white males, and these groups have the same distribution across neighborhoods, then measured segregation should not change. This new axiom is inconsistent with some of the axioms considered in the previous section. Since GDP is a stronger axiom than WGDP, the only candidate for an ordering that satisfies the axioms would be the one induced by the Atkinson index. But this ordering does not satifies GDP, as the following example shows.

Example 1 *Consider the following cities:* $X = \{(4,3),(2,3)\}$*, and* $Y = \{(2,2,3),(1,1,3)\}$*. City* Y *is obtained from* X *by splitting the first group into two identically distributed subgroups. But* $A(X) = 0.0144$ *while* $A(Y) = 0.0127$*.*

Accordingly, in order to use GDP we must weaken at least one of the previously used axioms. Scale Invariance is a sensible candidate. The combination of GDP and SI give rise to counterintuitive implications. For instance, suppose there are three equal-sized groups: blacks, Asians, and whites, and two neighborhoods. Suppose Asians are evenly distributed, most blacks are in the first neighborhood, and most whites are in the second:

$$\langle (100, 50, 10), (10, 50, 100) \rangle$$

[16]Note that X' has the same set N of neighborhoods as X and for each neighborhood $n \in N$, $T_g^n = T_{g_1}^n + T_{g_2}^n$.

(The order is black, Asian, and white.) Now double the number of Asians in each neighborhood:

$$\langle (100, 100, 10), (10, 100, 100) \rangle$$

By SI, segregation is unaffected. Next divide the new Asian group into two equal-sized groups that are both evenly distributed across neighborhoods:

$$\langle (100, 50, 50, 10), (10, 50, 50, 100) \rangle$$

GDP implies that segregation is unchanged. But taken together, the effect of the two operations is to create a duplicate group that is also evenly distributed across neighborhoods. This operation does seem as if it should lower the city's level of segregation. For this reason, we replace Scale Invariance by a weaker version:

Weak Scale Invariance (WSI) The segregation in a city is unchanged if the numbers of agents in all groups in all neighborhoods are multiplied by the same positive integer: for any city $X \in \mathcal{C}$ and any scalar $\alpha \geq 1$, if $\alpha X \in \mathcal{C}$ then $X \sim \alpha X$.

It is also necessary to modify Independence. Independence states that if X and Y have the same size and group distribution, then adjoining any city Z to X and to Y does not affect which city is more segregated. Since Z may not have the same group distribution as X and Y, adjoining Z will generally change the group distribution of the two cities. However, Scale Invariance would let us rescale the resulting cities so that their group distributions are the same as before. This means that in the presence of SI, Independence is equivalent to requiring that once the new cities are rescaled to restore their original group distributions, the segregation ordering is unaffected. When SI is dropped, we replace Independence by this alternative formulation. This makes sense since without SI, the segregation of a city might depend on its group distribution. Under the new axiom of Weak Independence, adjoining Z to X and to Y does not affect which city is more segregated *after the combined cities are rescaled to restore their original group distributions.*

Let ν be the "scaling" operation on cities that normalizes the size of each nonempty group to one. For instance, the result of applying ν to the city $\langle (1,2), (2,3) \rangle$ is the city $\langle \left(\frac{1}{3}, \frac{2}{5}\right), \left(\frac{2}{3}, \frac{3}{5}\right) \rangle$. We refer to a city in which all nonempty groups are of size 1 as a "normalized city". Weak Independence states that if X and Y are normalized cities, then adding a given set Z of neighborhoods to both cities and then scaling the result (which restores the cities' original group distributions) does not change which city is more segregated. This axiom is implied by IND and SI, but it does not itself imply IND or SI.

Weak Independence (WIND) Let $X, Y \in \mathcal{C}$ be two cities with the same set of nonempty groups G. Suppose that each nonempty group in X and Y has a size of one: $T_g(X) = T_g(Y) = 1$ for all groups $g \in G$. Then for all cities $Z \in \mathcal{C}^A$ such that $G(Z) \subset G$,

$$X \succcurlyeq Y \text{ if and only if } \nu(X \uplus Z) \succcurlyeq \nu(Y \uplus Z).$$

When Scale Invariance is imposed, Weak Independence is equivalent to Independence. Therefore, Theorems 1 and 2 would still hold if IND were replaced by WIND.

For technical reasons, we also modify Nontriviality as follows:

Dual Nontriviality Suppose the class $\mathcal{C}$ contains some cities with exactly T residents, where $T \geq 2$. Then there exist cities $X, Y \in \mathcal{C}$, each with exactly T residents, such that $X \succ Y$.

Let $\mathcal{C}^I \subset \mathcal{C}^A$ be the set of cities in which, for each nonempty group g, the number of members of group g is a positive integer: $T_g \in \mathbb{N}$. The following theorem states that the ordering represented by the Weighted Multigroup Atkinson Index (equation (2)) is the only one that satisfies the new set of axioms.

Theorem 3 *A segregation ordering on $\mathcal{C}^I$ satisfies GS, WSI, NDP, WIND, GDP, and DN if and only if it is represented by A^W on $\mathcal{C}^I$.*

Since each group has a weight equal to its proportion in the city, the weighted Atkinson index A^W is less sensitive to the introduction of small groups than the Atkinson index A. To illustrate, let $X = \langle (90, 10), (10, 90) \rangle$ be a city with 100 blacks and 100 whites, where blacks dominate the first neighborhood and whites the second. One can compute $A^W(X) = A(X) = 0.4$. Now suppose a black from the first neighborhood and a white from the second neighborhood are reclassified as belonging to the same mixed-race group. This yields the city $Y = \langle (89, 10, 1), (10, 89, 1) \rangle$. The reclassification causes a large drop in the Atkinson index to 0.29 while the weighted Atkinson index falls much less, to 0.39. This may make the weighted Atkinson index more suitable to studying multigroup segregation in the 2000 U.S. Census, which was the first to incorporate mixed-race groups.

6 Analysis of various indices and the independence of the axioms

It is natural to wonder whether the axioms in Theorems 2 and 3 are independent of each other. That is, for each axiom, is there an index that violates it yet that satisfies the other axioms? In this section, we show that this is indeed the case. Consequently, no axiom is superfluous: all are needed for our results to hold.

6.1 Independence of Axioms in Theorem 2

In this section we show that the axioms in Theorem 2 are independent. We also show that the Entropy index violates both SI and WGDP.

We begin with a simple sufficient condition for a segregation index to satisfy WGDP.

Lemma 1 *For any $X \in \mathcal{C}^A$, let $S(X) = \sum_{n \in N(X)} f(t^n)$ be a segregation index, where f is symmetric and satisfies $f(t) = f(\underbrace{t, \ldots, t}_{M\ times})$ for any $M \geq 1$. Then the segregation* *ordering on $\mathcal{C}^A$ that is represented by S on $\mathcal{C}^A$ satisfies WGDP.*

Proof. Let $X \in \mathcal{C}^A$. Let X' be the result of partitioning each group $g \in G(X)$ into $M \geq 2$ equal-sized groups, g_1 through g_M, where for all g, the M subgroups g_1 through g_M have the same distribution across neighborhoods as g itself. Since f is symmetric, we can reorder its arguments to obtain:

$$S(X') = \sum_{n \in N(X)} f(\underbrace{t^n, \ldots, t^n}_{M \text{ times}}) = \sum_{n \in N(X)} f\left(t^n\right) = S(X)$$

Q.E.D. ■

The following lemma gives sufficient conditions for a segregation index on $\mathcal{C}_K$ to satisfy GS, SI, NDP, and N.

Lemma 2 *For any $X \in \mathcal{C}_K$, let $S(X) = \sum_{n \in N(X)} f\left(t^n\right)$ be a segregation index. Then:*

1. *S satisfies SI.*
2. *If $f(t)$ is symmetric, then S satisfies GS.*
3. *If f is weakly convex and homogeneous of degree 1, then S satisfies NDP.*
4. *If f is not proportional to the sum of its arguments, then S satisfies N.*

Proof. For any positive scalar α, $t_g^n = \frac{T_g^n}{T_g} = \frac{\alpha T_g^n}{\alpha T_g}$; this establishes part 1. Parts 2 and 4 are trivial. For part 3, let X be a city and let m and n be two neighborhoods of X. Let X' be the city that results from combining m and n into a single neighborhood. Note that the vector of proportions of each group who are in the combined neighborhood is just $t^m + t^n$. Hence,

$$S(X) - S(X') = f(t^m) + f(t^n) - f\left(t^m + t^n\right) \tag{11}$$

By the weak convexity of f, this is nonnegative, so $X \succcurlyeq X'$. If one neighborhood (say m) is empty, then t^m is the zero vector; by homogeneity, $f(\mathbf{0}) = 0$, so by (11),

$X \sim X'$. Finally, if $\left(p_g^m\right)_{g \in G_K} = \left(p_g^n\right)_{g \in G_K}$, there must be a constant $\alpha > 0$ such that $t^m = \alpha t^n$; consequently,

$$S(X) - S(X') = f(\alpha t^n) + f(t^n) - f\left((\alpha + 1)\, t^n\right)$$

This is zero by homogeneity, so $X \sim X'$. Q.E.D. ■

We now consider the axioms in 2.

Claim 1 *The Multigroup Dissimilarity Index D satisfies GS, SI, NDP, N, and WGDP on $\mathcal{C}^A$, but fails IND.*

Proof. The function f in (3) is symmetric, weakly convex, homogeneous of degree one, and not proportional to the sum of its arguments. By Lemma 2, D satisfies GS, SI, NDP, and N. D satisfies WGDP by Lemma 1. As for IND, consider the following cities: $X = \langle (2,4), (2,0) \rangle$ and $Y = \langle (4,2), (0,2) \rangle$. It can be checked that $D(X) = D(Y) = 1/2$. Consider now the result of annexing to them the one-neighborhood city $Z = \langle (4,0) \rangle$. It can be checked that $D(X \uplus Z) = 3/4$ while $D(Y \uplus Z) = 1/2$. Hence, D violates IND. Q.E.D. ■

Claim 2 *The Entropy index H (equation (4)) satisfies GS, NDP, IND, and N on $\mathcal{C}^A$, but fails SI and WGDP.*

Proof. GS and N are trivial. To see that H satisfies IND, Suppose X and Y have the same group distribution and the same total population. Let G denote the set of nonempty groups in X and Y. Let Z be any other city such that $G(Z) \subset G$. Let $T' = T(X) + T(Z) = T(Y) + T(Z)$, $\eta = T(X)/T' = T(Y)/T'$, and

$$\begin{aligned} H_{\text{total}} &= H_{\text{total}}(X) = H_{\text{total}}(Y) \\ H'_{\text{total}} &= H_{\text{total}}(X \uplus Z) = H_{\text{total}}(Y \uplus Z) = \sum_{g \in G} P'_g \log_2 \left(\frac{1}{P'_g} \right) \end{aligned}$$

where P'_g is the proportion of residents in the combined cities $X \uplus Z$ and $Y \uplus Z$ who are in group g. We calculate

$$\begin{aligned} H_{\text{within}}(X \uplus Z) &= \sum_{n \in N(X)} \frac{T^n(X)}{T'} H_n + \sum_{n \in N(Z)} \frac{T^n(Z)}{T'} H_n \\ &= \eta H_{\text{within}}(X) + (1-\eta) H_{\text{within}}(Z) \end{aligned}$$

and likewise for $H_{\text{within}}(Y \uplus Z)$. Hence,

$$\begin{aligned} H(X \uplus Z) - H(Y \uplus Z) &= \frac{H_{\text{within}}(Y \uplus Z) - H_{\text{within}}(X \uplus Z)}{H'_{\text{total}}} \\ &= \eta \frac{H_{\text{within}}(Y) - H_{\text{within}}(X)}{H'_{\text{total}}} \end{aligned}$$

Now suppose that $H(X) \geq H(Y)$. This implies

$$H(X) - H(Y) = \frac{H_{\text{within}}(Y) - H_{\text{within}}(X)}{H_{\text{total}}} \geq 0$$

so that $H_{\text{within}}(Y) \geq H_{\text{within}}(X)$; but then $H(X \uplus Z) - H(Y \uplus Z) \geq 0$ as well.

We now verify that H satisfies NDP. Let $X \in \mathcal{C}^A$ be a city and let n be a neighborhood of X. Let X' be the city that results from dividing n into two neighborhoods, n_1 and n_2. If one of n_1 and n_2 is empty, say n_1, then $H_{n_1} = 0$ and $H_n = H_{n_2}$ so $H(X) = H(X')$. If n_1 and n_2 have the same group distributions, then $H_{n_1} = H_{n_2} = H_n$ and $T^n = T^{n_1} + T^{n_2}$, so $H(X) = H(X')$. Otherwise,

$$\begin{aligned} p_g^n &= \frac{T_g^{n_1} + T_g^{n_2}}{T^n} = \frac{T^{n_1}}{T^n}\frac{T_g^{n_1}}{T^{n_1}} + \frac{T^{n_2}}{T^n}\frac{T_g^{n_2}}{T^{n_2}} \\ &= \tau p_g^{n_1} + (1-\tau) p_g^{n_2} \text{ where } \tau = \frac{T^{n_1}}{T^n} \end{aligned}$$

Moreover, the entropy function $H_m = \sum_{g \in G(X)} p_g^m \log_2 \left(\frac{1}{p_g^m}\right)$ is a strictly concave function of $\left(p_g^m\right)_{g \in G(X)}$,[17] so $H_n > \tau H_{n_1} + (1-\tau) H_{n_2}$. Thus,

$$H(X) - H(X') = \frac{T^n H_n - T^{n_1} H_{n_1} - T^{n_2} H_{n_2}}{T H_{\text{total}}} > 0$$

[17] To see this, note that

$$\frac{d^2}{dp^2}\left(p \ln\left(\frac{1}{p}\right)\right) = \frac{d}{dp}\left(\ln\frac{1}{p} - 1\right) = -\frac{1}{p} < 0$$

and the sum of concave functions is concave.

As for WGDP, let $X \in \mathcal{C}^A$ and let X' be the result of partitioning each group $g \in G(X)$ into $M \geq 2$ equal-sized groups, g_1 through g_M, with the same distribution across neighborhoods as g itself. Note that

$$\begin{aligned} H_{\text{total}}(X') &= \sum_{g \in G(X')} P_g \log_2 \left(\frac{1}{P_g} \right) = \sum_{g \in G(X)} M \frac{P_g}{M} \log_2 \left(\frac{1}{P_g/M} \right) \\ &= H_{\text{total}}(X) + \log_2 (M) \end{aligned}$$

and

$$\begin{aligned} H_{\text{within}}(X') &= \sum_{n \in N(X)} \frac{T^n}{T} H_n(X') = \sum_{n \in N(X)} \frac{T^n}{T} \sum_{g \in G(X')} p_g^n \log_2 \left(\frac{1}{p_g^n} \right) \\ &= \sum_{n \in N(X)} \frac{T^n}{T} \left[H_n(X) + \log_2(M) \right] = H_{\text{within}}(X) + \log_2(M) \end{aligned}$$

so

$$H(X') = 1 - \frac{H_{\text{within}}(X) + \log_2(M)}{H_{\text{total}}(X) + \log_2 (M)} < H(X)$$

Q.E.D. ■

We now build an index that violates only GS: an index that is sensitive to how the groups are labeled. To do so we need to assign labels to the groups. One way is as follows. $\mathcal{C}^A$ is the set of cities with any number of nonempty groups. Let G^A denote the (infinite) set of groups in the cities of $\mathcal{C}^A$. Each group g in G^A will be assigned a label $\lambda(g)$, which equals b or w. Subgroups of a given group inherit that group's label. For instance, b might mark groups in which all members have some African blood; w then refers to groups in which no members have African blood. Since every group has one label or the other, we are excluding groups in which some members have African blood and some do not. This is permissible since $\mathcal{C}^A$ is the set of cities with any *number* of groups; it does not necessarily include all *possible* groups. Let $\beta \in (0, 1)$. For any group g, let the function $\varepsilon(g)$ equal β if the label of g is b and $1 - \beta$ if it is w. Let $\theta_G(g) = \varepsilon(g) \left[\sum_{g' \in G(X)} \varepsilon(g') \right]^{-1}$. ($\theta_G$ is a normalization of ε that sums to one: $\sum_{g \in G} \theta_G(g) = 1$.) Define the asymmetric Atkinson index with parameter β to be:

$$A_\beta(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G(X)} \left(t_g^n\right)^{\theta_{G(X)}(g)} \right)$$

This index is like the Multigroup Atkinson Index but it gives weight β to groups with label b and weight $1-\beta$ to groups that have label w. For $\beta = 1/2$, these weights are equal: $A_{1/2}$ is just the Multigroup Atkinson Index.

Claim 3 *For $\beta \neq 1/2$, the index A_β satisfies SI, NDP, IND, N, and WGDP on $\mathcal{C}^A$, but fails GS.*

Proof. The index fails GS since it is not invariant to permutations that do not preserve the groups' labels. The index can be rewritten

$$\begin{aligned} A_\beta(X) &= \sum_{n \in N(X)} f(t^n) \\ \text{where } f(t^n) &= \sum_{g \in G(X)} \theta_G(g) t_g^n - \prod_{g \in G(X)} \left(t_g^n\right)^{\theta_{G(X)}(g)} \end{aligned}$$

Note that f is convex and homogeneous of degree one, and is not proportional to the sum of its arguments. By Lemma 2, A_β satisfies SI, NDP, and N.

As for IND, let $X, Y \in \mathcal{C}_K$ be two cities with the same set G of nonempty groups, the same group distributions, and the same total populations. Let $Z \in \mathcal{C}^A$ such that $G(Z) \subset G$. We wish to show that $A(X) \geq A(Y)$ if and only if $A(X \uplus Z) \geq A(Y \uplus Z)$. Let $\gamma_g = \frac{T_g(X)}{T_g(X \uplus Z)} = \frac{T_g(Y)}{T_g(Y \uplus Z)}$ and $\eta_g = \frac{T_g(Z)}{T_g(X \uplus Z)} = \frac{T_g(Z)}{T_g(Y \uplus Z)}$. Note that a proportion $t_g^n \gamma_g$ of group-g residents of the city $X \uplus Z$ live in neighborhood $n \in N(X)$. Likewise, a proportion $t_g^n \eta_g$ of group-g residents of the city $X \uplus Z$ live in neighborhood $n \in N(Z)$.

Analogous statements are true for $Y \uplus Z$. Accordingly,

$$
\begin{aligned}
& A_\beta(X \uplus Z) \geq A_\beta(Y \uplus Z) \\
& \qquad \Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in G} \left(t_g^n \gamma_g\right)^{\theta_G(g)} \right) + \sum_{n \in N(Z)} \left(\prod_{g \in G} \left(t_g^n \eta_g\right)^{\theta_G(g)} \right) \leq \sum_{n \in N(Y)} \left(\prod_{g \in G} \left(t_g^n \gamma_g\right)^{\theta_G(g)} \right) + \sum_{n \in N(Z)} \Bigg(\\
& \qquad \Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in G} \left(t_g^n \gamma_g\right)^{\theta_G(g)} \right) \leq \sum_{n \in N(Y)} \left(\prod_{g \in G} \left(t_g^n \gamma_g\right)^{\theta_G(g)} \right) \\
& \qquad \Leftrightarrow \left(\prod_{g \in G} \left(\gamma_g\right)^{\theta_G(g)} \right) \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n\right)^{\theta_G(g)} \leq \left(\prod_{g \in G} \left(\gamma_g\right)^{\theta_G(g)} \right) \sum_{n \in N(Y)} \prod_{g \in G} \left(t_g^n\right)^{\theta_G(g)} \\
& \qquad \Leftrightarrow \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n\right)^{\theta_G(g)} \leq \sum_{n \in N(Y)} \prod_{g \in G} \left(t_g^n\right)^{\theta_G(g)} \\
& \qquad \Leftrightarrow A_\beta(X) \geq A_\beta(Y)
\end{aligned}
$$

This shows that the Atkinson index satisfies IND as well.

Regarding WGDP, let $X \in \mathcal{C}$ and let $G = G(X)$ and $N = N(X)$. Let X' be the result of partitioning each group $g \in G$ into $M \geq 2$ equal-sized groups, g_1 through g_M, where for all g, the M subgroups g_1 through g_M have the same distribution across neighborhoods as g itself. Let G' be the set of groups that results from this operation. For any new group $g \in G'$, let $\pi(g) \in G$ denote the parent group of g: the group in G of which g is a subgroup. For each $g \in G'$, $\varepsilon(g) = \varepsilon(\pi(g))$, so $\theta_{G'}(g) = \varepsilon(\pi(g)) \left[\sum_{g' \in G'} \varepsilon(\pi(g'))\right]^{-1} = M^{-1}\theta_G(\pi(g))$. Hence,

$$
\begin{aligned}
A_\beta(X') &= 1 - \sum_{n \in N} \left(\prod_{g \in G'} \left(t_g^n\right)^{\theta_{G'}(g)} \right) = 1 - \sum_{n \in N} \left(\prod_{g \in G'} \left(t_{\pi(g)}^n\right)^{M^{-1}\theta_G(\pi(g))} \right) \\
&= 1 - \sum_{n \in N} \left(\prod_{g \in G} \left[\left(t_g^n\right)^{M^{-1}\theta_G(g)} \right]^M \right) = A_\beta(X)
\end{aligned}
$$

Q.E.D. ∎

We have shown that GS and IND are independent of the other axioms. It remains to show that SI, NDP, N, and WGDP are each independent. To show that NDP is logically independent, note that $1-A$ satisfies all of the axioms but NDP. As for N, the trivial segregation order, which ranks all cities as equally segregated, violates N while satisfying all the other axioms. For WGDP, consider the index $S(X) = |G(X)|\, A(X)$. This index satisfies GS, SI, NDP, IND, and N, since it represents the same ordering as A does for any fixed number of groups. However, it clearly violates WGDP since subdividing each group into M equally distributed groups increases the index by a factor of M.

Finally, consider the index $S(X) = T(X)A(X)$. It clearly satisfies GS, NDP, and WGDP, since these axioms involve transformations that do not change a city's population. For all $K \geq 2$, $A(\overline{X}^K) = 1$, so $S(\overline{X}^K) = K$ while $S(2\overline{X}^K) = 2K$. This implies that S satisfies N but violates SI. As for IND, Let $X, Y \in \mathcal{C}^A$ be two cities with the same set G of nonempty groups, with the same group distributions and the same total populations. Let $Z \in \mathcal{C}^A$ be such that $G(Z) \subset G$. Then $T(X) = T(Y)$ and $T(X \uplus Z) = T(Y \uplus Z)$ so

$$
\begin{aligned}
S(X \uplus Z) \quad &\geq \quad S(Y \uplus Z) \\
&\iff T(X \uplus Z)A(X \uplus Z) \geq T(Y \uplus Z)A(Y \uplus Z) \\
&\iff A(X \uplus Z) \geq A(Y \uplus Z) \\
&\iff A(X) \geq A(Y) \\
&\iff T(X)A(X) \geq T(Y)A(Y) \\
&\iff S(X) \geq S(Y)
\end{aligned}
$$

so S satisfies IND.

The following table summarizes the above discussion.

	GS	SI	NDP	IND	N	WGDP
Atkinson: $A(X)$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$A_\beta(X)$ for $\beta \neq 1/2$		$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$T(X) * A(X)$	$\surd$		$\surd$	$\surd$	$\surd$	$\surd$
$1 - A(X)$	$\surd$	$\surd$		$\surd$	$\surd$	$\surd$
Dissimilarity: $D(X)$	$\surd$	$\surd$	$\surd$		$\surd$	$\surd$
Trivial index	$\surd$	$\surd$	$\surd$	$\surd$		$\surd$
$\lvert G(X)\rvert * A(X)$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	
Entropy: $H(X)$	$\surd$		$\surd$	$\surd$	$\surd$	

Table 1: Independence of the axioms

6.2 Independence of Axioms in Theorem 3

In this section we show that the axioms of Theorem 3 are independent....

A Proofs

We first state and prove some preliminary lemmas.

Lemma 3 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}_K$ that satisfies NDP and SI.*

1. *All cities in which every neighborhood is representative have the same degree of segregation under $\succcurlyeq$.*
2. *Any city in which every neighborhood is representative is weakly less segregated under $\succcurlyeq$ than any city in which some neighborhood is unrepresentative.*

Proof.

1. Consider any city Y in which every neighborhood is representative. Number the neighborhoods $1, ..., N$. For each $i = 1, ..., N$, let Y_i be city that results from Y when the first i neighborhoods of Y are combined into a single neighborhood. By NDP, for each $i = 1, ..., N-1$, $Y_i \sim Y_{i+1}$. Hence, by transitivity, $Y = Y_1 \sim$

Y_N. Y_N contains a single neighborhood. But by SI, any city with a single neighborhood is as segregated as any other city with a single neighborhood.

2. Consider any city X in which at least one neighborhood is unrepresentative. Number the neighborhoods $1, ..., N$. For each $i = 1, ..., N$, let X_i be city that results from X when the first i neighborhoods of X are combined into a single neighborhood. By NDP, for each $i = 1, ..., N-1$, $X_i \succcurlyeq X_{i+1}$. Hence, by transitivity, $X = X_1 \succcurlyeq X_N$. X_N contains a single neighborhood.

Q.E.D. ■

Lemma 4 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}_K$ that satisfies NDP and SI. All completely segregated cities have the same degree of segregation under $\succcurlyeq$, and are weakly more segregated than any city in which any neighborhood is mixed.*

Proof. Consider a completely segregated city X. Let X' be the city that results from X when, for each group $g \in G_K$, all neighborhoods that contain only members of group g are combined into a single neighborhood. (X' thus consists of K neighborhoods, each of which contains all the members of a single group.) By iteratively applying NDP, $X \sim X'$. By SI, X' is as segregated as any other city that consists of K neighborhoods, each of which contains all the members of a single group. This implies that all completely segregated cities have the same degree of segregation.

Now any city that has at least one mixed neighborhood can be converted into a completely segregated city by dividing each neighborhood n into K distinct neighborhoods, each of which includes all and only the members of a single group. By NDP, this procedure results in a weakly more segregated city. Q.E.D. ■

Lemma 5 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}_K$ that satisfies NDP, IND, N, and SI. Let $\overline{X} \in \mathcal{C}_K$ be the completely segregated city, $\langle (1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1) \rangle$, and let $\underline{X} \in \mathcal{C}_K$ be the completely integrated city $\langle (1, ..., 1) \rangle$. Then*

1. $\overline{X} \succ \underline{X}$;

2. *for any* $\alpha, \beta \in [0,1]$, $\alpha > \beta$,

$$\alpha\overline{X} \uplus (1-\alpha)\underline{X} \succ \beta\overline{X} \uplus (1-\beta)\underline{X}$$

Proof.

1. By N, there exist cities X and Y such that $X \succ Y$. By lemmas 3 and 4, $\overline{X} \succcurlyeq X \succ Y \succcurlyeq \underline{X}$, so $\overline{X} \succ \underline{X}$.

2. By NDP,

$$\alpha\overline{X} \uplus (1-\alpha)\underline{X} \sim \beta\overline{X} \uplus (\alpha-\beta)\overline{X} \uplus (1-\alpha)\underline{X}$$

and

$$\beta\overline{X} \uplus (1-\beta)\underline{X} \sim \beta\overline{X} \uplus (\alpha-\beta)\underline{X} \uplus (1-\alpha)\underline{X}$$

By part 1 and SI, $(\alpha-\beta)\overline{X} \succ (\alpha-\beta)\underline{X}$. Since the numbers of members of each group are equal in city $\overline{X}$ and in $\underline{X}$, they are also equal in city $(\alpha-\beta)\overline{X}$ and in $(\alpha-\beta)\underline{X}$. So by IND,

$$\beta\overline{X} \uplus (\alpha-\beta)\overline{X} \uplus (1-\alpha)\underline{X} \succ \beta\overline{X} \uplus (\alpha-\beta)\underline{X} \uplus (1-\alpha)\underline{X}$$

The result follows by transitivity.

Q.E.D. ■

Proof of Theorem 1. We first show that the Atkinson index satisfies all the axioms. We then show that any other index that satisfies all the axioms must represent the same ordering as the Atkinson index.

One can easily verify that A can be written $A(X) = \sum_{n \in N(X)} f^A(t^n)$ where $f^A(t)$ equals the arithmetic mean of the components of t less the geometric mean: $f^A(t) = \left(\frac{1}{K}\sum_{g \in G_K} t_g^n\right) - \left(\prod_{g \in G_K} t_g^n\right)^{1/K}$. Clearly, f^A is symmetric, weakly convex, homogeneous of degree 1, and is not proportional to the sum of its arguments. By Lemma 2, A satisfies GS, SI, NDP, and N.

As for IND, let $X, Y \in \mathcal{C}_K$ be two cities with the same set G of nonempty groups. Suppose X and Y have the same group distributions and the same total populations. Let $Z \in \mathcal{C}^A$ such that $G(Z) \subset G$. We wish to show that $A(X) \geq A(Y)$ if and only if $A(X \uplus Z) \geq A(Y \uplus Z)$. Let $\gamma_g = \frac{T_g(X)}{T_g(X \uplus Z)} = \frac{T_g(Y)}{T_g(Y \uplus Z)}$ and $\eta_g = \frac{T_g(Z)}{T_g(X \uplus Z)} = \frac{T_g(Z)}{T_g(Y \uplus Z)}$. Note that a proportion $t_g^n \gamma_g$ of group-g residents of the city $X \uplus Z$ live in neighborhood $n \in N(X)$. Likewise, a proportion $t_g^n \eta_g$ of group-g residents of the city $X \uplus Z$ live in neighborhood $n \in N(Z)$. Analogous statements are true for $Y \uplus Z$. Accordingly,

$$
\begin{aligned}
&A(X \uplus Z) \geq A(Y \uplus Z) \\
&\qquad \Leftrightarrow \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n \gamma_g\right)^{1/K} + \sum_{n \in N(Z)} \prod_{g \in G} \left(t_g^n \eta_g\right)^{1/K} \leq \sum_{m \in N(Y)} \prod_{g \in G} \left(t_g^m \gamma_g\right)^{1/K} + \sum_{m \in N(Z)} \prod_{g \in G} \left(t_g^m \eta_g\right)^{1/K} \\
&\qquad \Leftrightarrow \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n \gamma_g\right)^{1/K} \leq \sum_{m \in N(Y)} \prod_{g \in G} \left(t_g^m \gamma_g\right)^{1/K} \\
&\qquad \Leftrightarrow \left(\prod_{g \in G} \left(\gamma_g\right)^{1/K}\right) \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n\right)^{1/K} \leq \left(\prod_{g \in G} \left(\gamma_g\right)^{1/K}\right) \sum_{m \in N(Y)} \prod_{g \in G} \left(t_g^m\right)^{1/K} \\
&\qquad \Leftrightarrow \sum_{n \in N(X)} \prod_{g \in G} \left(t_g^n\right)^{1/K} \leq \sum_{m \in N(Y)} \prod_{g \in G} \left(t_g^m\right)^{1/K} \\
&\qquad \Leftrightarrow A(X) \geq A(Y)
\end{aligned}
$$

This shows that the Atkinson index satisfies IND as well.

Fix a number of groups $K \geq 2$. We now show that any ordering that satisfies GS, SI, NDP, IND, and N on $\mathcal{C}_K$ must be the Atkinson ordering. Let $\succcurlyeq$ be such an ordering. Let $\overline{X} \in \mathcal{C}_K$ be the completely segregated city, $\langle (1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1) \rangle$, and let $\underline{X} \in \mathcal{C}_K$ be the completely integrated city $\langle (1, ..., 1) \rangle$.

Claim 4 *Let* $t \in [0, 1]^K$ *and let*

$$X = \langle t, (1 - t_1, 0, ..., 0), (0, 1 - t_2, 0, ..., 0), ..., (0, ..., 0, 1 - t_K) \rangle.$$

Then, there exists a unique $\alpha_X \in [0, 1]$ *such that*

$$X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X}.$$

Further, this unique α_X *is* $\left(\prod_{g=1}^{K} t_g\right)^{1/K}$.

Proof. In this proof, all neighborhoods are K-tuples. For existence, there are two cases.

Case 1. Suppose $t_g = 0$ for some g. In this case we have to show that $\alpha_X = 0$ or, equivalently, that $X \sim \overline{X}$. By symmetry, we can assume WLOG that $t_1 = 0$. Therefore $t = (0, t_2, t_3, ..., t_K)$. Let σ_{12} be the permutation that relabels groups 1 and 2 into 2 and 1, respectively. Therefore, $\sigma_{12}t = (t_2, 0, t_3, ..., t_K)$. By GS,

$$\begin{aligned} X &= t * \underline{X} \uplus (\mathbf{1} - t) * \overline{X} \\ &\sim \sigma_{12}t * \underline{X} \uplus (\mathbf{1} - \sigma_{12}t) * \overline{X} = \widehat{X} \end{aligned}$$

For any $\beta \in (0,1)$, let $\gamma = (\beta, 1, ..., 1)$. By IND and SI,

$$\gamma * X \uplus (1-\gamma) * \overline{X} \sim \gamma * \widehat{X} \uplus (1-\gamma) * \overline{X}$$

In other words,

$$\gamma * \left(t * \underline{X} \uplus (\mathbf{1} - t) * \overline{X}\right) \uplus (1-\gamma) * \overline{X} \sim \gamma * \left(\sigma_{12}t * \underline{X} \uplus (\mathbf{1} - \sigma_{12}t) * \overline{X}\right) \uplus (1-\gamma) * \overline{X}$$

Hence, by NDP and GS,

$$\begin{aligned} (\gamma * t) * \underline{X} \uplus (\mathbf{1} - \gamma * t) * \overline{X} &\sim (\gamma * \sigma_{12}t) * \underline{X} \uplus (\mathbf{1} - \gamma * \sigma_{12}t) * \overline{X} \\ &\sim [\sigma_{12}(\gamma * \sigma_{12}t)] * \underline{X} \uplus (\mathbf{1} - [\sigma_{12}(\gamma * \sigma_{12}t)]) * \overline{X}. \end{aligned}$$

Therefore

$$\begin{aligned} X &= t * \underline{X} \uplus (\mathbf{1} - t) * \overline{X} \\ &= (0, t_2, t_3, ..., t_K) * \underline{X} \uplus (\mathbf{1} - (0, t_2, t_3, ..., t_K)) * \overline{X} \\ &\sim (0, \beta t_2, t_3, ..., t_K) * \underline{X} \uplus (\mathbf{1} - (0, \beta t_2, t_3, ..., t_K)) * \overline{X}. \end{aligned}$$

where the first two lines are definitions and the last one uses the fact that $(\gamma * t) = t$ and that $\sigma_{12}(\gamma * \sigma_{12}t) = \sigma_{12}\gamma * t = (0, \beta t_2, t_3, ..., t_K)$. We can repeat this procedure for $t_3, ..., t_K$ to obtain

$$\begin{aligned} X &\sim (0, \beta t_2, \beta t_3, ..., \beta t_K) * \underline{X} \uplus (\mathbf{1} - (0, \beta t_2, \beta t_3, ..., \beta t_K)) * \overline{X} \\ &\sim \beta t * \underline{X} \uplus (\mathbf{1} - \beta t) * \overline{X}. \end{aligned}$$

Now choose some constants $\beta, \beta' \in (0,1)$, $\beta > \beta'$, We obtain:

$$X \sim \beta t * \underline{X} \uplus (\mathbf{1} - \beta t) * \overline{X}$$

and

$$X \sim \beta' t * \underline{X} \uplus (\mathbf{1} - \beta' t) * \overline{X}$$

But by NDP,

$$\begin{aligned} &\beta' t * \underline{X} \uplus (\mathbf{1} - \beta' t) * \overline{X} \\ \sim\ &\beta' t * \underline{X} \uplus (\mathbf{1} - \beta t) * \overline{X} \uplus (\beta - \beta') t * \overline{X} \end{aligned}$$

and

$$\begin{aligned} &\beta t * \underline{X} \uplus (\mathbf{1} - \beta t) * \overline{X} \\ \sim\ &\beta' t * \underline{X} \uplus (\beta - \beta') t * \underline{X} \uplus (\mathbf{1} - \beta t) * \overline{X} \end{aligned}$$

so by IND,

$$(\beta - \beta') t * \overline{X} \sim (\beta - \beta') t * \underline{X}$$

so by SI,

$$t * \overline{X} \sim t * \underline{X}$$

Now by NDP and IND,

$$\overline{X} \sim t * \overline{X} \uplus (\mathbf{1} - t) * \overline{X} \sim t * \underline{X} \uplus (\mathbf{1} - t) * \overline{X} = X$$

so that $\alpha_X = 0$, as claimed

Case 2. Let

$$X = t * \underline{X} \uplus (\mathbf{1} - t) * \overline{X} = \langle (t_1, ..., t_K), (1 - t_1, 0, ..., 0), (0, 1 - t_2, 0, ..., 0), ..., (0, ..., 0, 1 - t_K) \rangle.$$

Suppose $t_g \in (0,1]$ for all g. Let $\alpha = \left(\prod_{g=1}^{K} t_g\right)^{1/K}$, and let

$Y = \alpha\underline{X} \uplus (1-\alpha)\overline{X} = \langle(\alpha, ..., \alpha), (1-\alpha, 0, ..., 0), (0, 1-\alpha, 0, ..., 0), ..., (0, ..., 0, 1-\alpha)\rangle.$

We shall show that $X \sim Y$ and therefore that α is the α_X we are looking for.
Let $\gamma_1 \in (0,1)$. For $g = 2, ..., K$, define $\gamma_g = \gamma_{g-1}\frac{t_{g-1}}{\alpha}$. Note that by definition of α,

$$\begin{aligned}\gamma_K &= \gamma_1 \prod_{g=1}^{K-1}\left(\frac{t_g}{\alpha}\right) = \gamma_1\left(\frac{\prod_{g=1}^{K-1} t_g}{\alpha^{K-1}}\right) = \gamma_1\left(\frac{1/t_K}{1/\alpha}\frac{\prod_{g=1}^{K} t_g}{\alpha^K}\right) = \gamma_1\left(\frac{1/t_K}{1/\alpha}\right) = \gamma_1\frac{\alpha}{t_K} \\ &\Longrightarrow \gamma_1 = \gamma_K\frac{t_K}{\alpha}\end{aligned}$$

Now choose γ_1 small enough that each $\gamma_g \leq 1$; this holds if $\max_{g\in\langle 2,...,K\rangle}\gamma_g = \max_{g\in\langle 2,...,K\rangle}\gamma_1 \prod_{j=2}^{g}\left(\frac{t_{j-1}}{\alpha}\right) \leq 1$.Denote by $\gamma = (\gamma_1, \ldots, \gamma_K)$ the K-tuple just built. Note than $\alpha\gamma$ is a permutation of $\gamma * t$. Now by definition of X and Y, by SI and IND, and by NDP

$$\begin{aligned}X &\sim Y \Leftrightarrow t * \underline{X} \uplus (\mathbf{1}-t)\overline{X} \sim \alpha\underline{X} \uplus (1-\alpha)\overline{X} \\ &\Leftrightarrow \gamma * \left(t * \underline{X} \uplus (\mathbf{1}-t)\overline{X}\right) \uplus (\mathbf{1}-\gamma)\overline{X} \sim \gamma * \left(\alpha * \underline{X} \uplus (\mathbf{1}-\alpha)\overline{X}\right) \uplus (\mathbf{1}-\gamma)\overline{X} \\ &\Leftrightarrow (\gamma * t) * \underline{X} \uplus (\mathbf{1}-\gamma * t)\overline{X} \sim (\alpha\gamma) * \underline{X} \uplus (\mathbf{1}-\alpha\gamma)\overline{X}.\end{aligned}$$

But the last two cities are equally segregated because $\alpha\gamma$ is a permutation of $\gamma * t$ and $\succcurlyeq$ satisfies GS. ■

Claim 5 *Let* $t^1, t^2 \in [0,1]^2$ *and let* $X = \langle t^1, t^2, (1-t^1_1-t^2_1, 0, ..., 0), (0, 1-t^1_2-t^2_2, 0, ..., 0), ..., (0, ..., 0, 1-t^1_K-t^2_K)\rangle$ *be a city. Then there is a unique* $\alpha_X \in [0,1]$ *such that*

$$X \sim \alpha_X\underline{X} \uplus (1-\alpha_X)\overline{X}. \tag{12}$$

Further, this unique α_X *is* $\left(\prod_{g=1}^{K} t^1_g\right)^{1/K} + \left(\prod_{g=1}^{K} t^2_g\right)^{1/K}$.[18]

[18] This is less than or equal to 1 since the geometric average of a set of numbers can be no greater

Proof. Assume first that $t_g^i \leq 1/2$ for $i = 1, 2$ and $g = 1, ..., K$. Uniqueness follows from Lemma 5, part 2, so we need only to show the existence of α_X. If $t_g^i = 0$ for any i or g then the result follows from Claim 4. Assume WLOG that $\prod_{g=1}^{K} t_g^1 \leq \prod_{g=1}^{K} t_g^2$. Define $\widetilde{t_g^i} = t_g^i/(1 - t_g^2)$ for $g = 1, ..., K$ and $i = 1, 2$. Note that $\prod_{g=1}^{K} \widetilde{t_g^1} \leq \prod_{g=1}^{K} \widetilde{t_g^2}$. Define $\tau = \left(\prod_{g=1}^{K} \frac{t_g^1}{t_g^2}\right)^{1/K} = \left(\prod_{g=1}^{K} \frac{\widetilde{t_g^1}}{\widetilde{t_g^2}}\right)^{1/K} \leq 1$. We can write

$X = \langle t^1, (1 - t_1^1 - t_1^2, 0, ..., 0), (0, 1 - t_2^1 - t_2^2, 0, ..., 0), ..., (0, ..., 0, 1 - t_K^1 - t_K^2) \rangle \uplus \langle t^2 \rangle$.

By SI

$$X \sim Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle \tag{13}$$

where[19]

$$\begin{aligned} Y &= \left\langle \left(\widetilde{t_1^1}, ..., \widetilde{t_K^1}\right), \left(1 - \widetilde{t_1^1}, 0, ..., 0\right), ..., \left(0, ..., 0, 1 - \widetilde{t_K^1}\right) \right\rangle \\ &= \widetilde{t^1} * \underline{X} \uplus (\mathbf{1} - \widetilde{t^1}) * \overline{X} \end{aligned}$$

By Claim 4,

$$Y \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y)\overline{X}. \tag{14}$$

where $\alpha_Y = \left(\prod_{g=1}^{K} \widetilde{t_g^1}\right)^{1/K}$. Define

$$Y' = \tau\widetilde{t^2} * \underline{X} \uplus (\mathbf{1} - \tau\widetilde{t^2}) * \overline{X}. \tag{15}$$

We must verify that all entries in Y' are nonnegative. This holds if $\tau\widetilde{t_g^2} \leq 1$ for all g. Since $t_g^2 \leq 1/2$ for all g, it follows that $\widetilde{t_g^2} \leq 1$; since $\tau \leq 1$ as well, it follows that $\tau\widetilde{t_g^2} \leq 1$.

---

than their arithmetic average:

$$\left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K} \leq \frac{1}{K}\sum_{g=1}^{K} t_g^1 + \frac{1}{K}\sum_{g=1}^{K} t_g^2 = \frac{1}{K}\sum_{g=1}^{K}\left(t_g^1 + t_g^2\right) \leq \frac{1}{K}\sum_{g=1}^{K} 1 = 1$$

[19] We must check that Y has no negative entries. Since X cannot have negative entries, it must be that $t_g^1 + t_g^2 \leq 1$ for all g. Since in addition $t_g^2 < 1$ for all g, it follows that $\frac{t_g^1}{1 - t_g^2} \leq 1$ for all g. Hence, all entries in Y are nonnegative.

Since $\left(\prod_{g=1}^{K} \tau\widetilde{t_g^2}\right)^{1/K} = \left(\prod_{g=1}^{K} \widetilde{t_g^1}\right)^{1/K} = \alpha_Y$, by Claim 4,

$$Y' \sim \alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X}. \tag{16}$$

It follows from (14) and (16) that $Y \sim Y'$. As a result,

$$\begin{aligned}
X &\sim Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by (13)} \\
&\sim Y' \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by IND} \\
&\sim \tau\widetilde{t^2} * \underline{X} \uplus (\mathbf{1} - \tau\widetilde{t^2}) * \overline{X} \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by (15)} \\
&\sim (\tau+1)\,\widetilde{t^2} * \underline{X} \uplus (\mathbf{1} - \tau\widetilde{t^2}) * \overline{X} && \text{by NDP} \\
&\sim (\tau+1)\,t^2 * \underline{X} \uplus (\mathbf{1} - (\tau+1)\,t^2) * \overline{X} && \text{by SI and definiton of } \widetilde{t^2}.
\end{aligned}$$

Therefore, using Claim 4, $X \sim \alpha_X \underline{X} \uplus (1-\alpha_X)\overline{X}$, where

$$\alpha_X = (\tau+1)\left(\prod_{g=1}^{K} t_g^2\right)^{1/K} = \left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K}.$$

Consider now the case of general $t^1, t^2 \in [0,1]^2$. Define $\widehat{t^i} = \frac{1}{2}t^i$ for $i = 1, 2$. Let

$$\widehat{X} = \langle \widehat{t^1}, \widehat{t^2}, (1-\widehat{t_1^1}-\widehat{t_1^2}, 0, ..., 0), (0, 1-\widehat{t_2^1}-\widehat{t_2^2}, 0, ..., 0), ..., (0, ..., 0, 1-\widehat{t_K^1}-\widehat{t_K^2}) \rangle$$

Each entry in each vector is at most one half. By the preceding argument, there is a unique $\widehat{\alpha}_X \in [0,1]$ such that

$$\widehat{X} \sim \widehat{\alpha}_X \underline{X} \uplus (1-\widehat{\alpha}_X)\overline{X}. \tag{17}$$

and this unique $\widehat{\alpha}_X$ is $\left(\prod_{g=1}^{K} \widehat{t_g^1}\right)^{1/K} + \left(\prod_{g=1}^{K} \widehat{t_g^2}\right)^{1/K}$. Further note that by NDP, $\widehat{X} \sim \frac{1}{2}X \uplus \frac{1}{2}\overline{X}$. Therefore

$$\begin{aligned}
\frac{1}{2}X \uplus \frac{1}{2}\overline{X} &\sim \widehat{\alpha}_X \underline{X} \uplus (1-\widehat{\alpha}_X)\overline{X} \\
&\sim \frac{1}{2}\,(2\widehat{\alpha}_X)\,\underline{X} \uplus (1-\frac{1}{2}\,(2\widehat{\alpha}_X))\overline{X} \\
&\sim \frac{1}{2}\,(2\widehat{\alpha}_X)\,\underline{X} \uplus \frac{1}{2}(1-(2\widehat{\alpha}_X))\overline{X} \uplus \frac{1}{2}\overline{X}
\end{aligned}$$

where the last line follows from NDP. Finally, by IND and SI

$$X \sim (2\widehat{\alpha}_X)\,\underline{X} \uplus (1-(2\widehat{\alpha}_X))\overline{X}$$

which means that the unique α_X that we are looking for is $2\widehat{\alpha}_X = \left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K}$. Q.E.D. ■

Lemma 6 *For every city X there is a unique $\alpha_X \in [0,1]$ such that*

$$X \sim \alpha_X \underline{X} \uplus (1-\alpha_X)\overline{X}.$$

Further, this unique α_X is $\sum_{n\in N(X)} \left(\prod_{g=1}^{K} t_g^n\right)^{1/K}$.[20]

Proof. We say that a neighborhood is a *ghetto* if all its residents belong to the same group. By SI it is enough to prove the statement for cities where all groups have a population measure of one. Also, by NDP we can restrict attention to cities where for each group there is at most one ghetto. The proof is by induction on the number of non-ghetto neighborhoods. Claims (4) and (5) already show the that the statement is true for cities with at most two non-ghetto neighborhoods. Assume that the statement of the theorem holds for all cities with $m-1$ non-ghetto neighborhoods, let

$$X = \langle t^1, \cdots, t^m, (1-\sum_{n=1}^{m} t_1^n, 0, ..., 0), (0, 1-\sum_{n=1}^{m} t_2^n, 0, ..., 0), ..., (0, ..., 0, 1-\sum_{n=1}^{m} t_K^n)\rangle$$

be a city with m non-ghetto neighborhoods, and let t^m be one of them. Then one can write

$$X = Y \uplus \langle t^m \rangle$$

where Y denotes X with neighborhood t^m removed. Y has $m-1$ non-ghetto neighborhoods. By SI

$$Y \uplus \langle t^m \rangle \sim \left[\left(\frac{1}{1-t_1^m}, ..., \frac{1}{1-t_K^m}\right) * Y\right] \uplus \left\langle \left(\frac{t_1^m}{1-t_1^m}, ..., \frac{t_K^m}{1-t_K^m}\right)\right\rangle$$

[20] By the reasoning given in footnote 18, α_X must lie between zero and one.

By the induction hypothesis, $\left(\frac{1}{1-t_1^m}, ..., \frac{1}{1-t_K^m}\right) * Y \sim \alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X}$ where

$$\alpha_Y = \sum_{n=1}^{m-1} \left(\prod_{g=1}^{K} \frac{t_g^n}{1-t_g^m} \right)^{1/K}.$$

Using (in order) IND, SI, and Claim 5,

$$\begin{aligned}
&\left[\left(\frac{1}{1-t_1^m}, ..., \frac{1}{1-t_K^m}\right) * Y\right] \uplus \left\langle \left(\frac{t_1^m}{1-t_1^m}, ..., \frac{t_K^m}{1-t_K^m}\right)\right\rangle \\
&\sim \alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X} \uplus \left\langle \left(\frac{t_1^m}{1-t_1^m}, ..., \frac{t_K^m}{1-t_K^m}\right)\right\rangle \\
&\sim (1-t_1^m, ..., 1-t_K^m) * \left(\alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X}\right) \uplus \langle t^m \rangle \\
&\sim \alpha_X \underline{X} \uplus (1-\alpha_X)\overline{X}
\end{aligned}$$

where

$$\begin{aligned}
\alpha_X &= \left(\prod_{g=1}^{K} [1-t_g^m]\right)^{1/K} \alpha_Y + \left(\prod_{g=1}^{K} t_g^m\right)^{1/K} \\
&= \left(\prod_{g=1}^{K} [1-t_g^m]\right)^{1/K} \sum_{n=1}^{m-1} \left(\prod_{g=1}^{K} \frac{t_g^n}{1-t_g^m}\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^m\right)^{1/K} \\
&= \sum_{n=1}^{m} \left(\prod_{g=1}^{K} t_g^n\right)^{1/K}
\end{aligned}$$

Q.E.D. ■

Now define the function S on the set of cities by $S(X) = 1 - \alpha_X$, where for each X, α_X is the unique number identified in Lemma 6. By Lemmas 5 and 6, for any cities X and Y, $X \succcurlyeq Y$ if and only if $S(X) \geq S(Y)$. This function (which is just $A(X)$) thus represents the relation $\succcurlyeq$. Q.E.D.

Proof of Proposition 1. By inspection, A satisfies the Cardinalization Principle. By Theorem 1, there is a unique segregation ordering $\succcurlyeq$ that satisfies the five axioms, and it is represented by A. Let S be another index that represents $\succcurlyeq$ and that satisfies the Cardinalization Principle. We must show that $S(X) = A(X)$ for any city $X \in \mathcal{C}_K$. By Lemma 6, there is an $\alpha_X \in [0,1]$ such that $X \sim X^{\alpha_X}$. Hence, $S(X) = S(X^{\alpha_X})$, which equals α_X by the Cardinalization Principle. But $A(X)$ also equals $A(X^{\alpha_X})$, which equals α_X by the Cardinalization Principle. Hence, $S(X) = A(X)$. Q.E.D.

Proof of Theorem 3. We first define two transformations of cities. Each produces a city in which each group has size one. The first does so by subdividing the groups; the second by scaling the groups.

1. For any city $X \in \mathcal{C}^I$, define the "flattened" city $\phi(X)$ in which each group g in X has been divided into T_g subgroups of size one, such that each subgroup has the same distribution across neighborhoods in $\phi(X)$ as g has in X. More formally: for each neighborhood $n \in N(X)$ the corresponding neighborhood n in $\phi(X)$ contains T_g groups for each nonempty group g in X . Each of these T_g groups contains T_g^n/T_g members. Note that the city $\phi(X)$ contains a total of T groups. For example, if $X = \{(1,2), (0,1)\}$, then
$$\phi(X) = \left\{ \left(1, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right), \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \right\}.$$

2. For any city $X \in \mathcal{C}$, let $\nu(X)$ be the "normalized" city that results from scaling each group so that its size is one. In the normalized city, the number of members of any group g in each neighborhood n equals T_g^n/T_g. For example, if $X = \{(1,2), (0,1)\}$, then
$$\nu(X) = \left\{ \left(1, \frac{2}{3}\right), \left(0, \frac{1}{3}\right) \right\}.$$

We first state, without proof, some simple properties of these transformations.

Lemma 7 *1. Assume that $\succcurlyeq$ satisfies GDP. Then for all $X \in \mathcal{C}^I$, $\phi(X) \sim X$.*

2. If $\succcurlyeq$ satisfies WGDP and WSI, then for all $X \in \mathcal{C}^I$ and for all positive integers α, $\phi(X) \sim \phi(\alpha X)$.[21]

[21] Note that $\phi(\alpha\phi(X)) = \phi(\alpha X)$. Then,by a sequential application of WSI, and WGDP,
$$\phi(X) \sim \alpha\phi(X) \sim \phi(\alpha\phi(X)) = \phi(\alpha X).$$

3. *For any city $X \in \mathcal{C}$, and for any permutation σ,*

$$\nu(\sigma(X)) = \sigma(\nu(X)).$$

4. *For all $X \in \mathcal{C}^I$, $\nu(\phi(X)) = \phi(X)$.*

5. *For all $X \in \mathcal{C}$, $\phi(\nu(X)) = \nu(X)$.*

Definition 2 *For any segregation ordering $\succcurlyeq$ on $\mathcal{C}^I$ define the associated scale invariant ordering on $\mathcal{C}^A$, denoted $\succcurlyeq^\psi$, as follows:*

$$\text{for any } X, Y \in \mathcal{C},\ X \succcurlyeq^\psi Y \text{ if and only if } \nu(X) \succcurlyeq \nu(Y). \tag{18}$$

Claim 6 *For any $\succcurlyeq$ defined on $\mathcal{C}^I$ that satisfies GS, WSI, NDP, GDP, WIND, and DN, the associated scale invariant ordering $\succcurlyeq^\psi$ on $\mathcal{C}^A$ satisfies GS, SI, NDP, WGDP, IND, and N.*

Proof. GS: Let X be a city with the set G of nonempty groups and let $\sigma : G \to G$ be a permutation. We must verify that $X \sim^\psi \sigma(X)$. This holds iff $\nu(X) \sim \nu(\sigma(X))$. But $\nu(\sigma(X)) = \sigma(\nu(X))$ by Lemma 7, part 3. The result follows since $\sim$ satisfies GS.

SI: Holds by construction.

NDP: Assume that Y is obtained from X by dividing some neighborhood in X into two neighborhoods. Then $\nu(Y)$ is obtained from $\nu(X)$ after the corresponding neighborhood in $\nu(X)$ is subdivided in the same way. Since $\succcurlyeq$ satisfies NDP, $\nu(Y) \succcurlyeq \nu(X)$; hence, by definition of $\succcurlyeq^\psi$, $Y \succcurlyeq^\psi X$. If the two neighborhoods in Y have the same group distribution or one is empty, then this is also true of $\nu(Y)$. In this case, $\nu(Y) \sim \nu(X)$ by NDP. By definition of $\succcurlyeq^\psi$, $Y \sim^\psi X$.

WGDP: Let X be a city in which the set of groups is G. Let X' be the result of partitioning each group $g \in G$ into M equal-sized groups, g_1 through g_M, where the subgroups g_1 through g_M have the same distribution across neighborhoods as g itself. By WSI, $\nu(X) \sim M\nu(X)$. Note that if we partition the groups of $M\nu(X)$ in the above way, we obtain $\nu(X')$. Therefore; by GDP, $M\nu(X) \sim \nu(X')$. Hence, $\nu(X) \sim \nu(X')$, which by definition of $\succcurlyeq^{\psi}$ implies $X' \sim^{\psi} X$.

IND: Let X and Y be two cities with the same set G of nonempty groups. Suppose X and Y have the same group distributions and the same total populations. Let $Z \in \mathcal{C}^A$ be such that $G(Z) \subset G$. We must show that

$$X \succcurlyeq^{\psi} Y \Longleftrightarrow X \uplus Z \succcurlyeq^{\psi} Y \uplus Z. \tag{19}$$

By definition of $\succcurlyeq^{\psi}$, equation (19) is equivalent to

$$\nu(X) \succcurlyeq \nu(Y) \Longleftrightarrow \nu(X \uplus Z) \succcurlyeq \nu(Y \uplus Z) \tag{20}$$

For each $g \in G(Z)$, define α_g to be $\frac{1}{T_g(X)} = \frac{1}{T_g(Y)}$. Let $\overrightarrow{\alpha} = (\alpha_g)_{g \in G(Z)}$. Then $\nu(X \uplus Z) = \nu\left(\nu\left(X\right) \uplus \omega(Z)\right)$ where $\omega(Z) = \overrightarrow{\alpha} * Z$. Hence, to show (20) it suffices to prove that

$$\nu(X) \succcurlyeq \nu(Y) \Longleftrightarrow \nu(\nu(X) \uplus \omega(Z)) \succcurlyeq \nu(\nu(Y) \uplus \omega(Z)) \tag{21}$$

This holds since $\succcurlyeq$ satisfies WIND and $G(\omega(Z)) = G(Z) \subset G$.

N: Since $\succcurlyeq$ satisfies DN, for any $K \geq 2$ there exist cities X and Y in $\mathcal{C}^I$, each with exactly K residents, such that $X \succ Y$. By GDP, $\phi(X) \succ \phi(Y)$; by part 4 of Lemma 7, $\nu\left(\phi(X)\right) \succ \nu\left(\phi(Y)\right)$, so $\phi(X) \succ^{\psi} \phi(Y)$. We have produced two cities with K nonempty groups that $\succ^{\psi}$ ranks differently. ■

Definition 3 *For any segregation ordering $\succcurlyeq$ on $\mathcal{C}^A$, define the associated flattening invariant ordering on $\mathcal{C}^I$, denoted $\succcurlyeq^{-\psi}$, as follows:*

$$\text{for any } X, Y \in \mathcal{C}^I,\ X \succcurlyeq^{-\psi} Y \text{ if and only if } \phi\left(X\right) \succcurlyeq \phi\left(Y\right) \tag{22}$$

Claim 7 *For any $\succcurlyeq$ defined on $\mathcal{C}^A$ that satisfies GS, SI, NDP, WGDP, IND, and N, the associated flattening invariant ordering $\succcurlyeq^{-\psi}$ on $\mathcal{C}^I$ satisfies GS, WSI, NDP, GDP, WIND, and DN.*

Proof. GS: for any city $X \in \mathcal{C}^I$ and any permutation $\sigma : G \to G$, we must show that $X \sim^{-\psi} \sigma(X)$. By (22), this holds iff $\phi(X) \sim \phi(\sigma(X))$. $\phi(\sigma(X))$ is the result of flattening the city X after its groups have been permuted by σ. This is equivalent to permuting the groups of the flattened city $\phi(X)$: there is a permutation $\widehat{\sigma}$ of the groups of $\phi(X)$ such that $\widehat{\sigma}(\phi(X)) = \phi(\sigma(X))$. Since $\succcurlyeq$ satisfies GS, $\phi(X) \sim \widehat{\sigma}(\phi(X)) = \phi(\sigma(X))$.

WSI: for any city $X \in \mathcal{C}^I$ and any integral scalar $\alpha \geq 1$, we must show that $X \sim^{-\psi} \alpha X$. By (22), this holds iff $\phi(X) \sim \phi(\alpha X)$, which follows from part 2 of Lemma 7.

NDP: Let $X \in \mathcal{C}^I$ be a city and let n be a neighborhood of X. Let X' be the city that results from dividing n into two neighborhoods, n_1 and n_2. We must show that if either (a) at least one of n_1 and n_2 is empty or (b) n_1 and n_2 have the same demographic distributions (i.e., $\left(p_g^{n_1}\right)_{g\in G} = \left(p_g^{n_2}\right)_{g\in G}$), then $\phi(X') \sim \phi(X)$ (and hence $X' \sim^{-\psi} X$ by (22)); otherwise, $\phi(X') \succcurlyeq \phi(X)$ (and hence $X' \succcurlyeq^{-\psi} X$ by (22)). Under the operation ϕ, each group g in a city is split up into T_g subgroups of size one but with the same distribution across neighborhoods: instead of T_g^n members of group g who live in neighborhood n, there are T_g^n/T_g members of each of T_g subgroups of g who live in n. Hence, the city $\phi(X')$ results from splitting neighborhood n of city $\phi(X)$ into two neighborhoods, n_1 and n_2, where the number of members of each subgroup g_m of g in neighborhood n_i for $i = 1, 2$ is just $T_g^{n_i}/T_g$. Since $\succcurlyeq$ satisfies NDP, $\phi(X') \succcurlyeq \phi(X)$. If (a) holds, then either n_1 or n_2 in $\phi(X')$ must be empty; if (b) holds, then n_1 and n_2 have the same demographic distributions in $\phi(X')$ since for each subgroup m of each group g of X,

$$p_{g_m}^{n_1} = \frac{T_g^{n_1}/T_g}{T^{n_1}} = \frac{p_g^{n_1}}{T_g} = \frac{p_g^{n_2}}{T_g} = \frac{T_g^{n_2}/T_g}{T^{n_2}} = p_{g_m}^{n_2}$$

(This uses the definition $p_g^{n_i} = T_g^{n_i}/T^{n_i}$.) If (a) or (b) holds, then $\phi(X') \sim \phi(X)$ as $\succcurlyeq$ satisfies NDP.

GDP: Let $X \in \mathcal{C}^I$ be a city in which the set of groups is G. Let $X' \in \mathcal{C}^I$ be the result of partitioning some group $g \in G$ into two groups, g_1 and g_2, where the number of members of each group is an integer. We must show that if $\left(t_{g_1}^n\right)_{n\in N} = \left(t_{g_2}^n\right)_{n\in N}$ then $\phi(X') \sim \phi(X)$ and hence $X' \sim^{-\psi} X$. But $\phi(X')$ and $\phi(X)$ are the same up to a permutation of groups. The result follows from GS.

WIND: Suppose $X, Y \in \mathcal{C}^I$ have the same set of nonempty groups, each of which has size one and let $Z \in \mathcal{C}^I$. We must show that

$$X \succcurlyeq^{-\psi} Y \text{ if and only if } \nu(X \uplus Z) \succcurlyeq^{-\psi} \nu(Y \uplus Z) \tag{23}$$

or equivalently, by definition of $\succcurlyeq^{-\psi}$, that

$$\phi(X) \succcurlyeq \phi(Y) \text{ if and only if } \phi(\nu(X \uplus Z)) \succcurlyeq \phi(\nu(Y \uplus Z))$$

But

$$\begin{aligned}
\phi(X) \quad &\succcurlyeq \quad \phi(Y) \Longleftrightarrow X \succcurlyeq Y && \text{since } X \text{ and } Y \text{ are already flat} \\
&\Longleftrightarrow \quad X \uplus Z \succcurlyeq Y \uplus Z && \text{by IND of } \succcurlyeq \\
&\Longleftrightarrow \quad \nu(X \uplus Z) \succcurlyeq \nu(Y \uplus Z) && \text{by SI of } \succcurlyeq \\
&\Longleftrightarrow \quad \phi(\nu(X \uplus Z)) \succcurlyeq \phi(\nu(Y \uplus Z)) && \text{byLemma7, part 5.}
\end{aligned}$$

DN: Since $\succcurlyeq$ satisfies N, for any $K \geq 2$ there exist cities X and Y in $\mathcal{C}$, each with exactly K groups, such that $X \succ Y$. By SI, $\nu(X) \succ \nu(Y)$; by part 5 of Lemma 7, $\phi\left(\nu(X)\right) \succ \phi\left(\nu(Y)\right)$, so $\nu(X) \succ^{-\psi} \nu(Y)$. We have produced two cities each with K residents that $\succ^{-\psi}$ ranks differently. Q.E.D. ■

To prove Theorem 3, we first show that if $\succcurlyeq$ is a segregation ordering on $\mathcal{C}^I$ that satisfies WSI, GS, NDP, GDP, WIND, and DN, then it must be represented by A^W. By Claim 6 its associated scale invariant ordering $\succcurlyeq^{\psi}$ satisfies GS, SI, NDP, WGDP, IND, and N. By Theorem 2, A represents $\succcurlyeq^{\psi}$ on $\mathcal{C}^A$. Also, by definition, $\succcurlyeq$ and $\succcurlyeq^{\psi}$coinicide for all cities each of whose groups have a number of residents equal to 1.

Let $X, Y \in \mathcal{C}^I$. Then,

$$\begin{aligned} X \succcurlyeq Y &\Leftrightarrow \phi(X) \succcurlyeq \phi(Y) \\ &\Leftrightarrow \phi(X) \succcurlyeq^{\psi} \phi(Y) \\ &\Leftrightarrow A(\phi(X)) \geq A(\phi(Y)) \\ &\Leftrightarrow A^W(X) \geq A^W(Y) \end{aligned}$$

where the first line follows from the fact that $\succcurlyeq$ satisfies GDP (Lemma 7, part 1), the second line from the fact that $\succcurlyeq = \succcurlyeq^{\psi}$on cities each of whose groups have a number of residents equal 1, and the third line from the fact that $\succcurlyeq^{\psi}$is represented by A. The last line follows from the fact that $A^W(\cdot) = A(\phi(\cdot))$.

It remains to show that there exists a segregation ordering on $\mathcal{C}^I$ that satisfies WSI, GS, NDP, GDP, WIND, and DN. Consider the segregation order on $\mathcal{C}^A$ defined by the Atkinson index:

$$X \succcurlyeq Y \Longleftrightarrow A(X) \geq A(Y)$$

By Theorem 2, $\succcurlyeq$ satisfies GS, SI, NDP, WGDP, IND, and N. Therefore, by Claim 7, $\succcurlyeq^{-\psi}$ on $\mathcal{C}^I$ satisfies WSI, GS, NDP, GDP, WIND, and DN. Q.E.D.

References

[1] Atkinson, A.B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* 2: 244-63.

[2] Collins, Chiquita, and David R. Williams. 1999. "Segregation and Mortality: The Deadly Effects of Racism." *Sociological Forum* 74: 495–523.

[3] Cotter, David A., JoAnn DeFiore, Joan M. Hermsen, Brenda M. Kowalewski, and Reeve Vanneman. 1997. "All Women Benefit: The Macro-Level Effect of Occupational Integration on Gender Earnings Equality." *American Sociological Review* 62:714–734.

[4] Cover, Thomas M., and Joy A. Thomas. 1991. *Elements of Information theory.* New York: Wiley.

[5] Cutler, David, and Edward Glaeser. 1997. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112: 827– 872.

[6] Cutler, David, and Edward Glaeser and Jacob Vigdor. 1999. "The Rise and Decline of the American Ghetto" *Journal of Political Economy* 107: 455– 506.

[7] Echenique, Federico, and Roland G. Fryer, Jr. 2005. "On the Measurement of Segregation." Mimeo.

[8] Hutchens, Robert. 2001. "Numerical Measures of Segregation: Desirable Properties and Their Implications." *Mathematical Social Sciences.* 42:13-29.

[9] Hutchens, Robert. 2003. "One Measure of Segregation." Forthcoming, *International Economic Review.*

[10] Jahn, Julius, and Schmidt, Calvin F., and Schrag, Clarence. 1947. "The Measurement of Ecological Segregation." *American Sociological Review* 12:293–303.

[11] James, David R., and Karl E. Taeuber. 1985. "Measures of Segregation." *Sociological Methodology* 15:1-32.

[12] Kinder, Donald R., and Tali Mendelberg. 1995. "Cracks in American Apartheid: The Political Impact of Prejudice among Desegregated Whites." *Journal of Politics* 57:402–424.

[13] Lewis, Gregory B. 1996. "Gender Integration of Occupations in the Federal Civil Service: Extent and Effects on Male-Female Earnings." *Industrial and Labor Relations Review* 49:472–483.

[14] David A. Macpherson and Barry T. Hirsh, 1995. "Wages and Gender Composition: Why do Women's jobs Pay Less?" " *Journal of Labor Economics* 13:426–471.

[15] Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Racial Segregation."" *Social Forces* 67:281–315.

[16] Meldrum, Christina, and Susan Eaton. 1994. "Resegregation in Norfolk, Virginia: Does Restoring Neighborhood Schools Work?" Report of the Harvard Project on School Desegregation.

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

[18] Orfield, Gary. 1995. "Desegregation, Resegregation, and Education." Part I of *In Pursuit of a Dream Deferred: Linking Housing and Education.* Minneapolis: Institute on Race and Poverty.

[19] Reardon, Sean F., and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." *Sociological Methodology* 32:33-67.

[20] Sakoda. J. 1981. "A Generalized Index of Dissimilarity." *Demography* 18:245-50.

[21] Schiller, Bradley. 1995. *The Economics of Poverty and Discrimination, 6th ed.* Englewood Cliffs, NJ: Prentice-Hall.

[22] Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures."

Econometrica 39:613-25.

[23] Shorrocks, A. F. 1984. "Inequality Decompositions by Population Subgroups." *Econometrica* 48:1369-86.

[24] South, Scott J., and Richard B. Felson. 1990. "The Racial Patterning of Rape." *Social Forces* 69:71–93

[25] Telles, Edward E. 1995. "Structural Sources of Socioeconomic Segregation in Brazilian Metropolitan Areas." *American Journal of Sociology* 100:1199–1223.

[26] Theil, Henri, and Anthony J. Finezza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts." *Journal of Mathematical Sociology* 1:187-194.