Measuring Segregation*

David M. Frankel † Oscar Volij ‡

September 5, 2005

Abstract

This paper gives an axiomatic foundation for multigroup segregation indices. We understand segregation to be the tendency of demographic groups to be distributed differently across locations (such as neighborhoods, schools, or occupations). We prove that there is a unique segregation index that satisfies a set of basic properties. One of these properties is invariance to changes in the relative sizes of the different groups (due, e.g., to differential population growth) that preserve each group's distribution across locations. We also characterize an alternative index that lacks this property, but that is more robust to changes in group definitions.

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

Keywords: Segregation, segregation indices.

*We thank seminar participants at University of the Basque Country, Brown University, University of Haifa, University of Iowa, Iowa State University, Rice University, Tel Aviv University, the 2004 Summer Meetings of the Econometric Society, and the Murat Sertel Memorial Conference on Economic Theory.

†Iowa State University. Email: dfrankel@econ.iastate.edu

‡Iowa State University and Ben-Gurion University. Email: oscar@volij.co.il

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.

The literature on segregation measurement has generated over 20 different indices (Massey and Denton [15]). While some papers have analyzed the properties of various indices, very few of them have provided a full characterization, and none of these have used purely ordinal axioms. Further, the existing characterizations are essentially valid for the two-group case. In this paper we provide a full ordinal characterization of two segregation indices for the multigroup case. The second index appears to be new to the literature.

Axiomatizations are important because they decompose an index into few basic and independent defining properties and thus facilitate the comparison of different measures. Ordinal axioms are more appealing than cardinal ones because they refer to bilateral comparisons and not to their specific functional representations. Multigroup segregation orderings are important because they allow us to study units (cities, school districts, etc.) with more than two groups and to compare units with different numbers of groups.

To begin, we need a working definition of segregation. An influential paper by James and Taeuber [11] defines segregation as "the differences in the distribution of social groups, such as blacks and whites, among units of social organization such as schools." In a later paper, Massey and Denton [15] discern 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.

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

[2] See Meldrum and Eaton [16], Orfield [18], and Schiller [23].

[3] See Cutler and Glaeser [4], Collins and Williams [2], and Kinder and Mendelberg [12], respectively.

> Minority members may be distributed so that they are overrepresented in some areas and underrepresented in others, varying on the characteristic of evenness.

Massey and Denton's other dimensions of segregation are isolation from the majority group, concentration in a small area, centralization in the urban core, and clustering in a contiguous enclave. Indices that measure the last three dimensions require detailed geographic information about a city's neighborhoods. Most indices of evenness and isolation do not require such information; they can usually be computed using data on neighborhood demographics alone.

Following James and Taeuber, we understand segregation to be the tendency of demographic groups to have different distributions across neighborhoods. This agrees with Massey and Denton's dimension of evenness. 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 of a segregation ordering. We then prove that there is a unique segregation ordering with these properties.

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 $1 - (.4)^{1/2}\,(.1)^{1/2} - (.6)^{1/2}\,(.9)^{1/2} = 0.065$.

One of the properties of this index is Scale Invariance: the index 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. Researchers have justified this property based on the argument that differences in segregation should not be due purely to differences in citywide demographic compositions.

Scale Invariance has strong implications: a group's weight in the segregation index cannot depend on its size. This property can be undesirable in some cases. A recent example is the introduction of tiny, mixed-race groups in the 2000 U.S. Census. To illustrate, Table 1 depicts a city with two neighborhoods, A and B. Neighborhood A contains 90 blacks and 10 whites; in neighborhood B, the numbers are reversed. The distributions of each group across neighborhoods appear in the second panel. Since the city is symmetric, its segregation index is $1-2\left(\frac{9}{10}\right)^{1/2}\left(\frac{1}{10}\right)^{1/2} = 0.4$. Now suppose a black from neighborhood A and a white from neighborhood B are reclassified as belonging to a new mixed-race group. Table 2 depicts the city as it now appears.

	Population			Row Distributions		
	A	B	Total	A	B	Total
Blacks	90	10	100	$\frac{9}{10}$	$\frac{1}{10}$	1
Whites	10	90	100	$\frac{1}{10}$	$\frac{9}{10}$	1

Table 1: Old demographic schema.

	Population			Row Distributions		
	A	B	Total	A	B	Total
Blacks	89	10	99	$\frac{89}{99}$	$\frac{10}{99}$	1
Whites	10	89	99	$\frac{10}{99}$	$\frac{89}{99}$	1
Mixed	1	1	2	$\frac{1}{2}$	$\frac{1}{2}$	1

Table 2: New demographic schema.

The segregation of the city drops to $1 - 2\left(\frac{89}{99}\right)^{1/3}\left(\frac{10}{99}\right)^{1/3}\left(\frac{1}{2}\right)^{1/3} = 0.29$. Since a group's size does not matter, the small mixed-race group receives the same weight of 1/3 in the geometric averages as each of the two larger groups. A minor change in the segregation schema leads to a large decline in measured segregation.

Our second axiomatization avoids this problem by replacing Scale Invariance with a very different axiom: the Group Division Property. An index with this property does not change when an existing group is subdivided into two groups that have the same neighborhood distribution. For instance, if whites are divided into white females and white males, and these groups have the same distribution across neighborhoods, the index is unaffected. When we replace Scale Invariance with the Group Division Property, we again obtain a unique index. It equals one minus the sum, over all neighborhoods, of the *weighted* geometric means of the percentages of each group who live in the neighborhood, where the weight of each group equals that group's proportion in the city. This index appears to be new to the literature.

With this weighted index, the change of the racial schema in our earlier example has only a small effect on measured segregation. Since the proportions of blacks and whites are equal in Table 1, the weighed index of this city is still equal to $1 - 2\left(\frac{90}{100}\right)^{1/2}\left(\frac{10}{100}\right)^{1/2} = 0.4$. However, the weighted index of the city in Table 2 is $1 - 2\left(\frac{89}{99}\right)^{99/200}\left(\frac{10}{99}\right)^{99/200}\left(\frac{1}{2}\right)^{2/200} = 0.394$. Since there are only two mixed-race residents in a city of 200, they receive a weight of only 2/200 in the weighted segregation index as opposed to 1/3 in the unweighted index. Consequently, the weighted index falls by much less than the unweighted index when the mixed-race group is introduced.

Our approach is to give a rigorous foundation for segregation that can be applied in a variety of contexts. This "context-free" approach may not suit everyone. Some researchers may desire axioms that are inspired by social welfare considerations. We view social welfare and segregation as orthogonal concepts. While the segregation of blacks and whites in the U.S. is seen by many as harmful, segregation in other contexts is often viewed as leading to peaceful coexistence.

Other researchers may desire axioms that are motivated by the underlying social processes that generate segregation. We agree that the causes of segregation are an important issue. However, segregation can have different causes in different contexts: racial discrimination, Tiebout sorting, a preference for living with members of the same group, and so on. Thus, we need a measure of segregation that does not depend directly on any particular cause. By analogy, height in humans is usually influenced by nutrition and state of health. But some people are tall because they are sick. Hence, one needs a measure of height that does not depend directly on a person's health status.

Another advantage of a context-free approach is that it lets us make meaningful comparisons of segregation in different settings. For instance, Farley [7] reports the following striking comparison: in Detroit in 1970, segregation between post-college educated black men and post-college educated white men was greater than segregation between post-college educated white men and white men with no education whatsoever. There is no reason to think that these two types of segregation have the same welfare implications or the same causes and effects. Yet we are wiser for being able to make the comparison. It tells us something useful.

The first paper to study segregation axiomatically is Philipson [19]. It provides an axiomatic characterization of a large family of segregation orderings that have an additively separable representation. The representation consists of a weighted average of a function that depends on the neighborhood's demographic distribution only. The papers that are most closely related to the present one are Hutchens [8, 9]. These papers study the measurement of segregation in the case of two demographic groups. Hutchens [8] characterizes the family of indices that 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 [24, 25]) 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.[4]

Another related paper is Echenique and Fryer [6]. They 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. The paper also provides group-specific indices that measure one group's isolation from the other groups.

Our view is that evenness and isolation are each important in different contexts. Evenness reflects the extent to which the choices of demographic groups differ from each other. In a residential context, it is related to issues of Tiebout sorting, housing discrimination, and the effects of tastes on housing choice. In the context of occupational segregation, evenness can be used to study how the different occupational distributions of men and women affect male-female wage inequality. Isolation measures the tendency of an ethnic group to have most of its contacts with members of the same group. While isolation may help a group to develop a unique culture and identity, it may also create disadvantages in the labor market. Hence, isolation is a more relevant dimension for researchers who are interested in such phenomena as culture, identity, and minority labor market outcomes. The two dimensions may also be causally related. An absence of cross-racial social contact may strengthen a person's preference for living in neighborhoods in which her group is a majority. In this way, isolation may lead to a lack of evenness. Conversely, the tendency of ethnic groups to have sharply different neighborhood distributions, whether or not the differences are voluntary, may lead to ethnically isolated social networks.

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 and provide some known examples of segregation indices that represent various orderings in Section 3. Section 4 proposes our first set of axioms.

[4] For example, Hutchens [9] proposes that the segregation of a city that is composed of two areas (each of which may contain several neighborhoods) should equal the between-area segregation plus a weighted sum of segregation within the areas.

Our results appear in Section 5. The unweighted segregation index is characterized in sections 5.1 and 5.2. The weighted segregation index is axiomatized in Section 5.3. In Section 6 we prove independence of the axioms in each axiomatization. Longer proofs are relegated to the appendix.

2 Notation

We understand segregation as the tendency of demographic groups to be distributed differently across locations, such as neighborhoods, schools, or occupations. Hence, segregation is a property of *sets of locations*. If the locations are neighborhoods, then these sets are typically cities; if they are schools, then the sets are usually school districts; and if the locations are occupations, then the sets might be industries or geographic areas.

The assignment of individuals to locations is taken as a primitive of the model. In some cases this assignment may be unambiguous, such as when the locations are schools. In other cases, such as when the locations are neighborhoods, there may be more than one natural assignment: census tracts, city blocks, etc. We do not take a position on which assignment is best. Rather, we study how segregation should be measured for a *given* assignment of individuals to locations.

For concreteness, we will refer to locations as neighborhoods and to sets of locations as cities. Formally, we define a city as follows:

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 neighborhood $n \in N$, a nonnegative number T_g^n (the number of members of group g that reside in neighborhood n), such that the total population of group g in the city is positive: for all $g \in G$, $\sum_{n \in N} T_g^n > 0$.

For instance, in the city X depicted in Table 3, $G(X) = \{Black, White\}$, $N(X) = \{A, B\}$, $T_{Black}^A = 5$, and so on.

We will sometimes use the following more compact notation. The expression $\langle (1,2), (3,1) \rangle$ denotes a city with two groups (e.g., blacks and whites) and two neighborhoods. The first neigh-

City X		
	A	B
Blacks	5	2
Whites	1	4

Table 3: Example of a city.

borhood, $(1, 2)$, 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:

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

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. Neighborhood n in city X is *representative* if the group 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*.

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. 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 nonnegative 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 $\overrightarrow{\alpha} = (2,3)$, then $\overrightarrow{\alpha} * X = \langle (2,6), (6,12) \rangle$. We sometimes apply the same operation to individual neighborhoods; e.g., $\overrightarrow{\alpha} * (1,2) = (2,6)$.

For any two cities X and Y, $X \uplus Y$ denotes the result of adjoining Y to X. Its neighborhoods consist of the neighborhoods of X and Y; its groups are simply the groups of X and Y.[5] For instance, consider the one-neighborhood cities $X = \langle (1,2) \rangle$, and $Y = \langle (1,2) \rangle$. If the two cities contain the same two groups (e.g., blacks and whites), then $X \uplus Y = \langle (1,2), (1,2) \rangle$. If instead X contains a black and two whites while Y consists of a black and two Asians, then the combined city will include three groups: $X \uplus Y = \langle (1,2,0), (1,0,2) \rangle$. (The order in each neighborhood is blacks, whites, Asians.)

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 $K \geq 1$ nonempty 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.[6]

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

We impose axioms not on the segregation index but on the underlying segregation ordering. These approaches are not equivalent. As in utility theory, a segregation ordering may be represented by more than one index, and there are segregation orderings that are not captured by any index.

[5] Formally, let $X = \left\langle (T_g^n)_{g \in G} \right\rangle_{n \in N}$ and $Y = \left\langle (T_g^n)_{g \in G'} \right\rangle_{n \in N'}$ with disjoint set of neighborhoods. $X \uplus Y$ denotes the city $\left\langle (T_g^n)_{g \in G \cup G'} \right\rangle_{n \in N \cup N'}$.

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

3.1 Examples of segregation indices

We now discuss several examples of segregation indices, each of which represents a particular ordering. We begin with the Atkinson index:

Atkinson

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

When X contains exactly two nonempty groups, this ordering coincides with that of the usual Atkinson index with parameter 1/2 (Massey and Denton [15, p. 286]).[7] The Atkinson index is derived from the income inequality measure of the same name (Atkinson [1]). The Atkinson ordering (the ordering represented by A) is the unique ordering that satisfies our first set of basic properties (Theorems 1 and 2).

We also define a weighted version of the Atkinson index:

Weighted Atkinson

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

The Atkinson index (1) equals one minus the sum of the *unweighted* geometric averages of the group proportions in each neighborhood. In contrast, the weighted Atkinson index 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 belong to that group. While the Atkinson index treats each group the same regardless of its size, the weighted Atkinson index gives more weight to larger groups. The weighted Atkinson ordering is the unique ordering that satisfies our second set of basic properties (Theorem 3).

The next index is a multigroup version of the Index of Dissimilarity:

Dissimilarity

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

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

In the case of two groups, the Dissimilarity index 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 Dissimilarity index was first discussed by Jahn et al [10]. It was used by Cutler, Glaeser, and Vigdor [5] to measure the evolution of segregation in American cities.

Some researchers have used a weighted version of the Dissimilarity index.

Weighted Dissimilarity

$$D^W(X) = \sum_{n \in N(X)} \sum_{g \in G(X)} P_g \left| t_g^n - P^n \right|$$

Weighted Dissimilarity was used by Rhode and Strumpf [21] to assess the empirical importance of the Tiebout sorting model.[8]

4 Axioms

We state our axioms with respect to an unspecified class of cities, $\mathcal{C}$. Later we will apply them to particular classes. 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 $\mathcal{C}$ is closed under the transformation τ if, for all cities X in the class, the transformed city $\tau(X)$ is also in $\mathcal{C}$. Let $\mathcal{C}$ be a class of cities that is closed under the transformation τ. Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$, with the associated "equally segregated" relation $\sim$. We will say that the segregation of a city (under $\succcurlyeq$) is invariant to the 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. Implicit in this is the requirement that $\mathcal{C}$ be closed under the transformation.

The first axiom, Group Symmetry, is an essential property of a context-free measure of segregation. It states that the level of segregation in a city does not depend on the identities of the city's demographic groups; it depends only on the number of each group who live in each neighborhood. For instance, the cities depicted in Table 4 are equally segregated.

[8] This index was introduced by Morgan [17] and Sakoda [22]. In the Dissimilarity index, the neighborhood functions f assign equal weight $1/|G(X)|$ to each group. Weighted Dissimilarity results from changing this weight to the proportion of the group in the city, P_g: $D^W(X) = \sum_{n \in N(X)} f^W(t^n)$ where $f^W(t^n) = \sum_{g \in G(X)} P_g \Big| t_g^n - \sum_{g' \in G(X)} P_{g'} t_{g'}^n \Big| = \sum_{g \in G(X)} P_g \left| t_g^n - P^n \right|$.

City 1			City 2		
	A	B		A	B
Blacks	5	2	Protestants	5	2
Whites	1	4	Catholics	1	4

Table 4: Group Symmetry implies that these cities are equally segregated.

City 2 results from relabeling the groups of city 1. More generally, we define a relabeling as follows:

Definition 2 Let X be a city. A relabeling of the groups $G(X)$ is a one to one function $\sigma : G' \to G(X)$, where G' is some set of groups such that $|G'| = |G(X)|$. Given a relabeling σ, we define $\sigma(X)$ to be a city with the same set of neighborhoods as X and with the set G' of nonempty groups, such that for all groups $g' \in G'$, the number of members of group g' in each neighborhood n of $\sigma(X)$ equals the number of members of group $\sigma(g')$ in the corresponding neighborhood n of X.

The axiom of Group Symmetry states that the segregation of a city is unaffected by any relabeling of its groups.

Group Symmetry (GS) Let X be a city and let σ be some relabeling of the groups $G(X)$. Then $X \sim \sigma(X)$.

Segregation refers to the tendency of different groups to have different distributions across neighborhoods. This leaves open the question of whether segregation should also depend on the absolute sizes of the groups. The next axiom, Scale Invariance, gives a negative answer to this question. This axiom implies, for instance, that the two cities in Table 5 are equally segregated.

City 1			City 2		
	A	B		A	B
Blacks	5	1	Blacks	5	1
Whites	2	4	Whites	20	40

Table 5: Scale Invariance implies that these cities are equally segregated.

Scale Invariance is one of the five requirements that Jahn *et al* [10] say a satisfactory measure of segregation should satisfy.[9]

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 $\overrightarrow{\alpha} = (\alpha_g)_{g \in G(X)}$, $X \sim \overrightarrow{\alpha} * X$.

In section 5.3, we study an alternative axiomatization that does not satisfy Scale Invariance.

The next axiom is the Neighborhood Division Property. The central idea of this axiom is that if the partition of a city into neighborhoods becomes finer, one may detect more segregation, but not less. The axiom states, more precisely, that if two neighborhoods are combined, then segregation is unaffected if the two neighborhoods have the same group distribution; otherwise, segregation weakly falls. Examples appear in Tables 6 and 7. In the first panel of each table there are three neighborhoods, A, B, and C; the second panel depicts the result of combining A and B into a single neighborhood. Assume that neighborhood C always contains the same unspecified numbers of blacks and whites. Since A and B have different proportions of blacks in Table 6, NDP implies only that combining them does not raise segregation. A case in which A and B have the same proportion of blacks appears in Table 7. In this case, NDP implies the stronger property that combining A and B has no effect on the level of segregation in the city. All indices described in the previous section satisfy NDP.

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

[9] 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 intermediate scores cover most of the possible range between the extremes of 0 and 100.

Property (3) is Scale Invariance.

City 1				City 1′		
	A	B	C		A&B	C
Blacks	1	2	T^C_{Black}	Blacks	3	T^C_{Black}
Whites	2	1	T^C_{White}	Whites	3	T^C_{White}

Table 6: The Neighborhood Division Property implies that combining neighborhoods A and B does not raise segregation (although it may lower it).

City 2				City 2′		
	A	B	C		A&B	C
Blacks	1	2	T^C_{Black}	Blacks	3	T^C_{Black}
Whites	2	4	T^C_{White}	Whites	6	T^C_{White}

Table 7: Since A and B have the same group distributions, the Neighborhood Division Property implies that combining them does not affect the level of segregation.

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., $p_g^{n_1} = p_g^{n_2}$ for all $g \in G$), then $X' \sim X$. Otherwise, $X' \not\succcurlyeq X$.

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, the city's level of segregation does not change. The second 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.[10] But while NDP also includes the case of more than two groups, the transfer principle cannot be easily generalized to this case.[11] 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 perhaps segregation has fallen.

The next axiom is Independence. It states that if two cities have the same group distribution and the same total population, then adjoining a given set of neighborhoods to each of them does not affect which of the two cities is more segregated. Consider, for instance, the two cities depicted in Table 8. The result of adjoining a given, unspecified set of additional neighborhoods to both cities appears in Table 9. The two original cities have the same total size and the same group distribution (3/8 black, 5/8 white). The axiom of Independence implies that adjoining the new neighborhoods preserves the segregation order of the cities. If the city in the left panel in Table 8 is more segregated than the city on the right, then the same is true in Table 9.

An intuition for Independence is as follows. Since the two original cities 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.

[10] 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: $p^{n_1}_{Black} > p^{n_2}_{Black}$. Neighborhood n in city X can be written $\left(T^{n_1}_{Black} + T^{n_2}_{Black}, T^{n_1}_{White} + T^{n_2}_{White}\right)$. Split this neighborhood into two neighborhoods with identical percents black: $n_1' = \left(\left[T^{n_1}_{Black} + T^{n_2}_{Black}\right] \frac{T^{n_1}_{White}}{T^{n_1}_{White} + T^{n_2}_{White}}, T^{n_1}_{White}\right)$ and $n_2' = \left(\left[T^{n_1}_{Black} + T^{n_2}_{Black}\right] \frac{T^{n_2}_{White}}{T^{n_1}_{White} + T^{n_2}_{White}}, T^{n_2}_{White}\right)$. 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 $p^{n_1}_{Black}$ and $p^{n_2}_{Black}$. 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, $X' \succ X$.

[11] One attempt to do so appears in Reardon and Firebaugh [20].

City 1			City 2		
	A	B		A	B
Blacks	3	0	Blacks	2	1
Whites	3	2	Whites	2	3

Table 8: Two cities with equal sizes and equal group distributions.

City 1′						City 2′					
	A	B	C	D	...		A	B	C	D	...
Blacks	3	0	T^C_{Black}	T^D_{Black}	...	Blacks	2	1	T^C_{Black}	T^D_{Black}	...
Whites	3	2	T^C_{White}	T^D_{White}	...	Whites	2	3	T^C_{White}	T^D_{White}	...

Table 9: Result of adding the neighborhoods C, D, etc. to the cities in Table 8.

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 the same total populations but very different group distributions: $X = \langle(1, 100)\rangle$ and $Y = \langle(100, 1)\rangle$. Since Y is just X with the groups relabeled, they are equally segregated by Group Symmetry. Now consider adjoining $Z = \langle(100, 1)\rangle$ to each of them. Each neighborhood in $Y \uplus Z = \langle(100, 1), (100, 1)\rangle$ is representative of the city, while the neighborhoods of $X \uplus Z = \langle(1, 100), (100, 1)\rangle$ are very unrepresentative of their city. Hence, the city $X \uplus Z$ seems to be more segregated than the city $Y \uplus Z$: in a sensible segregation ordering, the addition of Z changes the relative segregation of X and Y. To allow orderings with this property, the axiom of Independence applies only 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 total populations:

$$X = \langle (90000, 10000), (10000, 90000) \rangle$$
$$Y = \langle (1, 0), (0, 1) \rangle$$
$$Z = \langle (100, 100) \rangle$$

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 applies only to cities X and Y that have the same size.

Our final axiom, and the most trivial, is the axiom of Nontriviality. It states that there exist two cities, one strictly more segregated than the other. More precisely:

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

This section has three parts. Subsection 5.1 shows that there is a unique ordering that satisfies the preceding five axioms when the number of groups is fixed. This ordering is represented by the Atkinson Index. Subsection 5.2 extends this result to the case of a variable number of groups. This extension relies on a new axiom, the Weak Group Division Property. In subsection 5.3, we strengthen this axiom while weakening Scale Invariance. We show once again that there is a unique ordering. This ordering is represented by the Weighted Atkinson Index.

5.1 Fixed number of groups

In this section we consider the class $\mathcal{C}_K$ of cities that contain exactly K nonempty groups. Fix a city $X \in \mathcal{C}_K$. For any neighborhood n in X, let $t^n(X) = \left(t^n_g(X)\right)_{g \in G(X)}$ denote the vector of the proportions of each group that live in n. For example, in the city $X = \langle (3, 0), (7, 10) \rangle$, which is in

$\mathcal{C}_2$, 3/10 of the blacks and none of the whites live in the first neighborhood, so $t^1(X) = (3/10, 0)$. We will omit the argument "X" when the city is clear from the context.

Our first result is that the five axioms of Section 4 are jointly satisfied on $\mathcal{C}_K$ by a unique segregation ordering, which is represented by the Atkinson index A.

Theorem 1 *The Atkinson ordering on $\mathcal{C}_K$ is the only ordering that satisfies GS, SI, NDP, IND, and N 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 the following intuitive property. Let $\underline{X}^K$ be a city with $K \geq 1$ groups of unit size who all live in the same neighborhood: $\underline{X}^K = \langle(\underbrace{1, 1, \ldots, 1}_{K \text{ groups}})\rangle$. Let $\overline{X}^K$ be a city with $K \geq 2$ groups of unit size who all live in separate neighborhoods:

$$\overline{X}^K = \underbrace{\left\langle \Big(\underbrace{1, 0, \ldots, 0}_{K \text{ groups}}\Big), \Big(\underbrace{0, 1, 0, \ldots, 0}_{K \text{ groups}}\Big), ..., \Big(\underbrace{0, ..., 0, 1}_{K \text{ groups}}\Big) \right\rangle}_{K \text{ neighborhoods}}.$$

The property is as follows.

Definition 3 The index S satisfies the *Cardinalization Principle with K Groups* if the following conditions hold:

1. If $K = 1$, then $S\left(\underline{X}^1\right) = 0$.
2. If $K > 1$, then for any $\alpha \in [0, 1]$, $S\left(\alpha\overline{X}^K \uplus (1 - \alpha)\underline{X}^K\right) = \alpha$.

Part 1 states that a city with one group should have an index of zero. Part 2 considers a city with K equal sized groups, in which a proportion α of each group live in a completely segregated area and $1 - \alpha$ live in a completely integrated area. The principle states that the segregation index of such a city should be α. For instance, the Cardinalization Principle states that the index of the city in Table 10 should be 0.15.

Only the Atkinson index satisfies the Cardinalization Principle together with the five axioms of Theorem 1:

	A	B	C
Blacks	15	0	85
Whites	0	15	85

Table 10: A segregation index that satisfies the Cardinalization Principle must assign a value of 0.15 to this city.

Proposition 1 *For all $K \geq 1$, the Atkinson index is the unique segregation index on $\mathcal{C}_K$ that satisfies the Cardinalization Principle with K Groups and has an induced segregation ordering that satisfies GS, SI, NDP, IND, and N.*

The following proposition provides a useful interpretation of the Atkinson index: it can be rewritten as a weighted average of neighborhood-level segregation indices, where the weights are the neighborhoods' relative sizes. For $(c_1, ..., c_K) \in \mathbb{R}_+^K$, define $f^A(c_1, ..., c_K) = \big(\frac{1}{K}\sum_{i=1}^K c_i\big) - \big(\prod_{i=1}^K c_i\big)^{1/K}$.

Proposition 2 *The Atkinson Index can be written*

$$A(X) = \sum_{n \in N(X)} \frac{T^n}{T} f^A\Big(\frac{p_1^n}{P_1}, \dots, \frac{p_{|G|}^n}{P_{|G|}}\Big). \tag{4}$$

The segregation index of neighborhood n, $f^A\big(\frac{p_1^n}{P_1}, \dots, \frac{p_{|G|}^n}{P_{|G|}}\big)$, is nonnegative and equals zero if and only if neighborhood n is representative of city X.

Proof. One can verify that $\sum_{n \in N(X)} \frac{1}{K} \sum_{g \in G_K} t_g^n = 1$ by reversing the order of summation. Substituting this expression for 1 in (1),

$$A(X) = \sum_{n \in N(X)} \frac{1}{K} \sum_{g \in G_K} t_g^n - \sum_{n \in N(X)} \Bigg(\prod_{g \in G(X)} t_g^n\Bigg)^{\frac{1}{|G(X)|}} = \sum_{n \in N(X)} f^A(t^n)$$

Equation (4) follows since $t_g^n = \frac{p_g^n}{P_g}\frac{T^n}{T}$ and since f^A is homogeneous of degree one. Further, the neighborhood-level segregation index, $f^A\left(\frac{p_1^n}{P_1}, \dots, \frac{p_{|G|}^n}{P_{|G|}}\right)$, is nonnegative, since f^A is the arithmetic mean of its arguments less their geometric mean. It equals zero if and only if the ratios $\frac{p_g^n}{P_g}$ are all equal, in which case they must all equal one: the neighborhood must be representative of the city. This follows since the numerators and the denominators must both sum to one. Q.E.D.

This expresses the Atkinson index as a weighted average of neighborhood-level segregation indices, where the weight on neighborhood n, T^n/T, is the proportion of city residents who live in n. The neighborhood-level segregation index can be large even if the neighborhood is quite small. For instance, consider the city $X = \langle (1,0), (99,100) \rangle$. Substituting into (4),

$$\begin{aligned} A(X) &= \frac{1}{200} f\left(1/\frac{1}{2}, 0\right) + \frac{199}{200} f\left(\frac{99}{199}/\frac{1}{2}, \frac{100}{199}/\frac{1}{2}\right) \\ &= \frac{1}{200} * 1 + \frac{199}{200} * 0.0000126. \end{aligned}$$

The segregation index of the first neighborhood is 1; however, its weight in the average is only 1/200 since only one of the 200 city residents lives there. The segregation index of the second neighborhood, 0.0000126, is much lower; however, the neighborhood's weight is almost one since nearly all city residents live there.

5.2 Variable Number of Groups

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. 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, segregation does not change. For example, consider the two-neighborhood city $\langle (2,4), (6,8) \rangle$ with two demographic groups (say, blacks and whites). If we divide each group into two equal sized groups (e.g., males and females) which are identically distributed across neighborhoods, we obtain 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: $t^n_{g_m} = t^n_g$ for all $m = 1, ..., M$ and for all neighborhoods n.[12] If $X' \in \mathcal{C}$, then $X' \sim X$.

[12] 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

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

Theorem 2 *The Atkinson ordering on $\mathcal{C}^A$ is the unique ordering that satisfies GS, SI, NDP, IND, N, and WGDP on $\mathcal{C}^A$.*

Proof. By Theorem 1, for any K, the Atkinson ordering on $\mathcal{C}_K$ is the only ordering on $\mathcal{C}_K$ that satisfies the axioms GS, SI, NDP, IND, and N. We first show that the Atkinson ordering on $\mathcal{C}^A$ satisfies WGDP. Recall that $A(X) = 1 - \sum_{n \in N(X)} \Big(\prod_{g=1,\ldots,K} t_g^n \Big)^{1/K}$ 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

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

Thus, the Atkinson ordering 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.

$m \in \{1, \cdots, M\}$.

Theorem 2 shows that the Atkinson ordering is the only ordering that satisfies the five axioms of Theorem 1 plus WGDP when the number of groups is variable. Any strictly increasing transformation of A also represents the same ordering. However, for any $K \geq 2$, any such transformation must violate the Cardinalization Principle for K groups:

Proposition 3 *The Atkinson index is the unique representation of the Atkinson ordering on $\mathcal{C}^A$ that satisfies the Cardinalization Principle for some number of groups $K \geq 2$. Moreover, the Atkinson index satisfies the Cardinalization Principle for all $K \geq 1$.*

5.3 Axiom Set II: WSI and GDP

We understand "segregation" as the tendency of different races to have different distributions across neighborhoods. So far we have interpreted this to mean that segregation should be a function of these distributions alone: the relative sizes of the groups should not matter. This is the property of Scale Invariance. In some situations, however, we may be particularly interested in the tendency of *large* groups to be distributed differently from one another. To allow large groups to have more weight, we must drop Scale Invariance. But one implication of Scale Invariance does make sense: the overall size of a city should not matter. Accordingly, we replace Scale Invariance by the following weaker axiom.

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 scalar: for any city $X \in \mathcal{C}$ and any positive scalar α, if $\alpha X \in \mathcal{C}$ then $X \sim \alpha X$.

In order to obtain a unique index, we must strengthen one or more of the other axioms. A plausible candidate is the Weak Group Division Property, which requires that segregation remain unchanged when existing groups are subdivided in a particular way. This axiom is weak because it applies only when (i) all the existing groups are subdivided into the same number of subgroups, (ii) the subgroups of a given group have equal size, and (iii) the subgroups of a given group are identically distributed across neighborhoods.

In this section we drop conditions (i) and (ii) as they do not contribute to the axiom's plausibility. We retain condition (iii) as it is essential.[13] Accordingly, the following 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 retirees and white nonretirees, and these groups have the same distribution across neighborhoods, then measured segregation should not change (even if there are fewer retirees than nonretirees).

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: $t_{g_1}^n = t_{g_2}^n = t_g^n$ for all $n \in N$.[14] If $X' \in \mathcal{C}$, then $X' \sim X$.

The Atkinson ordering violates GDP. For example, the Atkinson index of the city $\langle (2,1), (0,1) \rangle$ is 0.29. GDP implies that this city is as segregated as the city $\langle (1,1,1), (0,0,1) \rangle$. However, the Atkinson index of the second city is 0.21.

With these changes to the axioms, we also need 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 requirement. 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.

[13] To see why, consider a city with two white inhabitants, each of whom occupies one neighborhood: $\langle (1), (1) \rangle$. This city is completely integrated. Assume the first white is, in fact, of Hispanic origin. If the demographic schema is modified to distinguish between Hispanic and non Hispanic whites, the city will now look very segregated: $\langle (1,0), (0,1) \rangle$. A reasonable segregation measure will be higher under the new schema. By retaining condition (iii), we do not rule out such measures.

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

In fact, we only need this property in a special case: when the nonempty groups in X and Y are all of size 1. We refer to such a city as a "normalized city". For any city $X \in \mathcal{C}^A$, 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 instance, the result of applying ν to the city $\langle(1,2),(2,3)\rangle$ is the city $\langle\left(\frac{1}{3},\frac{2}{5}\right),\left(\frac{2}{3},\frac{3}{5}\right)\rangle$. Weak Independence states that if X and Y are normalized cities with the same set G of nonempty groups and Z is a set of neighborhoods whose groups are all in G, then adjoining Z to both cities and then scaling the result to restore 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.

To facilitate the proof of our next theorem, we also modify Nontriviality slightly. Nontriviality gives conditions under which there are two cities with exactly K groups that are ranked differently. Dual Nontriviality gives conditions under which there are two cities with exactly T residents that are ranked differently.

Dual Nontriviality (DN) 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 $X \succ Y$.

In the class $\mathcal{C}^A$ of all cities, Dual Nontriviality is implied by Nontriviality (assuming Weak Scale Invariance).[15] However, this is not true for arbitrary classes of cities. For the purposes of proving Theorem 3, Dual Nontriviality is a more convenient formulation.

[15] Nontriviality implies that there are two cities $X, Y \in \mathcal{C}^A$ such that $X \succ Y$. For any T, we can scale the populations of X and Y so that they both have T residents. By Weak Scale Invariance, this rescaling does not change the segregation level of either city.

Until now we have let cities have group sizes that are not integers. Theorem 3 restricts to the set of cities with integral group sizes.[16] This restriction allows us to prove the theorem in a brief and intuitive way. Formally, we restrict attention to the set $\mathcal{C}^I \subset \mathcal{C}^A$ 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 weighted Atkinson ordering is the only one that satisfies the new set of axioms on $\mathcal{C}^I$.

Theorem 3 *The weighted Atkinson ordering on $\mathcal{C}^I$ is the unique ordering that satisfies GS, WSI, NDP, WIND, GDP, and DN on $\mathcal{C}^I$.*

Proof. The idea of the proof is as follows. Let $\mathcal{S}^I$ be the set of segregation orderings on $\mathcal{C}^I$ satisfying the axioms of Theorem 3, and let $\mathcal{S}^A$ be the set of segregation orderings on $\mathcal{C}^A$ satisfying the axioms of Theorem 2. We will define an isomorphism between these two sets. By Theorem 2, $\mathcal{S}^A$ is a singleton. By the existence of the isomorphism, $\mathcal{S}^I$ must also be a singleton. Finally, using properties of the isomorphism we will show that the unique ordering in $\mathcal{S}^I$ is represented by A^W.

We will use two transformations of cities. The first is the operation ν, defined earlier in this section, which scales each group so that its size is one. The second is a "flattening" transformation, ϕ, which divides each group g into T_g subgroups of size one, where each subgroup has the same distribution across neighborhoods. For example, suppose $X = \langle (1,2), (0,1) \rangle$. Since the first group in X already has size one, it is not subdivided. The second group has two members in the first neighborhood and one in the second, so it is divided into three equally distributed subgroups, each having 2/3 of a member in the first neighborhood and 1/3 in the second:

$$\phi(X) = \left\langle \left(1, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right), \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \right\rangle.$$

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.

[16] We can easily extend Theorem 3 to the class of cities where each group has a rational number of members. Indeed, by Weak Scale Invariance, any such city X is as segregated as the city cX where c is constant such that cT_g is an integer for each $g \in G(X)$. In order to further extend Theorem 3 to the class of all cities one would need a continuity axiom. We feel that there would be little to gain from doing this.

We now define a function ψ that transforms any ordering on $\mathcal{C}^I$ into an ordering on $\mathcal{C}^A$.

Definition 4 For any segregation ordering $\succcurlyeq$ on $\mathcal{C}^I$ define the associated ordering $\succcurlyeq^{\psi}$ on $\mathcal{C}^A$ as follows:

$$\text{for any } X, Y \in \mathcal{C}^A,\ X \succcurlyeq^{\psi} Y \text{ if and only if } \nu(X) \succcurlyeq \nu(Y). \tag{5}$$

Analogously, we define a function $-\psi$ that transforms any ordering on $\mathcal{C}^A$ into an ordering on $\mathcal{C}^I$.

Definition 5 For any segregation ordering $\succcurlyeq$ on $\mathcal{C}^A$, define the associated ordering $\succcurlyeq^{-\psi}$ on $\mathcal{C}^I$ as follows:

$$\text{for any } X, Y \in \mathcal{C}^I,\ X \succcurlyeq^{-\psi} Y \text{ if and only if } \phi(X) \succcurlyeq \phi(Y) \tag{6}$$

For any ordering $\succcurlyeq$ on $\mathcal{C}^I$, the associated ordering $\succcurlyeq^{\psi}$ satisfies SI, which is one of the axioms in Theorem 2. Claim 1 states a stronger property: if in addition, the ordering $\succcurlyeq$ satisfies all the axioms of Theorem 3, then the associated ordering $\succcurlyeq^{\psi}$ satisfies all the other axioms of Theorem 2 as well. That is, ψ is a function from $\mathcal{S}^I$ to $\mathcal{S}^A$:

Claim 1 *For any $\succcurlyeq$ in $\mathcal{S}^I$, the associated ordering $\succcurlyeq^{\psi}$ is in $\mathcal{S}^A$.*

Claim 2 states the converse: if the ordering $\succcurlyeq$ on $\mathcal{C}^A$ satisfies all the axioms of Theorem 2, then the associated ordering $\succcurlyeq^{-\psi}$ satisfies all the other axioms of Theorem 3. That is, $-\psi$ is a function from $\mathcal{S}^A$ to $\mathcal{S}^I$.

Claim 2 *For any $\succcurlyeq$ in $\mathcal{S}^A$, the associated ordering $\succcurlyeq^{-\psi}$ is in $\mathcal{S}^I$.*

The next step is to show that ψ is one-to-one.

Claim 3 *ψ is a one-to-one function from $\mathcal{S}^I$ to $\mathcal{S}^A$.*

Proof. Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two different orderings in $\mathcal{S}^I$. To show that ψ is one-to-one, we must show that $\succcurlyeq_1^{\psi}$ and $\succcurlyeq_2^{\psi}$ are different orderings. Since $\succcurlyeq_1$ and $\succcurlyeq_2$ differ, there are cities $X, Y \in \mathcal{C}^I$ for which $X \succcurlyeq_1 Y$ but $Y \succ_2 X$. As $\succcurlyeq_1$ and $\succcurlyeq_2$ satisfy GDP, we can assume without loss of generality that X and Y are normalized cities. (Otherwise, one can first apply ϕ to obtain normalized cities, without disrupting the cities' ranking under the two orderings.) But any ordering $\succcurlyeq$ in $\mathcal{S}^I$ agrees with its associated ordering $\succcurlyeq^\psi$ on the set of normalized cities. This implies that $X \succcurlyeq_1^\psi Y$ but $Y \succ_2^\psi X$; that is, $\succcurlyeq_1^\psi$ and $\succcurlyeq_2^\psi$ are different orderings. Q.E.D.

It follows that ψ is an isomorphism from $\mathcal{S}^I$ to $\mathcal{S}^A$ whose inverse is $-\psi$. To see why, note that $\mathcal{S}^A$ is a singleton by Theorem 2. Since $-\psi$ is a function from $\mathcal{S}^A$ to $\mathcal{S}^I$, the latter set must have at least one element. Since $\psi : \mathcal{S}^I \longrightarrow \mathcal{S}^A$ is one-to-one, $\mathcal{S}^I$ cannot have more than one element; else $\mathcal{S}^A$ would as well. So $\mathcal{S}^I$ is also a singleton. But then ψ must be an isomorphism from $\mathcal{S}^I$ to $\mathcal{S}^A$ whose inverse is $-\psi$.

Let $\succcurlyeq$ be the unique element of $\mathcal{S}^I$. It remains only to show that $\succcurlyeq$ is represented by A^W:

$$\begin{aligned}
X \succcurlyeq Y \quad &\Leftrightarrow \quad \phi(X) \succcurlyeq \phi(Y) && \text{as } \succcurlyeq \text{ satisfies GDP} \\
&\Leftrightarrow \quad \phi(X) \succcurlyeq^\psi \phi(Y) && \left\{ \begin{array}{l} \text{as } \succcurlyeq \text{ and } \succcurlyeq^\psi \text{ agree} \\ \text{on the set of normalized cities,} \\ \text{which includes } \phi(X) \text{ and } \phi(Y) \end{array} \right. \\
&\Leftrightarrow \quad A(\phi(X)) \geq A(\phi(Y)) && \text{as } \succcurlyeq^\psi \text{ is represented by } A \\
&\Leftrightarrow \quad A^W(X) \geq A^W(Y) && \text{as } A^W(\cdot) = A(\phi(\cdot)).\ ^{17}.
\end{aligned}$$

Q.E.D.

While any increasing transformation of the weighted Atkinson index also represents the same ordering, only the weighted Atkinson index also satisfies the Cardinalization Principle:

Proposition 4 *A^W is the unique segregation index on $\mathcal{C}^I$ whose induced segregation ordering satisfies GS, WSI, NDP, WIND, GDP, and DN, and that satisfies the Cardinalization Principle for some $K \geq 2$. Moreover, A^W satisfies the Cardinalization Principle for all $K \geq 1$.*

Proof. Identical to the proof of Proposition 3, with SI replaced by WSI. Q.E.D.

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

[17] For any city Z, $A(\phi(Z)) = 1 - \sum_{n \in N(Z)} \Big(\prod_{g \in G(Z)} \left(t_g^n\right)^{T_g(Z)} \Big)^{\frac{1}{T(Z)}} = A^W(Z)$

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

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 $f(t) = f(\underbrace{t, \ldots, t}_{M \text{ times}})$ for any $M \geq 1$. Then the segregation ordering on $\mathcal{C}^A$ that is represented by S 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 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:

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

Q.E.D.

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

Lemma 2 *For any $X \in \mathcal{C}_K$, let $S(X) = \sum_{n \in N(X)} f(t^n)$ 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.*

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. Part 2 is 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,

$$S(X) - S(X') = f(t^m) + f(t^n) - f(t^m + t^n) \tag{7}$$

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, $f(0,...,0) = 0$, so by (7), $X \sim X'$. Finally, if $p_g^m = p_g^n$ for all $g \in G_K$, 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 Theorem 2. We first show that IND is independent.

Claim 4 *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, and homogeneous of degree one. By Lemma 2, D satisfies GS, SI, and NDP. D satisfies WGDP by Lemma 1 as $f(t) = f(t,...,t)$. 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 verified that $D(X) = D(Y) = 1/2$. Consider now the result of annexing to them the one-neighborhood city $Z = \langle(4,0)\rangle$. One can verify that $D(X \uplus Z) = 3/4$ while $D(Y \uplus Z) = 1/2$. Hence, D violates IND. We leave it to the reader to check that D satisfies N. 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. We begin with two groups, b ("blacks") and w ("whites"). We define the set of groups to be the closure of $\{b, w\}$ under subdivision. For instance, this set could include the group "black females", but it does not include the group "females" since this group includes both blacks and whites. Let $\beta \in (0,1)$. For any group g, let the function $\varepsilon(g)$ equal β if g is a subgroup of b and $1-\beta$ if it is a subgroup of w. For any city X with set of groups G, let $\theta_G(g) = \varepsilon(g)\left[\sum_{g' \in G} \varepsilon(g')\right]^{-1}$. ($\theta_G$ 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:

$$A_\beta(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G(X)} \left(t_g^n\right)^{\theta_{G(X)}(g)} \right) \tag{8}$$

This index is like the 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 Atkinson index.

Claim 5 *The index A_β satisfies SI, NDP, IND, N, and WGDP on $\mathcal{C}^A$, but fails GS unless $\beta = 1/2$.*

We have shown that GS and IND are independent of the other axioms. It remains to show that NDP, N, WGDP, and SI are each independent. To show that NDP is independent, note that the index $1-A$ (defined by $(1-A)(X) = 1 - A(X)$) 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. For SI, 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. 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

$$
\begin{aligned}
S(X \uplus Z) \quad & \geq \quad S(Y \uplus Z) & \\
& \Leftrightarrow \quad T(X \uplus Z)A(X \uplus Z) \geq T(Y \uplus Z)A(Y \uplus Z) & \text{by definition of } S \\
& \Leftrightarrow \quad A(X \uplus Z) \geq A(Y \uplus Z) & \\
& \Leftrightarrow \quad A(X) \geq A(Y) & \text{as } A \text{ satisfies IND} \\
& \Leftrightarrow \quad T(X)A(X) \geq T(Y)A(Y) & \\
& \Leftrightarrow \quad S(X) \geq S(Y) & \text{by definition of } S
\end{aligned}
$$

so S satisfies IND. Finally, 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.

We have shown that the axioms are independent: each axiom has an index that violates it but that satisfies all other axioms. This is summarized in Table 11. A check mark indicates that the given index satisfies the given axiom; the absence of a check mark indicates that it does not.

6.2 Independence of Axioms in Theorem 3

We first show that WIND is independent. Consider the Weighted Dissimilarity Index defined in Section 3.1. It clearly satisfies GS, WSI, DN, and GDP. As for NDP, 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

	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	$\surd$	$\surd$	$\surd$	$\surd$		$\surd$
$\|G(X)\| * A(X)$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	

Table 11: Independence of the axioms of Theorem 2.

neighborhood is just $t^m + t^n$. Hence,

$$\begin{aligned} D^W(X) - D^W(X') &= \sum_{g \in G(X)} P_g \left|t_g^m - P^m\right| + \sum_{g \in G(X)} P_g \left|t_g^n - P^n\right| - \sum_{g \in G(X)} P_g \left|t_g^m + t_g^n - P^m - P^n\right| \\ &= \sum_{g \in G(X)} P_g \left(\left|t_g^m - P^m\right| + \left|t_g^n - P^n\right| - \left|t_g^m + t_g^n - P^m - P^n\right|\right) \end{aligned}$$

By the triangle inequality, $D^W(X) \geq D^W(X')$. If the group distributions in neighborhoods m and n are identical, then $t_g^m - P^m$ and $t_g^n - P^n$ have the same sign, so $D^W(X) = D^W(X')$. Hence, D^W satisfies NDP. As for WIND, consider the following cities: $X = \langle(1/2, 1), (1/2, 0)\rangle$ and $Y = \langle(1, 1/2), (0, 1/2)\rangle$. It can be checked that $D^W(X) = D^W(Y) = 1/2$. Consider now the result of combining each of them with the one-neighborhood city $Z = \langle(1, 0)\rangle$. It can be checked that $D^W(\nu(X \uplus Z)) = 3/4$ while $D^W(\nu(Y \uplus Z)) = 1/2$. Hence, D^W violates WIND while satisfying the other axioms: WIND is independent.

We now show that GS is independent, following an analogous argument in Section 6.1. We modify the construction of the asymmetric Atkinson index A_β (equation (8)) to produce a weighted version of this index. Let $\theta_G^W(g) = P_g \varepsilon(g) \left[\sum_{g' \in G(X)} P_{g'} \varepsilon(g')\right]^{-1}$, where ε is defined in Section 6.1. For any $\beta \in (0, 1)$, define:

$$A_\beta^W(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G(X)} \left(t_g^n\right)^{\theta_{G(X)}^W(g)} \right) \tag{9}$$

This index trivially satisfies DN and WSI. As for NDP, let X' be the result of splitting neighborhood n in city X into two neighborhoods, n_1 and n_2. Then:

$$A_\beta^W(X') - A_\beta^W(X) = \prod_{g \in G(X)} \left(t_g^n\right)^{\theta_{G(X)}^W(g)} - \prod_{g \in G(X)} \left(t_g^{n_1}\right)^{\theta_{G(X)}^W(g)} - \prod_{g \in G(X)} \left(t_g^{n_2}\right)^{\theta_{G(X)}^W(g)}$$

Since $\sum_{g\in G(X)} \theta^W_{G(X)}(g) = 1$, the products are concave functions of the vectors t^n, t^{n_1}, and t^{n_2}; since $t^n = t^{n_1} + t^{n_2}$, $A^W_\beta(X') - A^W_\beta(X) \geq 0$. If the neighborhood distributions in n_1 and n_2 are the same, then $t^{n_1}_g = \lambda t^n_g$ and $t^{n_2}_g = (1-\lambda)t^n_g$ for all g, where $\lambda = T^{n_1}/T^{n_2}$, so $A^W_\beta(X') - A^W_\beta(X) = 0$. Hence, A^W_β satisfies NDP. For WIND, let $X, Y \in \mathcal{C}^I$ be two normalized cities with the same set G of nonempty groups. Let $Z \in \mathcal{C}^A$ such that $G(Z) \subset G$. We wish to show that $A^W_\beta(X) \geq A^W_\beta(Y)$ if and only if $A^W_\beta(\nu(X \uplus Z)) \geq A^W_\beta(\nu(Y \uplus Z))$. But

$$\begin{aligned} A^W_\beta(X) &\geq A^W_\beta(Y) \Leftrightarrow A_\beta(X) \geq A_\beta(Y) \qquad \text{as } A_\beta \text{ agrees with } A^W_\beta \text{ on normalized cities} \\ &\Leftrightarrow A_\beta(X \uplus Z) \geq A_\beta(Y \uplus Z) \qquad \text{as } A_\beta \text{ satisfies IND} \\ &\Leftrightarrow A_\beta(\nu(X \uplus Z)) \geq A_\beta(\nu(Y \uplus Z)) \qquad \text{as } A_\beta \text{ satisfies SI} \\ &\Leftrightarrow A^W_\beta(\nu(X \uplus Z)) \geq A^W_\beta(\nu(Y \uplus Z)) \qquad \text{as } A_\beta \text{ agrees with } A^W_\beta \text{ on normalized cities} \end{aligned}$$

Regarding GDP, let $X \in \mathcal{C}$ and let $G = G(X)$ and $N = N(X)$. Let X' be the result of partitioning some group $g \in G$ into 2 groups, g_1 and g_2, each having the same distribution across neighborhoods as g itself. Then the only change in A^W_β is that for each n, the term $\left(t^n_g\right)^{\theta^W_{G(X)}(g)}$ in the product is replaced by $\left(t^n_{g_1}\right)^{\theta^W_{G(X)}(g_1)} \left(t^n_{g_2}\right)^{\theta^W_{G(X)}(g_2)}$. But $t^n_{g_1} = t^n_{g_2} = t^n_g$ and

$$\theta^W_{G(X)}(g_1) + \theta^W_{G(X)}(g_2) = \theta^W_{G(X)}(g)$$

so $A^W_\beta(X') = A^W_\beta(X)$: GDP holds. This proves that GS is independent.

Finally, analogous arguments to the ones used in Section 6.1 show that the indices $1 - A^W$, the trivial index, $|G(X)|A^W$, and $T(X)A^W$ satisfy all the axioms except NDP, DN, GDP, and WSI, respectively.

We have verified that the axioms of Theorem 3 are independent. This is summarized in Table 12.

7 Proofs

We first state and prove some preliminary lemmas.

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

1. *All cities in which every neighborhood is representative have the same degree of segregation under $\succcurlyeq$.*

	GS	WSI	NDP	WIND	DN	GDP
Weighted Atkinson: $A^W(X)$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
Asymmetric Weighted Atkinson: $A^W_\beta(X)$ for $\beta \neq 1/2$		$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$T(X) * A^W(X)$	$\surd$		$\surd$	$\surd$	$\surd$	$\surd$
$1 - A^W(X)$	$\surd$	$\surd$		$\surd$	$\surd$	$\surd$
Weighted Dissimilarity: $D^W(X)$	$\surd$	$\surd$	$\surd$		$\surd$	$\surd$
Trivial index	$\surd$	$\surd$	$\surd$	$\surd$		$\surd$
$\lvert G(X)\rvert * A^W(X)$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	

Table 12: Independence of the axioms of Theorem 3.

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 and GS, 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, SI, and GS. 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 and GS, 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, SI, and GS. Then*

1. $\overline{X}^K \succ \underline{X}^K$;
2. *for any $\alpha, \beta \in [0,1]$, $\alpha > \beta$,*

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

Proof.

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

2. By NDP,

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

and

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

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

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

The result follows by transitivity.

Q.E.D.

Proof of Theorem 1. Claim 5 implies that the Atkinson index A satisfies all the axioms of the theorem. We now show that it is the only index to do so.

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. Since $\succcurlyeq$ satisfies GS, it depends on the set of groups only via its cardinality. Hence, w.l.o.g., we can restrict attention to a fixed set of K groups.

Claim 6 *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 $X \sim \alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K$. Further, this unique α_X is $\left(\prod\limits_{g=1}^{K} t_g\right)^{1/K}$.*

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 $X \sim \overline{X}^K$. 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, $\sigma_{12}t = (t_2, 0, t_3, ..., t_K)$. Let $\mathbf{1}$ denote a vector of K ones. By GS,

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

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

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

In other words,

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

Hence, by NDP and GS,

$$\begin{aligned} (\gamma * t) * \underline{X}^K \uplus (\mathbf{1} - \gamma * t) * \overline{X}^K &\sim (\gamma * \sigma_{12}t) * \underline{X}^K \uplus (\mathbf{1} - \gamma * \sigma_{12}t) * \overline{X}^K \\ &\sim [\sigma_{12}(\gamma * \sigma_{12}t)] * \underline{X}^K \uplus (\mathbf{1} - [\sigma_{12}(\gamma * \sigma_{12}t)]) * \overline{X}. \end{aligned} \quad (10)$$

But note that since $(\gamma * t) = t$,

$$(\gamma * t) * \underline{X}^K \uplus (\mathbf{1} - \gamma * t) * \overline{X}^K = t * \underline{X}^K \uplus (\mathbf{1} - t) * \overline{X}^K = X.$$

Since $\sigma_{12}(\gamma * \sigma_{12} t) = \sigma_{12}\gamma * t = (0, \beta t_2, t_3, ..., t_K)$,

$$\sigma_{12}(\gamma * \sigma_{12} t) * \underline{X}^K \uplus (\mathbf{1} - [\sigma_{12}(\gamma * \sigma_{12} t)]) * \overline{X} \sim (0, \beta t_2, t_3, ..., t_K) * \underline{X}^K \uplus (\mathbf{1} - (0, \beta t_2, t_3, ..., t_K)) * \overline{X}^K.$$

Therefore, it follows from (10) that

$$X \sim (0, \beta t_2, t_3, ..., t_K) * \underline{X}^K \uplus (\mathbf{1} - (0, \beta t_2, t_3, ..., t_K)) * \overline{X}^K.$$

We can repeat this procedure for $t_3, ..., t_K$ to obtain

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

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

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

and

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

But by NDP,

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

and

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

so by IND,

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

and by SI,

$$t * \overline{X}^K \sim t * \underline{X}^K.$$

Now by NDP and IND,

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

so that $\alpha_X = 0$, as claimed.

Case 2. Suppose $t_g \in (0,1]$ for all g. Let $\alpha = \left(\prod_{g=1}^{K} t_g\right)^{1/K}$, and let

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

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,

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

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

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

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

Claim 7 *Let* $t^1, t^2 \in [0,1]^2$ *and let* $X = \langle t^1, t^2, (1-t_1^1-t_1^2, 0, ..., 0), (0, 1-t_2^1-t_2^2, 0, .., 0), ..., (0, ..., 0, 1-t_K^1-t_K^2) \rangle$ *be a city. Then there is a unique* $\alpha_X \in [0,1]$ *such that* $X \sim \alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K$. *Further, this unique* α_X *is* $\left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K}$.[18]

[18] This is less than or equal to 1 since the geometric average of a set of numbers can be no greater than their arithmetic average: $\left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K} \leq \frac{1}{K}\sum_{g=1}^{K} t_g^1 + \frac{1}{K}\sum_{g=1}^{K} t_g^2 = \frac{1}{K}\sum_{g=1}^{K} \left(t_g^1 + t_g^2\right) \leq \frac{1}{K}\sum_{g=1}^{K} 1 = 1.$

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 6. 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

$$X \sim Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle \tag{11}$$

where[19]

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

By Claim 6,

$$Y \sim \alpha_Y \underline{X}^K \uplus (1 - \alpha_Y) \overline{X}^K . \tag{12}$$

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

$$Y' = \tau \widetilde{t^2} * \underline{X}^K \uplus (\mathbf{1} - \tau \widetilde{t^2}) * \overline{X}^K . \tag{13}$$

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 $\tau \widetilde{t_g^2} \leq 1$. 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 6,

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

It follows from (12) and (14) that $Y \sim Y'$. As a result,

$$\begin{aligned} X &\sim Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by (11)} \\ &\sim Y' \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by IND} \\ &\sim \tau \widetilde{t^2} * \underline{X}^K \uplus (\mathbf{1} - \tau \widetilde{t^2}) * \overline{X}^K \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2}\right) \right\rangle && \text{by (13)} \\ &\sim (\tau + 1) \widetilde{t^2} * \underline{X}^K \uplus (\mathbf{1} - \tau \widetilde{t^2}) * \overline{X}^K && \text{by NDP} \\ &\sim (\tau + 1) t^2 * \underline{X}^K \uplus (\mathbf{1} - (\tau + 1) t^2) * \overline{X} && \text{by SI and definition of } \widetilde{t^2}. \end{aligned}$$

---

[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.

Therefore, using Claim 6, $X \sim \alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K$, 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

$$\widehat{X} \sim \widehat{\alpha}_X \underline{X}^K \uplus (1-\widehat{\alpha}_X)\overline{X}^K. \tag{15}$$

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

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

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

$$X \sim \left(2\widehat{\alpha}_X\right)\underline{X}^K \uplus (1-\left(2\widehat{\alpha}_X\right))\overline{X}^K$$

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

Lemma 6 *For every city X there is a unique $\alpha_X \in [0,1]$ such that $X \sim \alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K$. Further, this unique α_X is $\sum_{n\in N(X)}\left(\prod\limits_{g=1}^{K} t_g^n\right)^{1/K}$.*[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 (6) and (7) already

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

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

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

By the induction hypothesis, $\left(\frac{1}{1-t_1^m}, ..., \frac{1}{1-t_K^m}\right) * Y \sim \alpha_Y \underline{X}^K \uplus (1-\alpha_Y) \overline{X}^K$ 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 7,

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

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 $X \sim \alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K$. Hence, $S(X) = S(\alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K)$, which equals α_X by the Cardinalization Principle. But $A(X)$ also equals $A(\alpha_X \underline{X}^K \uplus (1-\alpha_X)\overline{X}^K)$, which equals α_X by the Cardinalization Principle. Hence, $S(X) = A(X)$. Q.E.D.

Proof of Proposition 3. 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:

$$\underline{X}^K \sim \underline{X}^{K'} \text{ for all } K, K' \geq 1 \tag{16}$$

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

$$\overline{X}^K \sim \overline{X}^{K'} \text{ for all } K, K' \geq 2 \tag{17}$$

Let A' be any increasing transformation of A that differs from A and let X be a city for which $A(X) \neq A'(X)$. There are two cases.

1. If X has one group, then $X \sim \underline{X}^1$ by NDP and SI. Hence, $A(X) = 0$, so $A'(X) \neq 0$. But $A'(X) = A'(\underline{X}^1)$, so $A'(\underline{X}^1)$ is not equal to zero. By (16), for any $K \geq 2$, $\underline{X}^K \sim \underline{X}^1$. Hence, $A'(\underline{X}^K) \neq 0$, so A' violates the Cardinalization Principle for K groups (setting $\alpha = 0$ in condition 2 and using NDP).

2. If X has $K' \geq 2$ groups, then by Lemma 6 there is an $\alpha \in [0,1]$ such that $X \sim \alpha\overline{X}^{K'} \uplus (1-\alpha)\underline{X}^{K'}$. By (16) and (17), $X \sim \alpha\overline{X}^K \uplus (1-\alpha)\underline{X}^K$ 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.

Q.E.D.

Our proofs of Claims 1 and 2 make use of several simple properties of the transformations ϕ and ν, which we state without proof.

Lemma 7 *1. 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]

2. For any city $X \in \mathcal{C}$, and for any relabeling σ, $\nu(\sigma(X)) = \sigma(\nu(X))$.

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

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

Proof of Claim 1. We must show that, for any $\succcurlyeq$ defined on $\mathcal{C}^I$ that satisfies GS, WSI, NDP, GDP, WIND, and DN, the associated ordering $\succcurlyeq^\psi$on $\mathcal{C}^A$ satisfies GS, SI, NDP, WGDP, IND, and N.

GS: Let X be a city with the set G of nonempty groups and let σ be a relabeling. We must verify that $X \sim^\psi \sigma(X)$. This holds if and only if $\nu(X) \sim \nu(\sigma(X))$. But $\nu(\sigma(X)) = \sigma(\nu(X))$ by Lemma 7, part 2. 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

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

$$\phi(X) \sim \alpha\phi(X) \sim \phi(\alpha\phi(X)) = \phi(\alpha X).$$

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

$$X \succcurlyeq^{\psi} Y \Longleftrightarrow X \uplus Z \succcurlyeq^{\psi} Y \uplus Z. \tag{18}$$

By definition of $\succcurlyeq^{\psi}$, equation (18) is equivalent to

$$\nu(X) \succcurlyeq \nu(Y) \Longleftrightarrow \nu(X \uplus Z) \succcurlyeq \nu(Y \uplus Z). \tag{19}$$

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 (19) it suffices to prove that

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

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 3 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. Q.E.D.

Proof of Claim 2. We must show that, for any $\succcurlyeq$ defined on $\mathcal{C}^A$ that satisfies GS, SI, NDP, WGDP, IND, and N, the associated ordering $\succcurlyeq^{-\psi}$on $\mathcal{C}^I$ satisfies GS, WSI, NDP, GDP, WIND, and DN.

GS: for any city $X \in \mathcal{C}^I$ and any relabeling σ, we must show that $X \sim^{-\psi} \sigma(X)$. By (6), this holds if and only if $\phi\left(X\right) \sim \phi\left(\sigma(X)\right)$. $\phi\left(\sigma(X)\right)$ is the result of flattening the city X after its groups have been relabeled by σ. This is equivalent to relabeling the groups of the flattened city $\phi(X)$: there is a relabeling $\widehat{\sigma}$ of the groups of $\phi\left(X\right)$ such that $\widehat{\sigma}\left(\phi\left(X\right)\right) = \phi(\sigma(X))$. Since $\succcurlyeq$ satisfies GS, $\phi\left(X\right) \sim \widehat{\sigma}\left(\phi\left(X\right)\right) = \phi(\sigma(X))$.

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 (6), this holds if and only if $\phi\left(X\right) \sim \phi\left(\alpha X\right)$, which follows from part 1 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 (6)); otherwise, $\phi(X') \succcurlyeq \phi(X)$ (and hence $X' \succcurlyeq^{-\psi} X$ by (6)). 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

$$X \succcurlyeq^{-\psi} Y \text{ if and only if } \nu(X \uplus Z) \succcurlyeq^{-\psi} \nu(Y \uplus Z) \tag{21}$$

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

$$\begin{array}{rcll} \phi(X) \succcurlyeq \phi(Y) & \Leftrightarrow & X \succcurlyeq Y & \text{since } X \text{ and } Y \text{ are already flat} \\ & \Leftrightarrow & X \uplus Z \succcurlyeq Y \uplus Z & \text{by IND of } \succcurlyeq \\ & \Leftrightarrow & \nu(X \uplus Z) \succcurlyeq \nu(Y \uplus Z) & \text{by SI of } \succcurlyeq \\ & \Leftrightarrow & \phi(\nu(X \uplus Z)) \succcurlyeq \phi(\nu(Y \uplus Z)) & \text{by Lemma 7, part 4.} \end{array}$$

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 4 of Lemma 7, $\phi(\nu(X)) \succ \phi(\nu(Y))$, so $\nu(X) \succ^{-\psi} \nu(Y)$. We have produced two cities each with K residents that $\succ^{-\psi}$ ranks differently. Q.E.D.

Proof of Claim 5. If $\beta \neq 1/2$, the index fails GS since it is not invariant to permutations that do not preserve the groups' labels. The index can be rewritten

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

Note that f is convex and homogeneous of degree one. By Lemma 2, A_β satisfies SI and NDP. We leave it to the reader to check that A_β satisfies 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,

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

This shows that A_β 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.

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

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

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

[7] Farley, Reynolds. 1977. "Residential Segregation in Urbanized Areas of the United States in 1970: An Analysis of Social Class and Racial Differences." *Demography* 14:497-518.

[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." *International Economic Review* 45:555–578.

[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] Philipson, Tomas. 1993. "Social Welfare and Measurement of Segregation." *Journal of Economic Theory* 60:322-334.

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

[21] Rhode, Paul W., and Koleman S. Strumpf. 2003. "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990." *American Economic Review* 93:1648-1677.

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

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

[24] Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures." *Econometrica* 39:613-25.

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