Scale-Invariant Measures of Segregation[*]

David M. Frankel

Iowa State University

Oscar Volij

Ben Gurion University

August 7, 2008

Abstract

We characterize measures of school segregation for any number of ethnic groups using a set of purely ordinal axioms that includes Scale Invariance: a school district's segregation ranking should be invariant to changes that do not affect the distribution of ethnic groups across schools. The symmetric Atkinson index is the unique such measure that treats ethnic groups symmetrically and that ranks a district as weakly more segregated if either (a) one of its schools is subdivided or (b) its students in a subarea are moved around so as to weakly raise segregation in that subarea. If the requirement of symmetry is dropped, one obtains the general Atkinson index. The role of Scale Invariance is illustrated by studying segregation among U.S. public schools from 1987/8 to 2005/6, a period in which ethnic groups became distributed more similarly across schools. While the Atkinson indices declined sharply, most other indices either rose or declined only slightly.

1 Introduction

Recent research supports the view that school segregation creates unequal opportunities: separate schools are not equal. By and large, students in schools with a higher proportion of minority students have lower educational attainment and subsequent wages (Boozer, Krueger, and Wolkin [4, pp. 303-6]; Hanushek, Kain, and Rivkin [16]; Hoxby [17]). Consistent with this, African Americans tend to have better educational outcomes in less segregated school districts (Card and Rothstein [5]; Guryan [15]).[1]

While the effects of school segregation are of great concern, there also rages, behind the scenes, an intense controversy. It is over a more fundamental issue: how segregation itself should be measured. One view of segregation, which we take in this paper, is that it relates to how *origins* affect *destinations*: how a student's ethnic origin affects which school she ends up attending. To measure segregation, we must therefore look at the extent to which ethnic groups are distributed differently across schools. This notion is favored by the sociologists James and Taeuber [22]. It corresponds to the first of Massey and Denton's [28] five dimensions of segregation, which they call "evenness".[2]

In order to be true to the notion of evenness, we require that a measure be Scale Invariant: that it rank a district based solely on how each ethnic group is distributed across the district's schools, and not on the proportions of different ethnic groups in the district. In explaining this property, Taeuber and James write:

> School segregation refers to racial variation in the distribution of students across schools. ... This concept of segregation does not depend on the relative proportions of blacks and whites in the [district], but only upon the relative distributions of students among schools.... [Taeuber and James [38, p. 134]]

---

[1]Cutler and Glaeser [10] find similar effects of residential segregation.

[2]Their second dimension is isolation of the minority group. The other three dimensions, which are more relevant in a residential context, are concentration in a small area, centralization in the urban core, and clustering in a contiguous enclave.

Scale Invariance is one of the five requirements that Jahn *et al* [21] say a satisfactory measure of segregation should satisfy.[3]

One important context in which Scale Invariance is desirable is school desegregation. How should a district's progress in desegregation be judged? To be fair, we should not penalize a district for factors that are out of its control, such as the size or growth rate of its minority population. Accordingly, a Scale Invariant measure of segregation is needed. This is not a trivial requirement: none of the commonly used segregation indices are Scale Invariant in the general multigroup case.[4]

Formally, we define a segregation ordering as a complete ordering on school districts: a ranking of districts from most segregated to least segregated. In addition to Scale Invariance, we require this ordering to satisfy a few other simple axioms. Group Symmetry requires that the segregation ordering be invariant to the renaming of the groups. The Weak School Division Property states that in a school district that contains a single school, building a new school to which some of the students are moved (a) cannot lower segregation in the district, and (b) leaves segregation unchanged if the ethnic distributions of the two resulting schools are identical. Finally, our Independence axiom states that if the students in a subset of schools are reallocated within that subset, then segregation in the whole district rises if and only if it rises within the subset.

These are all ordinal axioms: rules for how various operations should affect a district's place in the segregation ordering. In contrast, the prior literature has generally imposed cardinal axioms on the segregation index itself. Sometimes these properties have clear ordinal implications, but often they do not. For example, what are the ordinal implications of requiring that an index be additively separable? This problem is avoided by restricting

---

[3]They write: "a satisfactory measure of ecological segregation should ... not be distorted by the size of the total population, the proportion of Negroes, or the area of a city...." (Jahn *et al* [21]).

[4]The Gini index and the usual formulation of the Dissimilarity index are Scale Invariant only in the two-group case. The Entropy and Normalized Exposure indices are never Scale Invariant. See, e.g., the survey in Reardon and Firebaugh [33].

to ordinal axioms.

Our main results are as follows. First, the ordering that is captured by the symmetric Atkinson index is the unique nontrivial ordering that satisfies Group Symmetry, Scale Invariance, the Weak School Division Property, and Independence. This index is defined as one minus the sum, over all schools, of the geometric averages of the percentages of each group who attend the school.[5] For instance, suppose 40% of blacks, 70% of Hispanics, and 10% of whites attend school A while the remainder attend school B. The index equals $1 - (.4)^{1/3} (.7)^{1/3} (.1)^{1/3} - (.6)^{1/3} (.3)^{1/3} (.9)^{1/3}$.

We then drop the Group Symmetry axiom and characterize the set of orderings that satisfy the remaining axioms, with the addition of a technical Continuity axiom. Each such ordering is represented by an index of the following form: one minus the sum, over all the schools, of some *weighted* geometric average of the percentages of each group who attend the school. In the above example this would equal $1 - (.4)^b (.7)^h (.1)^w - (.6)^b (.3)^h (.9)^w$ where b, h, and w are arbitrary nonnegative weights that sum to one. We will refer to this as the general Atkinson index.

While the symmetric Atkinson index gives equal weight to all ethnic groups, the general Atkinson index leaves these weights up to the researcher. For instance, one may desire an index that is more sensitive to segregation between two large groups than two small groups. Most multigroup segregation indices accomplish this by weighting a group according to the group's relative size in a given district. This violates Scale Invariance since a group's weight varies from one district to another. An alternative is to use an Atkinson index with weights that vary across groups but not across districts. For instance, a group's weight might equal the proportion of students in the U.S. who belong to that group in some reference year. Since the weight on each group is constant across districts, Scale Invariance is satisfied.

_____

[5] This is an increasing transformation of the original symmetric Atkinson index of James and Taeuber [22] in the case of two ethnic groups. Hutchens [19] calls this version the Square Root Index. Since they are related by an increasing transformation, the two versions represent the same ordering of school districts (section 3).

We illustrate our results by studying changes in U.S. public school segregation from 1987/8 to 2005/6. We study the symmetric Atkinson index as well as two asymmetric Atkinson indices that are based, respectively, on the ethnic distributions at the beginning and end of the period. We first show, using Lorenz curves, that ethnic groups had more similar distributions across public schools in 2005/6 than in 1987/8.[6] By the evenness criterion, school segregation should have fallen. Indeed, the Atkinson indices fell considerably over the period. However, there was also another development: the proportion of whites in the student population fell from 71% to 58%. This increase in ethnic diversity has no effect on the Atkinson indices as they are Scale Invariant. However, it tempered or even reversed the declines of the multigroup versions of indices such as Dissimilarity, Gini, Entropy, and Normalized Exposure, none of which are Scale Invariant in the multigroup case.[7] As a result, none of these indices fell as much as the Atkinson indices, and some of them rose considerably.

The Atkinson segregation indices were introduced by James and Taeuber [22] and are based on the Atkinson family of inequality indices (Atkinson [2]). Massey and Denton [28] study properties of the Atkinson indices; Johnston, Poulsen, and Forrest [24] use them to study residential segregation. While this literature has focused on the case of two ethnic groups, we study the general multigroup case.

The first to study segregation axiomatically, Philipson [31], 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 a school's demographic distribution only. Hutchens [18] characterizes the family of segregation indices that satisfy a set of basic properties in the case of two ethnic groups. In a subsequent paper, Hutchens [19] strengthens one axiom and obtains the symmetric Atkinson

---

[6] There is just one exception: the Lorenz curve for blacks vs. whites for the two years intersect. However, they are very close to one another.

[7] There is a Scale Invariant version of the Dissimilarity index that does decline over the period (section 3.2). However, this version is not in common use.

index. While we assume properties of the underlying segregation ordering, Hutchens follows the inequality literature (e.g., Shorrocks [36, 37]) by imposing restrictions directly on the segregation index. Echenique and Fryer [13] characterize an index that uses social network data to measure the strength of an individual's isolation from members of other demographic groups. They also rely on cardinal axioms.

None of our axioms relate to comparisons between districts that have different numbers of ethnic groups. Thus, we can say nothing about how such districts should be ranked. Frankel and Volij [14] allow a variable number of ethnic groups by assuming the Group Division Property: if a given ethnic group is subdivided into two groups that have the same distribution across schools, then the segregation of the district should not change. However, they drop Scale Invariance as the two axioms together have implausible implications.[8]

The paper is organized as follows. Concepts and notation are defined in Section 2. Section 3 gives examples of segregation indices. Section 4 presents the axioms. Theoretical results appear in section 5 and empirical findings in section 6. We conclude in section 7. Proofs are relegated to an appendix.

2   Definitions

We assume a continuum population. This is a reasonable approximation when ethnic groups are large. In our examples, each "person" should be interpreted as representing some large, fixed number of students.

Formally, we define a (school) district as follows:

Definition 1 A *district X* consists of

- A nonempty and finite set of schools $\mathbf{N}$ and ethnic groups $\mathbf{G}$. (We will write $\mathbf{N}(X)$ when necessary to avoid ambiguity.)

---

[8] Together, the axioms permit one to scale a group up and then split it into identically distributed groups - thus making endless copies of the original group - without changing a district's segregation ranking.

- For each ethnic group $g \in \mathbf{G}$ and for each school $n \in \mathbf{N}$, a real number $T_g^n \geq 0$: the number of members of ethnic group g that attend school n.

We will sometimes specify a district as a list of ethnic compositions of schools in the district. For instance, $\langle (10, 20), (30, 10) \rangle$ denotes a district with two schools and two ethnic groups - say, blacks and whites. The first school, $(10, 20)$, contains ten blacks and twenty whites; the second, $(30, 10)$, contains thirty blacks and ten whites.

For any nonnegative scalar α, αX denotes the district in which the number of students in each group and school has been multiplied by α. If Y is another district, $X \uplus Y$ denotes the result of combining districts X and Y. For example, if $X = \langle (10, 20), (30, 10) \rangle$ and $Y = \langle (40, 50) \rangle$, then $2X = \langle (20, 40), (60, 20) \rangle$, and $X \uplus Y = \langle (10, 20), (30, 10), (40, 50) \rangle$.

The following notation will be useful:

$T_g = \sum_{n \in \mathbf{N}} T_g^n$: the number of students in ethnic group g in the district

$T^n = \sum_{g \in \mathbf{G}} T_g^n$: the total number of students who attend school n

$T = \sum_{g \in \mathbf{G}} T_g$: the total number of students in the district

$P_g = \dfrac{T_g}{T}$: the proportion of students in the district who are in ethnic group g

$P^n = \dfrac{T^n}{T}$: the proportion of students in the district who are in school n

$p_g^n = \dfrac{T_g^n}{T^n}$ (for $T^n > 0$): the proportion of students in school n who are in ethnic group g

$t_g^n = \dfrac{T_g^n}{T_g}$: the proportion of students in ethnic group g who attend school n

The *ethnic distribution of a district* X is the vector $P = (P_g)_{g \in \mathbf{G}}$ of proportions of the students in the district who are in each ethnic group. The *ethnic distribution of a nonempty school* n is the vector $p^n = (p_g^n)_{g \in \mathbf{G}}$ of proportions of students in school n who are in each ethnic group. A school is *representative* if it has the same ethnic distribution as the district that contains it.

3 Examples of Segregation Indices

3.1 Atkinson Indices

The Atkinson segregation indices were introduced by James and Taeuber [22] for the case of two ethnic groups. They are based on the Atkinson inequality indices (Atkinson [2]).[9] Let $\mathbf{w} = (w_1 \ldots w_K)$ be a vector of K nonnegative weights that sum to one. The general Atkinson index with weights $\mathbf{w}$, $A_{\mathbf{w}}$, is defined by

$$A_{\mathbf{w}}(X) = 1 - \sum_{n \in \mathbf{N}(X)} \prod_{g \in \mathbf{G}} \left(t_g^n\right)^{w_g} \tag{1}$$

The symmetric Atkinson index is obtained when all the weights are equal:

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

3.2 Unweighted Dissimilarity

In the case of two groups, the Index of Dissimilarity (Jahn *et al* [21]) equals the proportion of either group who would have to change schools in order to attain complete integration. This index was used by Cutler, Glaeser, and Vigdor [11] to measure the evolution of segregation in American cities. Its usual generalization to three or more groups gives more weight to larger groups and is due to Morgan [30] and Sakoda [35] (see section 6). Another possible generalization, which we call the Unweighted Dissimilarity index, is defined by

$$D^U(X) = \frac{1}{2(K-1)} \sum_{n \in \mathbf{N}(X)} f(t^n) \text{ where } f(t^n) = \sum_{g \in \mathbf{G}} \left| t_g^n - \sum_{g' \in \mathbf{G}} \frac{1}{K} t_{g'}^n \right| \tag{3}$$

This generalization gives the same weight to each ethnic group.

---

[9]In the case of two groups, the Atkinson index with weight $0 < \delta < 1$ on group one equals $1 - \left[\sum_{n \in N(X)} \left(t_1^n\right)^{\delta} \left(t_2^n\right)^{1-\delta}\right]^{\frac{1}{1-\delta}}$ (James and Taeuber [22, p. 9]). This index is difficult to generalize to more than two groups since the outer exponent, $\frac{1}{1-\delta}$, is the reciprocal of the weight on a particular ethnic group. Instead, we generalize $1 - \sum_{n \in N(X)} \left(t_1^n\right)^{\delta} \left(t_2^n\right)^{1-\delta}$. This is an increasing transformation of the original index and thus represents the same ordering.

3.3 Mutual Information

The *entropy* of the discrete probability distribution $q = (q_1, \ldots, q_K)$ is defined by[10]

$$h(q) = \sum_{k=1}^{K} q_k \log_2 \left(\frac{1}{q_k} \right).$$

The Mutual Information Index equals the entropy of a district's ethnic distribution minus the average entropy of the ethnic distributions of its schools:

$$M(X) = h(P) - \sum_{n \in \mathbf{N}(X)} P^n h(p^n)$$

where $P = (P_g)_{g \in \mathbf{G}}$ is the district ethnic distribution and $p^n = (p_g^n)_{g \in \mathbf{G}}$ is the ethnic distribution of school n. This index was first proposed by Theil [39] and is axiomatized in Frankel and Volij [14].

4 Axioms

We now introduce our axioms. We restrict attention to districts that have a fixed set of ethnic groups with positive membership. More precisely, let $\mathbf{G}$ be a set of $K \geq 2$ ethnic groups and let $\mathcal{C}$ be the set of all districts whose nonempty ethnic groups are exactly the groups in $\mathbf{G}$. A *segregation ordering* $\succcurlyeq$ *on* $\mathcal{C}$ is a complete and transitive binary relation on $\mathcal{C}$. The statement $X \succcurlyeq Y$ means "district X is at least as segregated as district Y." The relations $\sim$ and $\succ$ are derived from $\succcurlyeq$ in the usual way.[11]

A related concept is the segregation *index*: a function $S : \mathcal{C} \to \Re$ that assigns to each district a number that is interpreted as the district's segregation level. The index S *represents* the segregation ordering $\succcurlyeq$ if, for any two districts $X, Y \in \mathcal{C}$,

$$X \succcurlyeq Y \text{ if and only if } S(X) \geq S(Y) \tag{4}$$

Every index S induces a segregation ordering $\succcurlyeq$ that is defined by (4).

---

[10]When $q_k = 0$, the term $q_k \log_2(1/q_k)$ is assigned the value zero.

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

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.

A district's *segregation ranking* or simply its *segregation* is its place in the segregation ordering. We will sometimes say that if a transformation $\sigma : \mathcal{C} \to \mathcal{C}$ is applied to a district X, then "the segregation of the district is unchanged" or "the district's segregation ranking is unaffected." By this we mean that $\sigma(X) \sim X$. If this holds for all districts X, then we will say that the segregation in a district is invariant to the transformation σ.

Our first axiom, Scale Invariance, requires that an ordering be insensitive to changes in the size of an ethnic group that leave that group's distribution across schools unchanged. More precisely:

Scale Invariance (SI) For any district $X \in \mathcal{C}$, group $g \in \mathbf{G}$, and constant $\alpha > 0$, let X' be the result of multiplying the number of group-g students in each school n in district X by α. Then $X' \sim X$.

The axiom of Independence states that if the students in a subarea of a district are reallocated among schools within that subarea, then segregation in the district rises if and only if segregation in the subarea rises:

Independence (IND) Let $X, Y \in \mathcal{C}$ have equal populations and equal group distributions. Then for any $Z \in \mathcal{C}$, $X \uplus Z \succsim Y \uplus Z$ if and only if $X \succsim Y$.

Since X and Y have the same number of each ethnic group, $Y \uplus Z$ is the result of reallocating the students within the subarea X of the district $X \uplus Z$.[12] The Dissimilarity Index violates this principle. For instance, suppose a district is composed of two areas: $X = \langle (50, 100), (50, 0) \rangle$ and $Z = \langle (100, 0) \rangle$. Suppose that the students in the first area are

---

[12]IND does not require Y to have the same number of schools as X. Hence, the reallocation might be accompanied by new school construction or conversion of some schools to other uses.

reallocated to yield $Y = \langle (100, 40), (0, 60) \rangle$. The Dissimilarity Index within this area *rises* from 0.5 to 0.6, but the index for the full district *falls*, counterintuitively, from 0.75 to 0.6.[13] Independence rules out this behavior. In Section 5.2 we show that Independence is also a precondition for an index to be additively decomposable in a sense discussed by Hutchens [18].

The next axiom is the Weak School Division Property. This axiom states that a one-school district cannot become *less* segregated if the school is split into two new schools. In addition, if the new schools have identical ethnic distributions, then segregation is unchanged. Intuitively, since a district that contains a single school is not segregated at all, splitting the school cannot lead to lower segregation.[14] And if the new schools have the same ethnic distribution, then the new district is not segregated at all, like the original district.

Weak School Division Property (WSDP) Let $X \in \mathcal{C}$ be a district consisting of a single
school. Let X' be the district that results from subdividing this school into two schools,
n_1 and n_2. Then, $X' \succcurlyeq X$. Further, if n_1 and n_2 have the same group distributions
(i.e., $p_g^{n_1} = p_g^{n_2}$ for all $g \in \mathbf{G}$), then $X' \sim X$.

This axiom implies, for instance, that a district with 110 whites and ten blacks in a single school does not become more segregated if the ten blacks and an equal number of whites are relocated to a second school: the district $\langle (110, 10) \rangle$ is no more segregated than the district $\langle (100, 0), (10, 10) \rangle$. Of course, one can think of notions of "segregation" that would contradict this. A student in the school $(10, 10)$ might think that her new environment is more "integrated" since it has equal numbers of blacks and whites. No model can capture all possible notions of segregation. By Massey and Denton's [28] evenness criterion, segregation

---

[13]The two versions of the Dissimilarity index coincide in this example since there are only two ethnic groups. They equal the percentage of either group that must change schools in order for all schools to have the same ethnic distribution.

[14]Our motivating example uses schools as the basic locational unit, so it ignores ability tracking and other forms of within-school segregation. Our approach could easily be used to study these phenomena by redefining basic locational unit to be the classroom or the ability group.

in the district has indeed increased: ethnic groups are (trivially) distributed evenly across schools in the first district but not in the second.

WSDP is related to two properties that are discussed by James and Taeuber [22] and subsequent authors. The first is *organizational equivalence*: if a school is divided into two schools that have the same ethnic distribution, the district'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) student moves from one school to another school in which the proportion of blacks (whites) is higher, then segregation in the district rises. In the case of two ethnic groups, WSDP follows from organizational equivalence and the transfer principle.[15] But while WSDP applies directly with any number of groups, it is unclear what form the transfer principle should take with more than two groups.[16]

Group Symmetry states that the level of segregation in a district does not depend on the labeling of the district's demographic groups; it depends only on the number of people in each group who attend each school. For instance, if "blacks" are relabeled "whites" and vice-versa, then segregation does not change.

Group Symmetry (GS) The segregation in a district is invariant to any relabeling or reordering of the groups in the district.

We will consider axiomatizations both with and without this axiom.

Our next axiom, Continuity, will be needed only when Group Symmetry is dropped.

---

[15] A rough intuition runs as follows. The first part of WSDP is just organizational equivalence itself. As for the second part, any division of a school into two new schools with differing ethnic distributions can be broken into two steps. First, create two schools with the desired sizes but the *same* ethnic distribution. By organizational equivalence, segregation is unchanged. In the second step, swap black students with white students until the desired ethnic distributions are attained. Since each swap moves students to schools in which their groups are overrepresented, segregation must rise by the transfer principle.

[16] For instance, suppose a black student moves to a school that has higher proportions of both blacks and Asians but fewer whites. Since there are more blacks, one might argue (using the transfer principle) that segregation has gone up. On the other hand, blacks are now more integrated with Asians. One attempt to overcome this difficulty appears in Reardon and Firebaugh [33].

Continuity (C) For any districts $X, Y, Z \in \mathcal{C}$, the sets

$$\{c \in [0,1] : cX \uplus (1-c)Y \succsim Z\} \text{ and } \{c \in [0,1] : Z \succsim cX \uplus (1-c)Y\}$$

are closed.

Our final axiom states that there exist two districts, one strictly more segregated than the other. It is needed to rule out the trivial segregation ordering.

Nontriviality (N) There exist districts $X, Y \in \mathcal{C}$ such that $X \succ Y$.

5 Results

We first show that the general Atkinson index (equation (1)) satisfies all of the axioms other than Group Symmetry. This axiom is satisfied only by the symmetric Atkinson index (equation (2)).

Proposition 1 *Let* $\mathbf{w} = (w_1, \ldots, w_K)$ *be a list of* K *non-negative weights that add up to one. The ordering represented by the general Atkinson index* $A_{\mathbf{w}}$ *satisfies SI, IND, WSDP, N, and C. The ordering represented by the symmetric Atkinson index* A *also satisfies GS.*

The next two theorems are the main results of our paper. Theorem 1 states that our set of axioms, less Group Symmetry, fully characterizes the general Atkinson index.

Theorem 1 *Let* $\succsim$ *be an ordering on* $\mathcal{C}$ *that satisfies SI, IND, WSDP, N, and C. There are fixed weights* $w_g \geq 0$ *for* $g = 1, ..., K$, *adding up to one, such that* $\succsim$ *is represented by the general Atkinson index* $A_{\mathbf{w}}(X)$.

An easy implication of Theorem 1 is that if the requirement of Group Symmetry is added, then the weights w_g must all be equal. Hence, the symmetric Atkinson index represents the unique ordering that satisfies this larger set of axioms. It turns out that this is still true if Continuity is dropped. This is the following result.

Theorem 2 *The ordering represented by the symmetric Atkinson index on* $\mathcal{C}$ *is the only ordering that satisfies GS, SI, WSDP, IND, and N.*

5.1 Independence of the Axioms

Are the axioms in Theorems 1 and 2 independent of each other? In this section, we show that they are: for each of the axioms in each of the two theorems, there is an index that violates it yet that satisfies the other axioms. Consequently, all of the axioms are needed for our results to hold.

For any two different vectors $\mathbf{w}$ and $\mathbf{w}'$ of group weights (each summing to one), consider the following lexicographic ordering:

$$X \succcurlyeq_{\mathbf{w},\mathbf{w}'} Y \text{ iff } \begin{cases} A_{\mathbf{w}}(X) > A_{\mathbf{w}}(Y) \\ \text{or} \\ A_{\mathbf{w}}(X) = A_{\mathbf{w}}(Y) \text{ and } A_{\mathbf{w}'}(X) \geq A_{\mathbf{w}'}(Y) \end{cases}$$

This ordering first uses the general Atkinson index with weights $\mathbf{w}$ to rank districts. Any "ties" are broken using the general Atkinson index with weights $\mathbf{w}'$. The following proposition uses this index and the other indices defined in Section 3 to show that our axioms are independent of each other.

Proposition 2 *The axioms SI, WSDP, IND, N, and C are independent of each other, as are the axioms GS, SI, WSDP, IND, and N. In particular:*

- *The symmetric Atkinson index $A(X)$ satisfies all the axioms;*

- *any general Atkinson index with unequal weights satisfies all axioms but Group Symmetry;*

- *the Mutual Information index satisfies all axioms but Scale Invariance;*

- *$1 - A(X)$ satisfies all axioms but the Weak School Division Property;*

- *the Unweighted Dissimilarity index satisfies all axioms but Independence;*

- *the trivial index, which ranks all districts as equally segregated, satisfies all axioms but Nontriviality;*

14

- *the lexicographic index $\succcurlyeq_{\mathbf{w},\mathbf{w}'}$, for weights $\mathbf{w} \neq \mathbf{w}'$, satisfies all axioms but Group Symmetry and Continuity (so C is independent of SI, WSDP, IND, N).*

This proposition is summarized in Table 1. A check mark indicates that an index satisfies a given axiom; an $\times$ indicates that it does not.

	GS	SI	WSDP	IND	N	C
Symmetric Atkinson: $A(X)$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
$A_{\mathbf{w}}(X)$ for $\mathbf{w} \neq (1/K, \ldots 1/K)$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Mutual Information $M(X)$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
$1 - A(X)$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
Unweighted Dissimilarity: $D^U(X)$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$
Trivial index	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$
Lexicographic $\succcurlyeq_{\mathbf{w},\mathbf{w}'}$ for $\mathbf{w} \neq \mathbf{w}'$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$

Table 1: Independence of the axioms.

5.2 Additive Decomposability

It is often necessary to study segregation at several levels simultaneously. For instance, one may be interested in how much of the segregation between classrooms in a district is due to residential segregation and how much is due to ability tracking within schools. As a first approximation, one might want to decompose total segregation into between-school and within-school, between-classroom segregation. It turns out that only indices that satisfy Independence can be decomposed in this way. This includes the Atkinson indices but not, e.g., the Unweighted Dissimilarity index.

For any district Z, let the lower-case letter z denote the one-school district that results from combining the students of Z into a single school. Following Hutchens [18], we say that the segregation index S is *additively decomposable* if, for any (nonempty) districts X and Y,

$$S(X \uplus Y) = S(x \uplus y) + \alpha(x, y)S(X) + \beta(x, y)S(Y) \tag{5}$$

where $\alpha(x, y)$ and $\beta(x, y)$ are strictly positive numbers that depend only on the numbers of students in each ethnic group in districts X and Y. That is, the segregation of the combined district $X \uplus Y$ can be written as the sum of segregation between the districts, $S(x \uplus y)$, and the weighted sum of segregation within the districts X and Y, where the weights $\alpha(x, y)$ and $\beta(x, y)$ can depend on the sizes and ethnic distributions of X and Y but *not* on the allocations of students across schools within X and Y.

Proposition 3 *Suppose S is an additively decomposable segregation index. Then the ordering represented by S satisfies Independence.*

The general Atkinson index with weights $\mathbf{w} = (w_1, ..., w_K)$ satisfies (5). More generally, let $Z = X_1 \uplus \cdots \uplus X_t$, where each X_i is a district. Then it is straightforward to verify that

$$A_{\mathbf{w}}(Z) = A_{\mathbf{w}}(x_1 \uplus \cdots \uplus x_t) + \sum_{i=1}^{t} \alpha_i A_{\mathbf{w}}(X_i)$$

where x_i is the district that results from combining the students in X_i into a single school and $\alpha_i = \prod_{g \in \mathbf{G}} \left(\frac{T_g(x_i)}{T_g(z)} \right)^{w_g}$.

Frankel and Volij [14] discuss a stronger type of additive separability, in which the weight α_i equals the proportion of students who are in district i. This stronger property is not satisfied by the Atkinson indices or, indeed, by any of the other common segregation indices except Mutual Information (Frankel and Volij [14]).

6 Empirical School Segregation Patterns

In this section we study the change in school segregation in the U.S. between the 1987/8 and 2005/6 school years. The data source is the Common Core of Data (Sable, Gaviola, and Garofano [34]). The set of reporting schools expanded considerably over the period, making longitudinal comparisons hard to interpret. Accordingly, we restricted to the 62,519 schools that reported positive attendance in every school year from 1987/8 to 2005/6.[17] We

---

[17] To aid in matching, for 1987/8 through 1998/99 we used the 13-year longitudinal version of this database (McLaughlin [29]). For subsequent years, we used the annual files. Schools that closed for one or more

use four, mutually exclusive ethnic groups: Asians, (non-Hispanic) whites, (non-Hispanic) blacks, and Hispanics.[18]

We trace changes in three sets of indices. The first set is the Scale Invariant indices: the Atkinson indices and the Unweighted Dissimilarity Index (section 3). Indices not in this set fall, heuristically, into two groups. Each such index begins with a quantity that captures some intuitive notion of segregation. Sometimes this quantity itself is used as the segregation index: the index is unnormalized. This may be because the index already takes a maximum value of one, or because normalization destroys certain desirable properties. This set consists of the Card-Rothstein Index, the Clotfelter Index, and the Mutual Information index. In other cases, the intuitive quantity is normalized by dividing by the maximum value it can take, given the district's ethnic distribution. This set consists of the Gini Index, the Weighted Dissimiliarity Index, the Normalized Exposure Index, and the Entropy Index.[19]

Over the period we study, school segregation measures were affected by two important developments. First, the distributions of different ethnic groups across schools in the U.S. became increasingly similar. Accordingly, the Scale Invariant indices show steep declines in segregation over the full period. At the same time, ethnic diversity was growing significantly. Most strikingly, the proportion of Hispanic students rose from 10% in 1987/8 to 19.3% in 2005/6. This change dominated for the unnormalized Scale-Variant indices: they show large *increases* over the period. As for the normalized Scale-Variant indices, increased ethnic diversity led to offsetting increases in both the intuitive quantities on which these index are based, as well as the maximum possible values of these quantities. The end result was little discernible change in the indices themselves.

years and then reopened are excluded from our sample. Since parents and teachers may prefer not to move back after they have gotten used to new schools, the sense in which these are actually "the same schools" is open to debate.

[18]The CCD actually has five ethnic groups; the smallest, American Indian/Alaskan Native, is not represented in some school districts. Hence, we include this group with the second smallest group, Asians.

[19]Frankel and Volij [14] study which of our axioms, among others, are satisfied by these indices.

6.1 Unnormalized Scale-Variant Indices

The unnormalized Scale-Variant indices consist of the Mutual Information Index and two other indices. Clotfelter [7] uses the percentage of black students who attend schools in which at least some proportion κ of students are black or Hispanic. We follow Clotfelter [7] by using the two thresholds $\kappa = 0.5$ and $\kappa = 0.9$. Card and Rothstein [5] compute the average fraction black or Hispanic in the schools attended by the typical black and white student, and define their segregation index as the difference between these figures. Letting whites, blacks, and Hispanics be indexed by 1, 2, and 3, respectively, the Card-Rothstein Index equals[20]

$$CR = \sum_{n \in \mathbf{N}} \left(\frac{T_2^n}{T_2} - \frac{T_1^n}{T_1} \right) \frac{T_2^n + T_3^n}{T^n}$$

6.2 Normalized Scale-Variant Indices

We consider four normalized Scale-Variant indices. The Weighted Dissimilarity Index of Morgan [30] and Sakoda [35] is defined as follows:

$$D^W = \frac{1}{I} D' \text{ where } D' = \frac{1}{2} \sum_{g \in \mathbf{G}} \sum_{n \in \mathbf{N}} P^n \left| p_g^n - P_g \right| \tag{6}$$

and I is the Simpson Iteraction Index, $I = \sum_{g \in \mathbf{G}} P_g (1 - P_g)$ (Lieberson [26]). Intuitively, D' equals the minimum proportion of the population that would have to change schools, keeping school sizes fixed, in order for each school to be representative of the district. I is what this proportion would be under complete segregation. Hence, the Weighted Dissimilarity Index, D^W, is a normalization of D' that take a maximum value of 1.

The multigroup Gini Index of Reardon [32], is a generalization of the two-group Gini index of Jahn, Schmidt, and Schrag [21]:

---

[20]In our application, Asians are included in the total school population, T^n.

$$G = \frac{1}{I} G' \text{ where } G' = \frac{1}{2} \sum_{g \in \mathbf{G}} \sum_{m \in \mathbf{N}} \sum_{n \in \mathbf{N}} P^m P^n \left| p_g^m - p_g^n \right|$$

G' is a measure of the extent to which the proportion in a given group varies across schools. More precisely, it is the weighted sum, over all ethnic groups g and all school-pairs, of the absolute difference between the proportions in the two schools who are in group g. The Gini index results from dividing this measure by its maximum value, I.

The Normalized Exposure Index was originally proposed by Bell [3] for the case of two groups. Its multigroup version, formulated by James [23], is

$$P = \sum_{g \in \mathbf{G}} \sum_{n \in \mathbf{N}} P^n \frac{(p_g^n - P_g)^2}{1 - P_g}$$

In the case of two groups (say blacks and whites, denoted 1 and 2, respectively), the index equals $\frac{P_2 - E^*}{P_2}$ where $E^* = \frac{1}{T_1} \sum_{n \in \mathbf{N}} T_1^n p_2^n$ is the proportion white in the school attended by the average black student and P_2 (the proportion white in the district) is the maximum value of E^* given the district ethnic distribution. Thus, the two-group index measures the exposure of blacks to whites, normalized by the maximum possible such exposure. The index is symmetric: it also measures the normalized exposure of whites to blacks.

Finally, the Entropy Index H (Theil [40]; Theil and Finizza [41]) is the result of dividing the Mutual Information Index, M (section 3), by its maximum value, the entropy of the district ethnic distribution: $H = M/h(P)$.

6.3   Findings

We study total segregation among U.S. schools, essentially treating the U.S. as a single district and studying its evolution over time. In contrast, the literature has typically focused on averages of city-level segregation indices. Despite its intuitive appeal, this common approach is generally not well founded. In most cases, the resulting average does not equal the within-city component of total segregation between U.S. schools: most segregation

indices are *not* decomposable into a within-city average plus a between-city term.[21]   In addition, individual city-level indices often have an intuitive meaning that is lost when taking their average across cities. This is particularly true when, as is often the case, normalized indices are used.[22]

Multigroup segregation was affected by two trends during the period we study.   The first is a decline in segregation between pairs of ethnic groups.   Table 2 depicts the Lorenz curves between each pair of students in 1987/8 and 2005/6.   For each pair except blacks and whites, the Lorenz curve rose over the period, indicating that the groups' distributions across schools were becoming increasingly similar.   Indeed, pairwise Gini indices fell for all groups, as shown in panel 1 of Table 3.[23]   For blacks vs. whites, the two curves cross and are close to one another: the change in the Gini index is only -0.005. The 2005/6 curve is higher in schools with moderately high black percentages and lower in the schools with the highest proportions black. This indicates a relative movement of whites into the former set of schools and out of the latter set.

Over the period, there was another important development: minority groups grew in relative terms. In 1987/8, the proportions of Asians, blacks, and Hispanics in U.S. schools were 4.0%, 15.3%, and 10.0%, respectively.   By 2005/6, these percentages had grown to 5.7%, 16.8%, and 19.3%, respectively, while the percentage of whites had fallen from 70.8%

---

[21]Indices that violate Independence are not decomposable into within and between terms (section 5.2). These include the Gini Index, the Card-Rothstein Index, the Normalized Exposure Index with three or more ethnic groups, and the Weighted Dissimilarity Index (Frankel and Volij [14]).   Unweighted Dissimilarity is also in this category (section 5.1).

[22]To see the effects of normalization, consider the population-weighted average of the *unnormalized* dissimilarity index D' defined in equation (6). This average does have an intuition: it is the minimum proportion of students in the country who would have to change schools within their given cities, keeping school sizes fixed, in order for each school to be representative of its city. In contrast, it is hard to find any intuition for the average across cities of the normalized index, D^W.

[23]The pairwise Gini index equals the area between the 45 degree line and the Lorenz curve, as a fraction of the total area that lies below the 45 degree line.

to 58.2%. These figures appear in panel 2 of in Table 3.

Panel 3 of Table 3 shows changes in the Scale Invariant segregation indices. These indices are insensitive to changes in the ethnic distribution, so they would be expected to decline in response to the increasing similarity in the ethnic groups' distributions across schools (Table 2). They consist of the Unweighted Index of Dissimilarity (UIOD), the symmetric Atkinson index (ATKSYM), and two different versions of the general Atkinson index: one in which a group's weight equals its share in the universe of districts in 1987/8 (ATK87) and one in which its weight equals its share in this universe in 2005/6 (ATK05). As anticipated, all of these indices declined over the period.

Unnormalized Scale-Variant indices are shown in panel 4 of Table 3. These indices consist of the Mutual Information Index (M), the Clotfelter Index with thresholds of 50% (Cl50) and 90% (Cl90), and the Card-Rothstein Index (CR). Falling bilateral segregation should cause these indices to fall; on the other hand, increased ethnic diversity tends to have the opposite effect. Over the 18-year period, the second effect dominates: all indices show increases.

For instance, the increase in the Mutual Information index shows that a randomly selected student's school now conveys more information about her race (Frankel and Volij [14]). This is driven by the fact that there is now more information to convey: since ethnic diversity has increased, the initial uncertainty about a random student's race is now greater. Similarly, the increase in the Clotfelter indices shows that a higher percentage of black students now attend schools in which at least a given threshold (50% and 90%) of students are black or Hispanic. This is driven by the increase in the proportions of blacks and Hispanics in the U.S. student population. Finally, the increase in the Card-Rothstein index shows that the absolute difference in the proportion of minorities in the school attended by the typical black vs. white student has grown. This is driven by two factors: growth in the proportion of minority students in the U.S., combined with the lack of progress in black-white integration (Table 2).

Normalized Scale-Variant indices appear in panel 5 of Table 3. This set consists of the

Gini Index (GINI), the Weighted Index of Dissimilarity (WIOD), the Normalized Exposure Index (NEXP), and the Entropy Index (ENT). Each index equals the ratio of some intuitive quantity to its maximum possible value. These indices show little change over the period. In each case, greater ethnic diversity caused offsetting increases in the intuitive quantity as well as in its maximum possible value, with little change in the ratio.

In particular, the Weighted Index of Dissimilarity equals the proportion of students who would have to change schools to attain perfect integration (D'), divided by the maximum possible value of this proportion (I, the Simpson interaction index). While D' rose from 0.301 to 0.370, I also rose, from 0.464 to 0.593 (Table 3, last panel). Their ratio, WIOD, fell only slightly, from 0.648 to 0.624. Similarly, the Gini Index fell slightly, from 0.818 to 0.793. It equals the ratio of the weighted-average absolute difference in ethnic group proportions (G'), which rose from 0.38 to 0.47, divided by the maximum possible value of this weighted average (also the Simpson interaction index, I), which also rose, from 0.464 to 0.593 (panel 6 of Table 3). The Entropy Index fell from 0.456 to 0.429; it equals the expected information a randomly selected student's school reveals about her race (the Mutual Information Index), which rose from 0.586 to 0.679, divided by total information present in a random student's race (the entropy of the overall ethnic distribution), which rose from 1.285 to 1.582 (panel 6 of Table 3).[24]

7   Conclusion

In this paper we have provided an axiomatic foundation for the Atkinson family of segregation orderings in the multigroup setting, using a parsimonious set of purely ordinal axioms. We have shown that the ordering represented by the symmetric Atkinson index is the only (nontrivial) segregation ordering that satisfies Group Symmetry, Scale Invariance, the Weak School Division Property, and Independence. We also showed that a (nontrivial) segre-

---

[24]The Normalized Exposure Index lacks such a simple interpretation since different normalization factors are applied to different terms in the sum.

gation ordering is represented by an index in the general Atkinson family if and only if it satisfies Scale Invariance, the Weak School Division Property, Independence, and a technical continuity property.

While the Atkinson indices are Scale-Invariant, most other multigroup segregation indices are not. We illustrate the role of Scale Invariance by studying changes in the total segregation of U.S. public school students from 1987/8 to 2005/6. For each pair of ethnic groups except blacks and whites, the Lorenz curve rose: the groups's distributions across schools became more similar. This constitutes a decline in bilateral segregation according to Massey and Denton's [28] criterion of evenness. Consistent with this, the Atkinson indices fell considerably over the period. On the other hand, there were large increases in the proportions of all minority groups, especially Hispanics, who overtook blacks as the second-largest ethnic group. This growth in ethnic diversity caused indices that are not Scale-Invariant to rise or remain essentially unchanged.

A Proofs

For any district X and any nonnegative constant c, let cX denote the district that results from multiplying the number of members of each group in each school of X by c. For any district X and any vector of nonnegative scalars $\overrightarrow{\alpha} = (\alpha_g)_{g \in \mathbf{G}}$, let $\overrightarrow{\alpha} * X$ denote the district in which the number of members of group g in school 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 schools; e.g., $\overrightarrow{\alpha} * (1,2) = (2,6)$.

We first define a slight strengthening of WSDP:

School Division Property (SDP) Let $X \in \mathcal{C}$ be any district and let n be a school in X. Let X' be the district that results from X if school n is subdivided into two schools, n_1 and n_2. Then, $X' \succcurlyeq X$. Further, if n_1 and n_2 have the same group distributions (i.e., $p_g^{n_1} = p_g^{n_2}$ for all $g \in \mathbf{G}$), then $X' \sim X$.

SDP follows from WSDP and IND:

Lemma 1 *Suppose the segregation ordering $\succcurlyeq$ satisfies Independence and the Weak School Division Property. Then $\succcurlyeq$ also satisfies the School Division Property.*

Proof. Let Y denote the district X less the school n:

$$X = Y \uplus \langle n \rangle$$
$$X' = Y \uplus \langle n_1, n_2 \rangle \, .$$

By WSDP, $\langle n_1, n_2 \rangle \succcurlyeq \langle n \rangle$. By IND, $Y \uplus \langle n_1, n_2 \rangle \succcurlyeq Y \uplus \langle n \rangle$. If n_1 and n_2 have the same population distribution then the symbol $\succcurlyeq$ can be replaced by $\sim$. Q.E.D.

We now state and prove some additional lemmas.

Lemma 2 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$ that satisfies SDP and SI.*

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

2. *Any district in which every school is representative is weakly less segregated under $\succcurlyeq$ than any district in which some school is unrepresentative.*

Proof.

1. Consider any district Y in which every school is representative. Number the schools $1, ..., N$. For each $i = 1, ..., N$, let Y_i be the district that results from Y when the first i schools of Y are combined into a single school. By SDP, 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 school. But by SI, any district with a single school is as segregated as any other district with a single school.

2. Let Y be a district in which every school is representative and consider any district X in which at least one school is unrepresentative. The above reasoning yields $X \succcurlyeq X_N$. X_N contains a single school, so it is representative. Therefore, by 1, $X \succcurlyeq Y$.

24

Q.E.D.

Lemma 3 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$ that satisfies SDP and SI. All completely segregated districts have the same degree of segregation under $\succcurlyeq$, and are weakly more segregated than any district in which any school is mixed.*

Proof. Consider a completely segregated district X. Let X' be the district that results from X when, for each group $g \in \mathbf{G}$, all schools that contain only members of group g are combined into a single school. (X' thus consists of K schools, each of which contains all the members of a single group.) By iteratively applying SDP, $X \sim X'$. By SI, X' is as segregated as any other district that consists of K schools, each of which contains all the members of a single group. This implies that all completely segregated districts have the same degree of segregation.

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

Let $\underline{X}$ be a district with K groups of unit size who all attend in the same school: $\underline{X} = \left\langle (\underbrace{1, 1, \ldots, 1}_{K \text{ groups}}) \right\rangle$. Let $\overline{X}$ be a district with K groups of unit size who all attend separate schools:

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

We say that a school is a *ghetto* if all its students belong to the same group. We first state and prove some auxiliary results about districts with a single non-ghetto school. For any scalar α, let $X(\alpha)$ denote the district $\alpha \underline{X} \uplus (1 - \alpha)\overline{X}$. City $X(\alpha)$ contains one school with α students of each group, and K ghettos, each with $1 - \alpha$ students. Similarly, for any vector $\mathbf{t} = (t_1, ..., t_K) \in [0, 1]^K$, let $X(\mathbf{t})$ denote the district

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

City $X(\mathbf{t})$ consists of the non-ghetto school $\mathbf{t}$, and for each group g, one ghetto with $1 - t_g$ students of group g.

Lemma 4 *Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$ that satisfies SDP, IND, N, and SI. Then*

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

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

Proof.

1. By N, there exist districts X and Y such that $X \succ Y$. By lemmas 2 and 3, $\overline{X} \succcurlyeq X \succ Y \succcurlyeq \underline{X}$, so $\overline{X} \succ \underline{X}$.

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

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

The result follows from the fact that, by SDP,

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

and

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

Q.E.D.

Lemma 5 *Let $\mathbf{t}, \mathbf{v} \in [0, 1]^K$, such that $\mathbf{t} \le \mathbf{v}$. Then, $X(\mathbf{t}) \succcurlyeq X(\mathbf{v})$. If $\mathbf{t} = (t_1, ..., t_K) \in (0, 1)^K$ then $\overline{X} \succ X(\mathbf{t}) \succ \underline{X}$.*

Proof. Let $\mathbf{t}, \mathbf{v} \in [0, 1]^K$, such that $\mathbf{t} \le \mathbf{v}$. Applying SDP twice, we obtain

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

26

Assume now that $\mathbf{t} = (t_1, ..., t_K) \in (0,1)^K$, and let $\bar{t} = \max\{t_1, ..., t_K\}$, $\underline{t} = \min\{t_1, ..., t_K\}$. Then,

$$
\begin{aligned}
\overline{X} &\succ X(\underline{t}) &&\text{by Lemma 4}\\
&\succcurlyeq X(\mathbf{t}) &&\text{since } (\underline{t}, \ldots, \underline{t}) \leq \mathbf{t}\\
&\succcurlyeq X(\bar{t}) &&\text{since } \mathbf{t} \leq (\bar{t}, \ldots, \bar{t})\\
&\succ \underline{X} &&\text{by Lemma 4.}
\end{aligned}
$$

Q.E.D.

Lemma 6 *For any two vectors* $\mathbf{t}, \mathbf{v} \in [0,1]^K$ *and for any* $\gamma \in (0,1]$,

1. $\mathbf{v}{*}X(\mathbf{t}) \uplus (1 - \mathbf{v}) * \overline{X} \sim X(\mathbf{v} * \mathbf{t})$

2. $\gamma X(\mathbf{t}) \uplus (1 - \gamma)\overline{X} \sim X(\gamma \mathbf{t})$

3. *If for some* $\alpha \in [0,1]$, $X(\mathbf{t}) \sim X(\alpha)$, *then* $X(\mathbf{v} * \mathbf{t}) \sim X(\alpha \mathbf{v})$

4. *If for some* $\alpha \in [0,1]$, $X(\mathbf{t}) \sim X(\alpha)$, *then* $X(\gamma \mathbf{t}) \sim X(\gamma \alpha)$

Proof.

1. By definition of $X(\mathbf{t})$ and by SDP,

$$
\begin{aligned}
\mathbf{v}{*}X(\mathbf{t}) \uplus (1 - \mathbf{v}) * \overline{X} &= \mathbf{v}* \left(\mathbf{t}{*}\underline{X} \uplus (1 - \mathbf{t}) * \overline{X}\right) \uplus (1 - \mathbf{v}) * \overline{X}\\
&\sim (\mathbf{v} * \mathbf{t}) {*}\underline{X} \uplus (1 - \mathbf{v} * \mathbf{t}) * \overline{X}\\
&= X(\mathbf{v} * \mathbf{t}).
\end{aligned}
$$

2. The proof is analogous to the previous one.

3. Now, if for some $\alpha \in [0,1]$, $X(\mathbf{t}) \sim X(\alpha)$, then, by SI and IND

$$
\mathbf{v}{*}X(\mathbf{t}) \uplus (1 - \mathbf{v}) * \overline{X} \sim \mathbf{v}{*}X(\alpha) \uplus (1 - \mathbf{v}) * \overline{X}
$$

which, by the previous steps, implies $X(\mathbf{v} * \mathbf{t}) \sim X(\alpha \mathbf{v})$.

4. The proof is analogous to the previous one.

Q.E.D.

27

A.1 Proof of Proposition 1

That the ordering represented by $A_{\mathbf{w}}$ satisfies N is obvious. It also satisfies C, since $A_{\mathbf{w}}$ is a continuous function. The fact that it satisfies SI follows from the fact that for any positive scalar α, $t_g^n = \frac{T_g^n}{T_g} = \frac{\alpha T_g^n}{\alpha T_g}$. We now show that the ordering satisfies IND and WSDP.

Proof.

IND Let $X, Y \in \mathcal{C}$ be two districts with the same group distributions, and the same total populations, and let $Z \in \mathcal{C}$ be another district. 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 students of the district $X \uplus Z$ attend school $n \in N(X)$. Likewise, a proportion $t_g^n \eta_g$ of group-g students of the district $X \uplus Z$ attend school $n \in N(Z)$. Analogous statements are true for $Y \uplus Z$. Accordingly,

$$A_{\mathbf{w}}(X \uplus Z) \geq A_{\mathbf{w}}(Y \uplus Z)$$

$$\Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \gamma_g \right)^{w_g} \right) + \sum_{n \in N(Z)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \eta_g \right)^{w_g} \right)$$

$$\leq \sum_{n \in N(Y)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \gamma_g \right)^{w_g} \right) + \sum_{n \in N(Z)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \eta_g \right)^{w_g} \right)$$

$$\Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \gamma_g \right)^{w_g} \right) \leq \sum_{n \in N(Y)} \left(\prod_{g \in \mathbf{G}} \left(t_g^n \gamma_g \right)^{w_g} \right)$$

$$\Leftrightarrow \left(\prod_{g \in \mathbf{G}} \left(\gamma_g \right)^{w_g} \right) \sum_{n \in N(X)} \prod_{g \in \mathbf{G}} \left(t_g^n \right)^{w_g} \leq \left(\prod_{g \in \mathbf{G}} \left(\gamma_g \right)^{w_g} \right) \sum_{n \in N(Y)} \prod_{g \in \mathbf{G}} \left(t_g^n \right)^{w_g}$$

$$\Leftrightarrow \sum_{n \in N(X)} \prod_{g \in \mathbf{G}} \left(t_g^n \right)^{w_g} \leq \sum_{n \in N(Y)} \prod_{g \in \mathbf{G}} \left(t_g^n \right)^{w_g}$$

$$\Leftrightarrow A_{\mathbf{w}}(X) \geq A_{\mathbf{w}}(Y)$$

WSDP Let X be a district with a single school and let $X' = \langle (t_g)_{g \in \mathbf{G}}, (1 - t_g)_{g \in \mathbf{G}} \rangle$ the district that results from dividing X into two schools. Then, since $A_{\mathbf{w}}$ maps districts to the unit interval, $A_{\mathbf{w}}(X') \geq 0 = A_{\mathbf{w}}(X)$. Further, if the two schools of X' have the same group distribution, then $t_g = t_{g'}$ for all $g, g' \in \mathbf{G}$, then $A_{\mathbf{w}}(X') = 1 - \prod_{g \in \mathbf{G}} t_g^{w_g} - \prod_{g \in \mathbf{G}} (1 - t_g)^{w_g} = 0$ since the weights w_g add up to one.

GS Since the geometric average is a symmetric function, A satisfies GS.

Q.E.D.

A.2   Proof of Theorem 1

Let $\succcurlyeq$ be a segregation ordering on $\mathcal{C}$ that satisfies C, SDP, IND, N, and SI. We first build an index that represents $\succcurlyeq$. Later we show that the index has the requisite form.

Lemma 7 *For any district X, there is a unique $\alpha_X \in [0,1]$ such that $X \sim X(\alpha_X)$.*

Proof.   By C, $\left\{\alpha \in [0,1] : \alpha\underline{X} \uplus (1-\alpha)\overline{X} \succcurlyeq X\right\}$ and $\left\{\alpha \in [0,1] : X \succcurlyeq \alpha\underline{X} \uplus (1-\alpha)\overline{X}\right\}$ are closed sets. Any α_X satisfies $X \sim X(\alpha_X)$ if and only if it is in the intersection of these two sets. The sets are each nonempty by Lemmas 2 and 3. Their union is the whole unit interval since $\succcurlyeq$ is complete. Since the interval $[0,1]$ is connected, the intersection of the two sets must be nonempty. By Lemma 4, their intersection cannot contain more than one element. Thus, their intersection contains a single element α_X.   Q.E.D.

Let X and Y be two districts, and let α_X, and α_Y be the respective scalars identified in Lemma 7. Then, by Lemma 4, $X \succcurlyeq Y$ if and only if $1 - \alpha_X > 1 - \alpha_Y$, which implies that the index $S : \mathcal{C} \to [0,1]$ defined by $S(Z) = 1 - \alpha_Z$ represents $\succcurlyeq$.

We will now show that the index S has the requisite form.

Proposition 4 *For each group g there is a fixed constant $w_g \geq 0$ such that for any $\beta \in (0,1]$,*

$$X((1,\ldots 1, \beta, 1, \ldots 1)) \sim X(\beta^{w_g})$$

where $(1,\ldots 1, \beta, 1, \ldots 1)$ is a vector with β in the gth place and ones elsewhere.

Proof. For any scalar β and any group g, let $\boldsymbol{\beta}_g$ denote the vector $(1,\ldots 1, \beta, 1, \ldots 1)$ with β in the gth place and ones elsewhere. Let $h_g : (0,1] \to \Re_+$ be the function defined by

$$X(\boldsymbol{\beta}_g) \sim X(h_g(\beta)).$$

29

That is, for each $\beta \in (0, 1]$, $h_g(\beta)$ is the unique scalar δ identified in Lemma 7 such that $X(\boldsymbol{\beta}_g) \sim X(\delta)$. Let $\beta, \beta' \in (0, 1]$. By Lemma 6 (parts 3 and 4), $X(\boldsymbol{\beta}_g * \boldsymbol{\beta}'_g) \sim X(h_g(\beta')\boldsymbol{\beta}_g) \sim X(h_g(\beta)h_g(\beta'))$. Therefore, h_g satisfies the functional equation

$$h_g(\beta\beta') = h_g(\beta)h_g(\beta') \text{ for all } \beta, \beta' \in (0, 1]. \tag{7}$$

Further, by Lemma 5, if $\beta < \beta'$, then $X(h_g(\beta)) \sim X(\boldsymbol{\beta}_g) \succcurlyeq X(\boldsymbol{\beta}'_g) \sim X(h_g(\beta'))$, which, by Lemma 4 implies that $h_g(\beta) \leq h_g(\beta')$. Therefore, h_g is nondecreasing (and as a result it is continuous at at least one point). The substitution $\beta = e^{-u}$, $\beta' = e^{-v}$, $h_g(e^{-u}) = f(u)$, transforms (7) into

$$f(u + v) = f(u)f(v) \text{ for all } u, v \geq 0.$$

Therefore, by Theorem 1 in Aczél [1, pp. 38-39], either (a) f is identically 0, or (b) $f(0) = 1$ and, for all $u > 0$, $f(u) = 0$, or (c) there is w_g such that $f(u) = e^{w_g u}$. This means that either h_g is identically 0, or $h_g(1) = 1$, and $h_g(\beta) = 0$ for $\beta \in (0, 1)$, or there is w_g such that $h_g(\beta) = \beta^{w_g}$.

The function h_g cannot be identically 0 because then, $\underline{X} = X(1, \ldots, 1) \sim X(h_g(1)) \sim X(0) = \overline{X}$, which contradicts nontriviality. Further, the function h_g cannot be such that $h_g(\beta) = 0$ for $\beta \in (0, 1)$, because by Lemma 5 $h_g(\beta) \geq \beta$. Therefore, using the fact that h_g is nondecreasing there is $w_g \geq 0$ such that $h_g(\beta) = \beta^{w_g}$.   Q.E.D.

Proposition 5 *There are fixed, non-negative weights $w_g \geq 0$ for $g = 1, ..., K$ such that for any $\mathbf{t} \in [0, 1]^K$ the unique $\alpha \in [0, 1]$ that satisfies $X(\mathbf{t}) \sim X(\alpha)$ is given by $\prod\limits_{g=1}^{K} (t_g)^{w_g}$. Further, the weights add up to one.*

Proof. Case 1: $\mathbf{t} \in (0, 1]^K$.

Let $\mathbf{t} = (t_1, \ldots, t_K) \in (0, 1]^K$. By Proposition 4,

$$X((1, 1, \ldots 1, t_g, 1, \ldots 1)) \sim X(t_g^{w_g}) \text{ for all } g = 1, \ldots K.$$

Note that $\mathbf{t} = (t_1, 1, ..., 1) * (1, t_2, 1, \ldots 1) * (1, \ldots 1, t_k)$. Then, repeated applications of Lemma 6 then yields

$$X(\mathbf{t}) = X\left(\prod_{g=1}^{K} t_g^{w_g}\right).$$

In order to complete the proof of case 1, we need to show that the weights w_g add up to one. Consider the district $X = X(\alpha)$ where $\alpha \in (0,1)$. By the previous conclusion $X \sim X\left(\prod_{g=1}^{K} \alpha^{w_g}\right)$. By Lemma 4 $\left(\prod_{g=1}^{K} \alpha^{w_g}\right) = \alpha$ which implies that the weights w_g add up to one.

Case 2: $\mathbf{t} \in [0,1]^K \backslash (0,1]^K$.

By Lemma 7 there is an $\alpha \in [0,1]$ such that $X(\mathbf{t}) \sim X(\alpha)$. We need to show that $\alpha = 0$. Let $\mathbf{t}(\varepsilon) = (t_1(\varepsilon), ..., t_K(\varepsilon))$ be the school that results from $\mathbf{t}$ after replacing the 0 components by $\varepsilon > 0$. Since $\mathbf{t} \in (0,1]^K$, by Case 1, $X(\mathbf{t}(\varepsilon)) \sim X(\alpha(\varepsilon))$ where $\alpha(\varepsilon) = \prod_{g=1}^{K} t_g(\varepsilon)^{w_g}$. By Lemma 5, $X(\mathbf{t}) \succcurlyeq X(\mathbf{t}(\varepsilon))$ which implies, $X(\alpha) \succcurlyeq X(\alpha(\epsilon))$. By Lemma 4, $\alpha(\epsilon)) > \alpha \geq 0$. Since $\alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$, we obtain that $\alpha = 0$. Q.E.D.

We now show that the statement of the theorem holds for districts with two non-ghetto schools.

Proposition 6 *Let $\mathbf{t}^1, \mathbf{t}^2 \in [0,1]^K$ and let $X = \langle \mathbf{t}^1, \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, ..., 0), ..., (0, ..., 0, 1 - t_K^1 - t_K^2) \rangle$ be a district. There is a unique $\alpha_X \in [0,1]$ that satisfies $X \sim X(\alpha_X)$. It is given by $\alpha_X = \prod_{g=1}^{K} \left(t_g^1\right)^{w_g} + \prod_{g=1}^{K} \left(t_g^2\right)^{w_g}$, where the weights w_g are those found in Proposition 5.*

Proof. Uniqueness of α_X follow from Lemma 4, so it is enough to show that $\alpha_X = \prod_{g=1}^{K} \left(t_g^1\right)^{w_g} + \prod_{g=1}^{K} \left(t_g^2\right)^{w_g}$ satisfies $X \sim X(\alpha_X)$. Assume first that $t_g^i \leq 1/2$ for $i = 1, 2$ and $g = 1, ..., K$. First suppose that $t_g^i = 0$ for some i and g. Assume WLOG that $t_1^2 = 0$. Then by Proposition 5 and SI,

$$\langle \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, ..., 0), ..., (0, ..., 0, 1 - t_K^1 - t_K^2) \rangle \sim \langle (1 - t_1^1, 0, ..., 0), ..., (0, ..., 0, 1 - t_K^1) \rangle$$

so by IND, $X \sim X(\mathbf{t}^1)$. The result then follows from Proposition 5.

Now suppose that $\mathbf{t}^1, \mathbf{t}^2 \in (0,1]^K$. Assume WLOG that $\prod_{g=1}^{K} \left(t_g^1\right)^{w_g} \leq \prod_{g=1}^{K} \left(t_g^2\right)^{w_g}$. Define $\widetilde{t_g^i} = t_g^i / (1 - t_g^2)$ for $g = 1, ..., K$ and $i = 1, 2$. Note that $\prod_{g=1}^{K} \left(\widetilde{t_g^1}\right)^{w_g} \leq \prod_{g=1}^{K} \left(\widetilde{t_g^2}\right)^{w_g}$. Define $\tau = \prod_{g=1}^{K} \left(\frac{t_g^1}{t_g^2}\right)^{w_g} = \prod_{g=1}^{K} \left(\frac{\widetilde{t_g^1}}{\widetilde{t_g^2}}\right)^{w_g} \leq 1$. We can write

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

31

By SI

$$X \sim Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2} \right) \right\rangle \tag{8}$$

where[25]

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

By Proposition 5,

$$Y \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}. \tag{9}$$

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

$$Y' = \tau \widetilde{t^2} * \underline{X} \uplus \left(\mathbf{1} - \tau \widetilde{t^2} \right) * \overline{X}. \tag{10}$$

We must verify that all entries in Y' are nonnegative. This holds if $\tau \widetilde{t_g^2} \le 1$ for all g. Since $t_g^2 \le 1/2$ for all g, it follows that $\widetilde{t_g^2} \le 1$; since $\tau \le 1$ as well, it follows that $\tau \widetilde{t_g^2} \le 1$. Since $\prod_{g=1}^{K} \left(\tau \widetilde{t_g^2} \right)^{w_g} = \prod_{g=1}^{K} \left(\widetilde{t_g^1} \right)^{w_g} = \alpha_Y$, by Proposition 5,

$$Y' \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}. \tag{11}$$

It follows from (9) and (11) that $Y \sim Y'$. As a result,

$$
\begin{aligned}
X \ &\sim \ Y \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2} \right) \right\rangle & \text{by (8)} \\
&\sim \ Y' \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2} \right) \right\rangle & \text{by IND} \\
&\sim \ \tau \widetilde{t^2} * \underline{X} \uplus \left(\mathbf{1} - \tau \widetilde{t^2} \right) * \overline{X} \uplus \left\langle \left(\widetilde{t_1^2}, ..., \widetilde{t_K^2} \right) \right\rangle & \text{by (10)} \\
&\sim \ (\tau + 1) \widetilde{\mathbf{t}^2} * \underline{X} \uplus \left(\mathbf{1} - \tau \widetilde{\mathbf{t}^2} \right) * \overline{X} & \text{by SDP} \\
&\sim \ (\tau + 1) \mathbf{t}^2 * \underline{X} \uplus \left(\mathbf{1} - (\tau + 1) \mathbf{t}^2 \right) * \overline{X} & \text{by SI and definition of } \widetilde{\mathbf{t}^2}.
\end{aligned}
$$

Therefore, using Proposition 5, $X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X}$, where

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

---

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

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

$$\widehat{X} = \langle \widehat{\mathbf{t}^1}, \widehat{\mathbf{t}^2}, (1 - \widehat{t}_1^1 - \widehat{t}_1^2, 0, ..., 0), (0, 1 - \widehat{t}_2^1 - \widehat{t}_2^2, 0, ..., 0), ..., (0, ..., 0, 1 - \widehat{t}_K^1 - \widehat{t}_K^2) \rangle.$$

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

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

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

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

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

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

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

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

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

Proposition 7 *For every district $X \in \mathcal{C}$ there is a unique $\alpha_X \in [0,1]$ such that $X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X}$. Further, this unique α_X is $\sum_{n \in N(X)} \prod_{g=1}^K \left(t_g^n \right)^{w_g}$, where the weights w_g are those found in Proposition 5.*

Proof. By SI it is enough to prove the statement for districts where all groups have a population measure of one. Also, by SDP we can restrict attention to districts where for each group there is at most one ghetto. The proof is by induction on the number of non-ghetto schools. Propositions 5 and 6 already show the that the statement is true for districts with at most two non-ghetto schools. Assume that the statement of the theorem holds for all districts with $m - 1$ non-ghetto schools, let

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

be a district with m non-ghetto schools. Then one can write

$$X = Y \uplus \langle \mathbf{t}^m \rangle$$

where Y denotes X with school $\mathbf{t}^m$ removed. Y has $m-1$ non-ghetto schools. By SI

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

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

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

Using (in order) IND, SI, and Proposition 6,

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

$$\sim \alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X} \uplus \left\langle \left(\frac{t_1^m}{1-t_1^m}, ..., \frac{t_K^m}{1-t_K^m} \right) \right\rangle$$

$$\sim (1-t_1^m, ..., 1-t_K^m) * \left(\alpha_Y \underline{X} \uplus (1-\alpha_Y)\overline{X} \right) \uplus \langle \mathbf{t}^m \rangle$$

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

where

$$\alpha_X = \prod_{g=1}^{K} \left(1-t_g^m \right)^{w_g} \alpha_Y + \prod_{g=1}^{K} \left(t_g^m \right)^{w_g}$$

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

$$= \sum_{n=1}^{m} \prod_{g=1}^{K} \left(t_g^n \right)^{w_g}.$$

Q.E.D.

This completes the proof of Theorem 1.

A.3   Proof of Theorem 2

Proposition 1 implies that the symmetric Atkinson index A satisfies all the axioms of the theorem. We now show that it is the only index to do so. We now show that any ordering

that satisfies GS, SI, SDP, IND, and N on $\mathcal{C}$ must be the symmetric Atkinson ordering. Let $\succcurlyeq$ be such an ordering.

Proposition 8 *Let* $\mathbf{t} = (t_1, \ldots, t_K) \in [0, 1]^K$ *and let* $X = X(\mathbf{t})$. *Then, there exists a unique* $\alpha_X \in [0, 1]$ *such that* $X \sim X(\alpha_X)$. *Further, this unique* α_X *is* $\left(\prod_{g=1}^{K} t_g \right)^{1/K}$.

Proof. Uniqueness follows from Lemma 7. For existence, there are two cases.

Case 1: Suppose $t_g = 0$ for some g. In this case we have to show that $\alpha_X = 0$ or, equivalently, that $X \sim \overline{X}$. By GS, we can assume w.l.o.g. that $t_1 = 0$. Therefore $\mathbf{t} = (0, t_2, t_3, \ldots, t_K)$. Let σ_{12} be the permutation that relabels groups 1 and 2 into 2 and 1, respectively. Therefore, $\sigma_{12}\mathbf{t} = (t_2, 0, t_3, \ldots, t_K)$. Let $\mathbf{1}$ denote a vector of K ones. By GS,

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

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

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

Hence, by SDP and GS,

$$(\boldsymbol{\gamma} * \mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \boldsymbol{\gamma} * \mathbf{t}) * \overline{X} \sim (\boldsymbol{\gamma} * \sigma_{12}\mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \boldsymbol{\gamma} * \sigma_{12}\mathbf{t}) * \overline{X}$$
$$\sim [\sigma_{12}(\boldsymbol{\gamma} * \sigma_{12}\mathbf{t})] * \underline{X} \uplus (\mathbf{1} - [\sigma_{12}(\boldsymbol{\gamma} * \sigma_{12}\mathbf{t})]) * \overline{X}. \qquad (13)$$

But note that since $(\boldsymbol{\gamma} * \mathbf{t}) = \mathbf{t}$, and $\sigma_{12}(\boldsymbol{\gamma} * \sigma_{12}\mathbf{t}) = (0, \beta t_2, t_3, \ldots, t_K)$, we can write (13) as

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

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

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

namely,

$$X \sim \beta\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \beta\mathbf{t}) * \overline{X} \qquad \text{for all } \beta \in (0, 1). \qquad (14)$$

Now choose some constants $\beta, \beta' \in (0,1)$, $\beta > \beta'$. It follows from (14) that

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

Since $\beta\mathbf{t} = \underbrace{\beta'\mathbf{t}}_{\geq 0} + \underbrace{(\beta - \beta')\mathbf{t}}_{\geq 0}$, and $\mathbf{1} - \beta'\mathbf{t} = \underbrace{(\beta - \beta')\mathbf{t}}_{\geq 0} + \underbrace{(\mathbf{1} - \beta\mathbf{t})}_{\geq 0}$, by SDP

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

Note that $(\mathbf{1} - \beta\mathbf{t}) = \underbrace{(\beta - \beta')(\mathbf{1} - \mathbf{t})}_{\geq 0} + \underbrace{[(1 - \beta)\mathbf{1} + \beta'(\mathbf{1} - \mathbf{t})]}_{\geq 0}$, so we can subdivide $(\mathbf{1} - \beta\mathbf{t}) * \overline{X}$ in the above expression using SDP again and get

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

By IND,

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

Finally by SI, $\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \sim \mathbf{t} * \overline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} = \overline{X}$, as claimed. Q.E.D.

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

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

We shall show that $X \sim Y$ and therefore that α is the α_X we are looking for.

Let $\gamma_1 \in (0,1)$. For $g = 2, ..., K$, define $\gamma_g = \gamma_{g-1}\frac{t_{g-1}}{\alpha}$. Note that by definition of α,

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

Now choose γ_1 small enough that each $\gamma_g \leq 1$; this holds if

$$\max_{g \in \langle 2, ..., K\rangle} \gamma_g = \max_{g \in \langle 2, ..., K\rangle} \gamma_1 \prod_{j=2}^{g}\left(\frac{t_{j-1}}{\alpha}\right) \leq 1.$$

36

Denote by $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ the K-tuple just built. Note that $\alpha\boldsymbol{\gamma}$ is a permutation of $\boldsymbol{\gamma} * \mathbf{t}$. Now by definition of X and Y, by SI and IND, and by SDP

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

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

Proposition 9 *Let* $\mathbf{t}^1, \mathbf{t}^2 \in [0, 1]^K$ *and let* $X = \langle \mathbf{t}^1, \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, ..., 0), (0, 1 - t_2^1 - t_2^2, 0, .., 0), ..., (0, ..., 0, 1 - t_K^1 - t_K^2) \rangle$ *be a district. Then there is* $\alpha_X \in [0, 1]$ *such that* $X \sim X(\alpha_X)$. *Further,* α_X *is* $\left(\prod_{g=1}^{K} t_g^1\right)^{1/K} + \left(\prod_{g=1}^{K} t_g^2\right)^{1/K}$.[26]

Proof. The proof is almost identical to the proof of Proposition 6. The only difference is that here the weights are $w_g = 1/K$, and instead of relying on Proposition 5 one needs to rely on the analogous Proposition 8. Q.E.D.

Proposition 10 *For every district* X *there is a unique* $\alpha_X \in [0, 1]$ *such that* $X \sim \alpha_X\underline{X} \uplus (1 - \alpha_X)\overline{X}$. *Further, this unique* α_X *is* $\sum_{n \in N(X)} \left(\prod_{g=1}^{K} t_g^n\right)^{1/K}$.[27]

Proof. The proof is almost identical to the proof of Proposition 9. The only difference is that here the weights are $w_g = 1/K$, and instead of relying on Proposition 5 and 6 one needs to rely on the analogous Propositions 8 and 9. This ends the proof of the theorem. Q.E.D.

---

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

[27] By the reasoning given in footnote 26, α_X must lie between zero and one.

A.4 Proof of Proposition 2

GS: The fact that Group Symmetry is independent of the other axioms follows directly from Theorems 1 and 2: for any vector of weights $\mathbf{w} \neq (1/K, \ldots 1/K)$, the Atkinson index $A_\mathbf{w}$ represents a segregation order that satisfies SI, WSDP, IND, N, and C, but fails GS.

WSDP: To see that WSDP is independent of the other axioms note that since the symmetric Atkinson ordering satisfies GS, SI, IND, C and N, so does the order represented by the index $1 - A$ (defined by $(1 - A)(X) = 1 - A(X)$). It is clear that this order does not satisfy WSDP.

N: The trivial segregation order, which ranks all districts as equally segregated, violates N while satisfying all the other axioms.

IND: Consider the Unweighted Dissimilarity index D^U. It is clear it satisfies N and GS. It satisfies C since it is represented by a continuous function. SI follows from the fact that for any positive scalar α, $t_g^n = \frac{T_g^n}{T_g} = \frac{\alpha T_g^n}{\alpha T_g}$. WSDP holds since $D^U(X) \geq 0$ for all districts X, and $D^U(X) = 0$ if all the schools of X are representative. As for IND, consider the following districts: $X = \langle (2,4), (2,0) \rangle$ and $Y = \langle (4,2), (0,2) \rangle$. One computes $D^U(X) = D^U(Y) = 1/2$. Consider now the result of annexing to them the one-school district $Z = \langle (4,0) \rangle$. One can verify that $D^U(X \uplus Z) = 3/4$ while $D^U(Y \uplus Z) = 1/2$. Hence, D^U violates IND.

SI: The Mutual Information index M clearly violates SI. Since the entropy function is symmetric, M satisfies GS. Since M is continuous, it also satisfies C. That it satisfies WSDP follows from the fact that $M(X) \geq 0$ for all districts X, and that $M(X) = 0$ if all the schools of X are representative. For a proof that the mutual information ordering satisfies IND, see Frankel and Volij [14].

C: Let $\mathbf{w} = (w_g)_{g=1}^K$ and $\mathbf{w}' = (w_g')_{g=1}^K$ be two different vectors of weights that each sum to one. It is easy to verify that $\succcurlyeq_{\mathbf{w},\mathbf{w}'}$ satisfies SI, IND, WSDP, and N since $A_\mathbf{w}$ and $A_{\mathbf{w}'}$ do. It clearly violates GS since at least one weight vector must be asymmetric. In addition, it violates C. To see why, let X and Y be two districts with different group distributions such that $A_\mathbf{w}(X) = A_\mathbf{w}(Y) < 1$ and $A_{\mathbf{w}'}(X) < A_{\mathbf{w}'}(Y)$. Let $c \in (0,1)$ and consider the district $cX \uplus (1-c)Y$. Let $\gamma_g = \frac{cT_g(X)}{cT_g(X) + (1-c)T_g(Y)}$ and $\eta_g = 1 - \gamma_g$. Note that a proportion $t_g^n(X)\gamma_g$ of group-g students of the district $cX \uplus (1-c)Y$ attend school $n \in N(X)$. Likewise, a

proportion $t_g^n(Y)\eta_g$ of group-g students of the district $cX \uplus (1-c)Y$ attend school $n \in N(Y)$. Therefore, we can write

$$1 - A_{\mathbf{w}}(cX \uplus (1-c)Y) = \sum_{n \in N(X)} \prod_{g \in \mathbf{G}} (t_g^n(X)\gamma_g)^{w_g} + \sum_{n \in N(Y)} \prod_{g \in \mathbf{G}} (t_g^n(Y)\eta_g)^{w_g}$$

$$= \sum_{n \in N(X)} \prod_{g \in \mathbf{G}} \left(t_g^n(X)\right)^{w_g} \left(\gamma_g\right)^{w_g} + \sum_{n \in N(Y)} \prod_{g \in \mathbf{G}} \left(t_g^n(X)\right)^{w_g} \left(\eta_g\right)^{w_g}$$

$$= \left(\prod_{g \in \mathbf{G}} (\gamma_g)^{w_g}\right) \sum_{n \in N(X)} \prod_{g \in \mathbf{G}} \left(t_g^n(X)\right)^{w_g} + \left(\prod_{g \in \mathbf{G}} (\eta_g)^{w_g}\right) \sum_{n \in N(Y)} \prod_{g \in \mathbf{G}} \left(t_g^n(Y)\right)^{w_g}$$

$$= (1 - A_{\mathbf{w}}(X)) \prod_{g \in \mathbf{G}} (\gamma_g)^{w_g} + (1 - A_{\mathbf{w}}(Y)) \prod_{g \in \mathbf{G}} (\eta_g)^{w_g} .$$

Since the group distributions of X and Y are not the same, there are groups $g, g' \in \mathbf{G}$ with $\gamma_g \neq \gamma_{g'}$. (Otherwise, for all groups g, γ_g equals a constant λ, which implies $\frac{T_g(X)}{T_g(Y)} = \frac{\lambda(1-c)}{c(1-\lambda)}$. Hence, X and Y must have the same group distribution, a contradiction.) Therefore, the geometric average $\prod_{g \in \mathbf{G}} (\gamma_g)^{w_g}$ is strictly lower than the corresponding arithmetic average, and the same is true for $\prod_{g \in \mathbf{G}} (1 - \gamma_g)^{w_g}$. As a result,

$$1 - A_{\mathbf{w}}(cX \uplus (1-c)Y) < (1 - A_{\mathbf{w}}(X)) \sum_{g \in \mathbf{G}} w_g \gamma_g + (1 - A_{\mathbf{w}}(Y)) \sum_{g \in \mathbf{G}} w_g \eta_g.$$

(By assumption, $A_{\mathbf{w}}(X)$ and $A_{\mathbf{w}}(Y)$ are strictly less than one.). Since $A_{\mathbf{w}}(X) = A_{\mathbf{w}}(Y)$, and since c was arbitrary chosen from $(0, 1)$, we obtain that $A_{\mathbf{w}}(cX \uplus (1-c)Y) > A_{\mathbf{w}}(Y)$ for all $c \in (0, 1)$. Consequently the set

$$\{c \in [0, 1] : cX \uplus (1-c)Y \succcurlyeq_{\mathbf{w}, \mathbf{w}'} Y\}$$

equals $[0, 1)$, which is not closed. Q.E.D.

A.5 Proof of Proposition 3

Let $X, Y \in \mathcal{C}$ be two districts with the same group distributions, and the same total populations. Let $Z \in \mathcal{C}$ be another district. We wish to show that $S(X) \geq S(Y)$ if and only if

$S(X \uplus Z) \geq S(Y \uplus Z)$. Note that

$$S(X \uplus Z) = S(x \uplus z) + \alpha(x, z)S(X) + \beta(x, z)S(Z) \text{ by (5)}$$
$$= S(y \uplus z) + \alpha(y, z)S(X) + \beta(y, z)S(Z) \text{ since } x = y$$

while $S(Y \uplus Z) = S(y \uplus z) + \alpha(y, z)S(Y) + \beta(y, z)S(Z)$ by (5). Since $\alpha(x, y) > 0$ by assumption, $S(X \uplus Z) - S(Y \uplus Z)$ is proportional to $S(X) - S(Y)$.

References

[1] Aczél, J. 1966. "Lectures on Functional Equations and their Applications." Academic Press, NewYork and London.

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

[3] Bell, Wendell. 1954. "A Probability Model for the Measurement of Ecological Segregation." Social Forces 32:357-364.

[4] Boozer, Michael A., Alan B. Krueger, and Shari Wolkin. 1992. "Race and School Quality Since Brown v. Board of Education." Brookings Papers on Economic Activity: Microeconomics 1992, pp. 269-338.

[5] Card, David, and Jesse Rothstein. 2007. "Racial Segregation and the Black-White Test Score Gap." Journal of Public Economics 91:2158-2184.

[6] Clotfelter, Charles T. 1979. "Alternative Measures of School Desegregation: A Methodological Note." Land Economics 54: 373-380.

[7] Clotfelter, Charles T. 1999. "Public School Segregation in Metropolitan Areas." Land Economics 75:487-504.

[8] Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. "Achievement and Segregation in Secondary Schools: A Further Look at Public and Private School Differences." Sociology of Education 55:162-182.

[9] 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.

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

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

[12] Duncan, Otis D., and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indices." American Sociological Review 20:210-217.

[13] Echenique, Federico, and Roland G. Fryer, Jr. 2007. "A Measure of Segregation Based on Social Interactions." Quarterly Journal of Economics, 122:441-485.

[14] Frankel, David M., and Oscar Volij. 2007. "Measuring Segregation." Mimeo.

[15] Guryan, Jonathan. 2005. "Desegregation and Black Dropout Rates." American Economic Review 94:919-943.

[16] Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." NBER Working Paper No. W8741.

[17] Hoxby, Caroline M. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.

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

[19] Hutchens, Robert. 2003. "One Measure of Segregation." International Economic Review 45:555–578.

[20] Ioannides, Yannis M., and Linda D. Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality." Journal of Economic Literature XLII:1056-1093.

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

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

[23] James, Franklin J. 1986. "A New Generalized 'Exposure-based' Segregation Index." Sociological Methods and Research 14:301-316.

[24] Johnson, Ron, Michael Poulsen, and James Forrest. 2007. Ethnic and Racial Segregation in U.S. Metropolitan Areas, 1980-2000: The Dimensions of Segregation Revisited. Urban Affairs Review 42:479-504.

[25] 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.

[26] Lieberson, Stanley. 1969. "Measuring Population Diversity." American Sociological Review 34:850-862.

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

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

[29] McLaughlin, Don. 2003. "CCD Data File: Thirteen-Year Longitudinal Common Core of Data, Non-Fiscal Survey Database, School Years 1986-87 1998-99." Mimeo, American Institutes for Research.

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

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

[32] Reardon, Sean F. 1998. "Measures of Racial Diversity and Segregation in Multigroup and Hierarchically Structured Populations." Presented at the Annual Meeting of the Eastern Sociological Society, Philadelphia, PA.

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

[34] Sable, J., Gaviola, N., and Garofano, A. 2007. Documentation to the NCES Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 200506 (NCES 2007-365). U.S. Department of Education, Washington, D.C.: National Center for Education Statistics.

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

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

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

[38] Taeuber, Karl E., and David R. James. 1982. "Racial Segregation Among Public and Private Schools." Sociology of Education 55:133-143.

[39] Theil, Henri. 1971. "Principles of Econometrics." Wiley & Sons. New York.

[40] Theil, Henri. 1972. "Statistical Decomposition Analysis." North Holland.

[41] Theil, Henri., and A. J. Finizza 1971. "A Note on the Measurement of Racial Integration in Schools." Journal of Mathematical Sociology 1:187–1193.

[42] Zoloth, Barbara S. 1976. "Alternative Measures of School Segregation." Land Economics 52:278-298.

White vs. Black

White vs. Hispanic

White vs. Asian

Black vs. Hispanic

Black vs. Asian
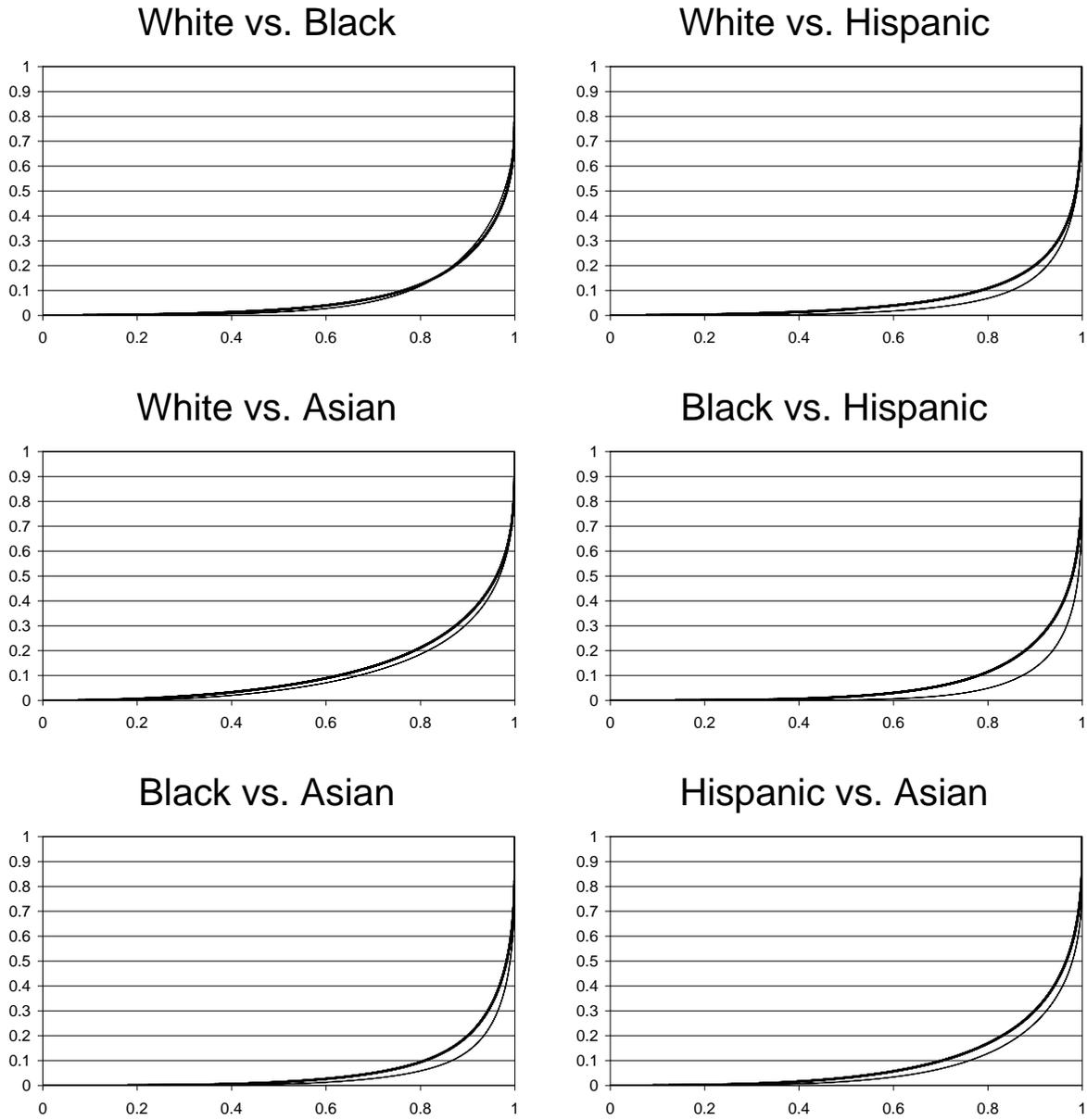

Hispanic vs. Asian

Table 2: Lorenz Curves for Pairs of Ethnic Groups, 1987-8 (thin curve) and 2005-6 (thick curve). Universe is set of U.S. public schools that report positive numbers of students in Common Core of Data for all years from 1987-8 to 2005-6. Schools are sorted in decreasing order of the ratio of the number students in the first group to the number of students in the second group. As each school is considered in turn, the cumulative proportions of students in each group are plotted. The first group appears on the horizontal axis; the second group, on the vertical axis.

	School Year		
	1987/88	**2005/06**	**Change**
1. PAIRWISE GINI COEFFICIENTS			
White-Black	0.851	0.846	-0.005
White-Hispanic	0.899	0.861	-0.038
White-Asian	0.791	0.760	-0.031
Black-Hispanic	0.918	0.854	-0.064
Black-Asian	0.910	0.872	-0.038
Hispanic-Asian	0.846	0.803	-0.043
2. PUBLIC SCHOOL ETHNIC DISTRIBUTION			
Asians	4.0%	5.7%	1.7%
Whites	70.8%	58.2%	-12.6%
Blacks	15.3%	16.8%	1.5%
Hispanics	10.0%	19.3%	9.3%
TOTAL	100.0%	100.0%	0.0%
3. SCALE INVARIANT INDICES			
UIOD (Unweighted Dissimilarity)	0.639	0.599	-0.040
ATKSYM (Asymmetric Atkinson)	0.640	0.528	-0.112
ATK05 (Atkinson; 2005/06 weights)	0.614	0.472	-0.142
ATK87 (Atkinson; 1987/88 weights)	0.573	0.406	-0.167
4. UNNORMALIZED SCALE-VARIANT INDICES			
M (Mutual Information)	0.586	0.679	0.093
Cl50 (Clotfelter with 50% threshold)	0.572	0.681	0.108
Cl90 (Clotfelter with 90% threshold)	0.268	0.325	0.057
CR (Card-Rothstein)	0.461	0.482	0.021
5. NORMALIZED SCALE-VARIANT INDICES			
GINI	0.818	0.793	-0.024
WIOD (Weighted Dissimilarity)	0.648	0.624	-0.024
NEXP (Normalized Exposure)	0.440	0.446	0.006
ENT (Entropy Segregation Index)	0.456	0.429	-0.027
6. MISCELLANEOUS			
D' (Unnormalized Dissimilarity)	0.301	0.370	0.069
G' (Unnormalized GINI)	0.380	0.470	0.091
I (Simpson Interaction Index)	0.464	0.593	0.129
h(P) (Entropy of US Ethnic Distribution)	1.285	1.582	0.297
Number of Schools	62519	62519	0.000
Number of Students (millions)	33.29	34.08	0.796

Table 3: Summary Statistics for U.S. Public Schools, 1987/8 and 2005/6. Universe is set of U.S. public schools that report positive numbers of students in Common Core of Data for all school years from 1987/8 to 2005/6.