

ENTROPY-BASED SEGREGATION INDICES

Ricardo Mora*
Javier Ruiz-Castillo*



Recent research has shown that two entropy-based segregation indices possess an appealing mixture of basic and subsidiary but useful properties. It would appear that the only fundamental difference between the mutual information or M index, and the entropy information or H index, is that the second is a normalized version of the first. This paper introduces another normalized index in that family, the H^ index, which captures segregation as the tendency of racial groups to have different distributions across schools. More importantly, the paper shows that applied researchers may do better using the M index than using either H or H^* in two circumstances: (1) if they are interested in the decomposability of the measurement of segregation, and (2) if they are interested in a margin-free measurement of segregation changes. The shortcomings of the H and H^* indices are illustrated below by means of numerical examples, as well as with school segregation data by ethnic group in the U.S. public school system between 1989 and 2005.*

1. INTRODUCTION

Segregation measures describe differences in the distribution of two or more demographic groups (genders, racial/ethnic groups) over a set of

The authors acknowledge financial support from the Spanish DGI, Grants ECO2009-11165 and SEJ2007-67436. We thank the editor and two anonymous referees for valuable comments. Direct correspondence to Ricardo Mora at ricmora@eco.uc3m.es.

*Universidad Carlos III de Madrid

organizational units (occupations, neighborhoods, schools). As with the measurement of other complex, multifaceted phenomena in the social sciences—such as income inequality or economic poverty—it should come as no surprise that there exists a plethora of indicators capturing different aspects of the same phenomenon (surveys include James and Taeuber [1985]; Massey and Denton [1988]; and Flückiger and Silber [1999]). In some circumstances, this multiplicity of potential measures does not cause any practical problem. In most applications, however, different indices will lead to different conclusions, making it relevant to seek criteria to discriminate between the admissible alternatives.

As in the income inequality literature, one way to select an adequate segregation measure is to study which properties different indices satisfy. For example, in many practical situations it is important to study segregation at several levels simultaneously. For that purpose, it is convenient to use additively decomposable segregation indices that for any partition of organizational units into *clusters* or demographic groups into *supergroups* allow us to express overall segregation as the sum of a between-groups term and a within-groups term.¹ This paper studies in depth three additively decomposable segregation indices that are related to the entropy concept first imported from information theory to the social sciences by Theil (1967, 1971):

1. The mutual information, or *M* index, an unbounded index first proposed by Theil (1971) and whose ordinal ranking has been recently characterized by Frankel and Volij (forthcoming).
2. The entropy, information or *H* index, a normalization of the *M* index by the ethnic group entropy, which was first introduced by Theil and Finizza (1971) and Theil (1972) for the two-group case,

¹ Examples of clusters in the school segregation context are the set of public or private schools in a country, or the sets of schools in major regions, states, cities, school districts, or neighborhoods. In the occupational segregation context, we can have clusters of occupations in professional categories, economic activity sectors, or two- or three-digit occupations. Of course, supergroups can be defined only in a multigroup segregation context. Examples in a school or residential context can be seen when precisely defined ethnic categories, such as Mexican or Puerto Rican, are aggregated into a major category such as Hispanic. In an occupational context, supergroups appear when different categories of female and male workers are aggregated into people of both genders of different age and/or educational attainment.

and was later extended to the multigroup case by Reardon and Firebaugh (2002).

3. The H^* index, a normalization of the M index by the organizational unit entropy, which is proposed in this paper for the first time.

In empirical contexts where it is advisable to use decomposable segregation indices, such as the entropy-based ones, a key question arises: Which index should be used? This is an important issue in a scenario in which, except for Frankel and Volij (forthcoming) in school segregation and Fuchs (1975), Mora and Ruiz-Castillo (2003, 2004), and Herranz, Mora, and Ruiz-Castillo (2005) in occupational segregation, the authors who have used an entropy-based index have preferred the H index.² Taking as reference the school segregation problem in the multiracial case, this paper establishes the practical and conceptual advantages of the M index in multilevel studies of segregation and its trends for the following reasons.

1. Assume, for example, that we want to assess the degree to which overall school segregation is due to racial differences across school districts, or how much is due to segregation within a large supergroup consisting of all minority races in the "United States". As pointed out in the income inequality literature, these deceptively simple questions raise a number of conceptual and methodological problems (Shorrocks 1988:435). This paper shows that the empirical questions usually asked in decomposability analysis receive the less ambiguous answers that are possible in a segregation context when the segregation measure satisfies two strong decomposability properties. These properties require that the within-groups term is the weighted average of segregation in each cluster or supergroup with weights equal to their demographic shares. However, as soon as these properties are imposed on segregation measures we are left solely with the M index (Frankel and Volij

² Theil and Finizza (1971) introduce the H index for the study of school segregation in the two-group case. Reardon, Yun, and McNulty (2000) distinguish between the central city and the suburbs in a study of within-cities school segregation in the multigroup case, while Miller and Quigley (1990) and Fisher (2003) on the one hand and Iceland (2002) on the other study within-cities and within-regions residential segregation. Fisher et al. (2004), who offer the only contribution on residential segregation that develops a full multilevel approach using the H index, only report pair wise comparisons of racial/ethnic groups.

forthcoming)³, which hence becomes the only index that provides unambiguous answers in decomposability analysis.

2. It turns out that the H and H^* indices—like all bounded segregation measures—violate these strong decomposability properties (Frankel and Volij forthcoming, Claim 2). However, Reardon et al. (2000) show that the H index satisfies some weaker decomposability properties, while we show that this is also the case for the H^* index. The decomposition of organizational units into clusters according to the H index and the decomposition of demographic groups into supergroups according to the H^* index are free from ambiguities. This paper establishes that, unfortunately, this is not the case for the decomposition into supergroups according to the H index, as well as the decomposition into clusters according to the H^* index. Moreover, the weights in all the decompositions for the H and the H^* indices are not invariant to changes in the within-group distributions, leading to additional problems of interpretation. The shortcomings of the decompositions of the H and H^* indices are illustrated below by means of numerical examples, as well as school segregation data by ethnic group in the U.S. public school system between 1989 and 2005.

3. One well-known problem with M and its normalized versions H and H^* is that they are not margin free. First, they violate the composition invariance property (*I1* hereafter), satisfied by the segregation indices used by sociologists and economists in a majority of empirical studies. An index violates *I1* if it changes when the number of people in a given demographic group is multiplied by the same positive constant throughout all organizational units. Second, they violate the occupational invariance property (*I2* hereafter), discussed in the literature on occupational segregation by gender in the 1980s. An index does not satisfy *I2* if it changes when the number of people in a given organizational unit is multiplied by the same constant throughout all demographic groups. Therefore, the three entropy-based indices mix up segregation changes with changes in the marginal distributions in segregation comparisons over time or across space. However, the M index admits two decompositions that isolate one term that captures segregation changes net of the impact of pure demographic factors (Mora and

³ Similar results are obtained for the class of relative income inequality entropy indices for different versions of the decomposability properties (Bourguignon 1979; Shorrocks 1980, 1984, 1988; Foster 1983).

Ruiz-Castillo 2009). This paper presents the first evidence showing the advantages of using the M index rather than the H and H^* indices to deal with these issues by means of numerical examples, and in the context of inter-temporal changes of school segregation in the U.S. public school sector between 1989 and 2005.

The rest of this paper is organized into six sections. Section 2 introduces the notation, and establishes that, for strongly school and group decomposable segregation indices, the empirical questions usually asked in decomposition analysis are free of ambiguities. Section 3 introduces the entropy-based segregation indices. Section 4 disentangles the different problems of interpretation that plague the weak decomposability properties satisfied by the H and the H^* indices. Section 5 discusses the invariance properties, while Section 6 briefly discusses the normalization issue. Section 7 concludes the discussion.

2. NOTATION AND STRONG DECOMPOSABILITY PROPERTIES

2.1. Notation

It would be useful to refer to a specific segregation problem. The case discussed throughout the paper is the school segregation problem. Assume that a city X consists of N schools, indexed by $n = 1, \dots, N$. Each student belongs to any of G racial groups, indexed by $g = 1, \dots, G$. However, given the racial diversity existing in many countries, this paper studies the multigroup case where $G \geq 2$. The data available can be organized into the $G \times N$ matrix

$$X = \{t_{gn}\} = \begin{bmatrix} t_{11} & \dots & t_{1N} \\ \vdots & \ddots & \vdots \\ t_{G1} & \dots & t_{GN} \end{bmatrix},$$

where t_{gn} is the number of individuals of racial group g attending school n , so that $t = \sum_{n=1}^N \sum_{g=1}^G t_{gn}$ is the total student population.

The information contained in the joint absolute frequencies of racial groups and schools, t_{gn} , is usually summarized by means of numerical indices of segregation. Let $\Xi(G, N)$ be the set of all cities with

G groups and N schools. A segregation index S is a real valued function defined in $\Xi(G, N)$, where $S(X)$ provides the extent of school segregation for any city $X \in \Xi(G, N)$. The following notation will be useful:

- $p_{gn} = t_{gn} / t$: proportion of students in group g and school n in the city,
- $p_{g|k}$: proportion of students in group g whose school n is located in school district k ,
- p_k : proportion of students in the city whose schools are located in school district k ,
- p_g : proportion of students in the city who belong to group g ,
- $p_{n|l}$: proportion of students in supergroup l who study in school n ,
- p_l : proportion of students who belong to supergroup l in the city,
- $p_{g|n}$: proportion of students in school n who belong to group g ,
- $p_{n|g}$: proportion of students in group g who study in school n , and
- p_n : proportion of students who study in school n in the city.

While lowercase p denotes a proportion, capital P denotes the vector of proportions that describes the associated discrete distributions. For example, P_{gn} will refer to the joint ethnic and school discrete distribution of city X . In the sections that follow, the discussion will be restricted to indices that capture a *relative* view of segregation in which all that matters is the joint distribution—that is, those indices that admit a representation as a function of P_{gn} .⁴

2.2. Strong School Decomposability

In many research situations it is useful to partition organizational units into clusters of different sizes. For example, we may want to assess the

⁴ This property, satisfied by most segregation indices, is referred to as *size invariance* in James and Taeuber (1985) and as *weak scale invariance* in Frankel and Volij (forthcoming). For a study that focuses on translation invariant segregation indices that represent an absolute view of segregation, see Chakravarty and Silber (1992).

degree to which overall school segregation is due to racial differences across school districts. Consider then a partition of the N schools into $K < N$ school districts, $X = X^1 \cup \dots \cup X^k \dots \cup X^K$, where X^k is the set of schools that belong to district k . In addition, let \tilde{X}^k refer to the district in which all schools in district k have been combined into a single school with conditional racial distribution $P_{g|k}$.

Following Frankel and Volij (forthcoming), a school segregation index S is said to be *strongly school decomposable (SSD)* if and only if for any partition $X = X^1 \cup \dots \cup X^k \dots \cup X^K$ of the schools into K clusters overall segregation, $S(X)$, can be written as

$$S(X) = S(\tilde{X}^1 \cup \dots \cup \tilde{X}^K) + \sum_{k=1}^K p_k S(X^k). \quad (1)$$

Therefore, if a school segregation index is *SSD*, then overall segregation can be expressed as the sum of two terms, one that captures between-groups segregation and one that captures within-groups segregation and is equal to the weighted average of segregation levels within each of the clusters, with weights equal to the demographic importance of each cluster.

For any partition of schools into clusters, we have to make sure that three magnitudes are well defined: (1) the contribution to overall segregation of any individual cluster; (2) the part of overall segregation accounted for by segregation within all clusters; and (3) the amount of segregation that can be attributed to racial differences across clusters of different sizes.

In the first place, note that if we are merely interested in ranking clusters' segregation levels, the decomposability requirement is quite inessential. However, if the analysis involves comparisons between clusters and overall levels, then decomposability can be very useful indeed. As pointed out in the field of income inequality, a problem arises in the different interpretations that can be placed in statements like " x percent of overall segregation is attributed to cluster k " (Shorrocks 1980, 1984, 1988). Fortunately, *SSD* implies a satisfactory way of assigning segregation contributions to the clusters. For, it seems natural, when equation (1) holds for any partition of N schools into K clusters, to define the contribution to overall segregation of cluster k by

$$C_k = p_k S(X^k). \quad (2)$$

It is easy to check that this definition for C_k is consistent with the other two obvious interpretations of the sentence "contribution to segregation of cluster k ." First, consider the situation in which the original frequencies of students across races and schools in the city are replaced by one frequency in which all schools in cluster k are incorporated into a single school. Since in this case $S(\tilde{X}^k) = 0$, then from equation (1) we can immediately see that

$$C_k = S(X) - S(X^1 \cup \dots \cup X^{k-1} \cup \tilde{X}^k \cup X^{k+1} \cup \dots \cup X^K),$$

That is, the contribution C_k can also be interpreted as the amount by which overall segregation falls if the segregation within cluster k is eliminated. Second, consider the situation by which the original joint frequencies are replaced by one in which all clusters except k become single school clusters. Since in this situation $S(\tilde{X}^j) = 0$, for all $j \neq k$, it follows that

$$C_k = S(\tilde{X}^1 \cup \dots \cup \tilde{X}^{k-1} \cup X^k \cup \tilde{X}^{k+1} \cup \dots \cup \tilde{X}^K) - S(\tilde{X}^1 \cup \dots \cup \tilde{X}^K).$$

That is, C_k can also be interpreted as the amount by which overall segregation increases if segregation within cluster k is introduced starting from the position of zero segregation within each cluster. Therefore, under **SSD** it is possible to provide the same answer to different interpretations of what is meant by the contribution of each cluster to overall segregation. Consequently, the problem of unambiguously comparing individual clusters' contributions is solved. For example, the ratio $S(X^k)/S(X)$ is greater than, equal to, or smaller than one whenever cluster k 's contribution to the overall segregation level, $C_k/S(X)$, is greater than, equal to, or smaller than its demographic importance given by p_k .

In the second place, we must examine the contribution made to overall segregation by all clusters taken together, C . This question admits two sensible interpretations. First, a natural response is to compute the reduction in overall segregation that would arise if the segregation within all clusters were eliminated. In the partition into K clusters C will be

$$C = S(X) - S(\tilde{X}^1 \cup \dots \cup \tilde{X}^K).$$

A second interpretation would consist of the sum of the individual contributions defined in expression (2), that is,

$$\sum_{k=1}^K C_k \equiv \sum_{k=1}^K p_k S(X^k).$$

We can immediately see that for any segregation measure S satisfying **SSD**, $C = \sum_{k=1}^K C_k$ so that both interpretations provide the same answer.

Finally, consider the possibility of partitioning the set of schools in a country into clusters of different size, say regions, cities, or school districts. An empirical question must then be addressed 'How much segregation can be attributed to racial differences across regions as opposed to other geographical levels.' This may be interpreted two ways: (1) by how much segregation would fall if racial differences across clusters were the only source of school segregation, or (2) by how much segregation would fall if racial differences at the cluster level were eliminated. Interpretation (1) suggests a comparison of overall segregation with the amount that would arise if segregation within each of K clusters were made equal to zero but racial differences across districts remained the same. As shown earlier, for measures satisfying **SSD** this would eliminate the total within-groups term and leave only the between-groups contribution, so that $S(X) = S(\tilde{X}^1 \cup \dots \cup \tilde{X}^K)$. Interpretation (2) suggests a comparison of overall segregation with the segregation level that would result if all clusters had the same racial composition, equal to the one for the nation as a whole, but the segregation within each cluster remained unchanged. Unfortunately, in contrast to the situation for relative measures of income inequality, this conceptual experiment is not possible for measures of segregation, a difficulty that deserves an explanation.

For any partition of an income distribution, any decomposable inequality index allows expressing overall income inequality as the sum of a between- and a within-groups term, where the between-groups term is the inequality of the distribution where each individual is assigned the mean income of the subgroup to which she belongs. In this situation, starting from an income distribution x and a partition of the population

into subgroups, there is no difficulty in constructing a new income distribution y satisfying two conditions: (1) the mean income of any subgroup is equal to the mean income for the entire population, so that the between-groups inequality of distribution is equal to zero, and (2) income inequality within each subgroup is preserved. It is thus easy to see that the difference between income inequality in the initial situation, say $I(x) = B(x) + W(x)$, and income inequality in the second situation, $I(y) = B(y) + W(y) = 0 + W(x)$, is equal to the between-groups term:

$$I(x) - I(y) = B(x) + W(x) - W(x) = B(x).$$

That is, according to interpretation (2), between-groups income inequality is the amount by which overall income inequality is reduced when the differences between subgroup income means are eliminated by making them equal to the population income mean.⁵

The corresponding conceptual exercise in the segregation case is logically impossible. Starting from $X = X^1 \cup \dots \cup X^k \dots \cup X^K$, let us attempt to construct another city Y satisfying two conditions.

1. The racial composition of every cluster k in Y is the same as the one for the original population as a whole—that is, $p_{g|k} = p_g$ for all k and g , so that there is no between-groups segregation in Y . In this case, overall segregation in Y coincides with the within-groups term.
2. The level of segregation within each cluster remains as in the original city, so that the within-groups term in Y coincides with the one in X . Hence, overall segregation in Y coincides with within-groups segregation in X .

If this operation were possible, it is easy to see that, as in the income inequality case, the difference between overall segregation in X and in Y would be equal to the between-groups term. However, under condition (1) within-group segregation in Y results from the comparison

⁵ As a matter of fact, the answers to interpretations (1) and (2) coincide and are equal to the between-groups term only when the weights in the within-groups term do not depend on the subgroup means. This is only the case for one of the members of the entropy family of income inequality indicators: the mean logarithmic deviation (Shorrocks 1980).

between the racial distributions at school level with the racial distribution in the original city; but this comparison is what is involved in computing overall segregation in X . Therefore, within-groups segregation in Y is equal to overall segregation in the original city, which contradicts the fact that overall segregation in Y coincides with within-groups segregation in X . This contradiction arises because it is generally impossible in the segregation context to eliminate the between-groups segregation maintaining the existing within-groups segregation as the former affects the latter. Nevertheless, this does not preclude the investigation of the original question about which geographical level accounts for a greater percentage of overall segregation. For any segregation measure satisfying *SSD*, the size of the between-groups term at each geographical level provides a clear answer, if only in the sense of interpretation (1).

2.3. Strong Group Decomposability

In many research situations it is useful to partition demographic groups into supergroups. For example, we may want to assess the degree to which overall school segregation is due to segregation within a large supergroup consisting of all minority races in the US. Consider a partition of G groups in city X into $L < G$ supergroups, $X = X_1 \cup \dots \cup X_L$, where X_l is the set of groups that belongs to supergroup l . In addition, let \tilde{X}_l be the supergroup in which all groups in supergroup l have been combined into a single group with conditional school distribution $P_{n|l}$

Following Frankel and Volij (forthcoming), a school segregation index S is said to be *strongly group decomposable (SGD)*, if and only if for any partition $X = X_1 \cup \dots \cup X_L$ of the G groups into L supergroups overall segregation, $S(X)$, can be written as

$$S(X) = S(\tilde{X}_1 \cup \dots \cup \tilde{X}_L) + \sum_{l=1}^L p_l S(X_l). \quad (3)$$

Therefore, if a school segregation index is *SGD* then for any partition of the racial groups into supergroups, overall city segregation can be expressed as the sum of two terms, one that captures between-supergroups segregation, and another that captures within-supergroups segregation and is equal to the weighted average of segregation within each of the

supergroups, with weights equal to the supergroups' demographic importance.

This definition also implies a satisfactory way of assigning segregation contributions to the supergroups. For, when equation (3) holds, the definition $C_l = p_l S(X_l)$ is consistent with all the obvious interpretations of the concept "contribution to segregation by supergroup l ": the amount by which overall segregation falls if the segregation within supergroup l is eliminated, or the amount by which overall segregation increases if segregation within supergroup l is introduced starting from the position of zero segregation within each supergroup. Reflecting a similarity with the case of the partition of schools into clusters, an index satisfying *SGD* provides a satisfactory answer to the question of how much segregation would fall if school differences across supergroups were the only source of segregation. However, it is logically impossible to eliminate the between-supergroups segregation maintaining the existing within-supergroups segregation as the latter is affected by the former.

3. ENTROPY-BASED SEGREGATION INDICES

3.1. Preliminaries

Before we present the entropy-based indices of segregation, the concept of entropy of a distribution must be introduced. Consider a discrete random variable x that takes Q probability values, indexed by $q = 1, \dots, Q$. Let p_q be the probability of the q th value with $p_q \geq 0$ and $\sum_{q=1}^Q p_q = 1$. For instance, if x is the ethnic group of a randomly selected student, then p_q is the proportion of students in the city who are in the q th group. The entropy of the Q values of variable x is the real value function defined as

$$E(P_q) = - \sum_{q=1}^Q p_q \log(p_q) = \sum_{q=1}^Q p_q \log \left(\frac{1}{p_q} \right),$$

with $0 \log(1/0) = 0$.⁶ Heuristically, the information brought about by observing the actual value of x is the opposite of the logarithm of its

⁶ The base of the logarithm is irrelevant, providing essentially a unit of measure. In this paper the natural logarithm will be used.

likelihood, $-\log(p_q) = \log(1/p_q)$: the observation of an unlikely value brings about a large amount of information once observed. Therefore, the entropy is a measure of the expected information for the value of variable x brought about by an observation.

3.2. The M Index

The M index is defined as follows. Suppose that a student is drawn randomly from the city, so that the expected information of learning her race is measured by the entropy of the city's ethnic distribution, $E(P_g)$. If we were informed about the school the student attends, the expected information from learning her race would now be measured by the entropy of her school's ethnic distribution, $E(P_{g|n})$. If the schools in the city are all segregated, then the latter entropy will tend to be lower because the student's school conveys some information about her race. The M index equals this change in entropy, $E(P_{g|n}) - E(P_g)$, averaged over the students in the city:

$$M = \sum_{n=1}^N p_n (E(P_g) - E(P_{g|n})). \quad (4)$$

The M index thus captures segregation viewed as the extent to which schools have different racial compositions from the population as a whole. This notion of segregation corresponds to differences in the column percentages in city X .

Note that $p_{g|n} p_n = p_{n|g} p_g$ so that $\log(p_g) - \log(p_{g|n}) = \log(p_n) - \log(p_{n|g})$: The information obtained about race from learning about the school the student attends equals the information gained about the school the student attends when learning about her race. Hence, the M index also equals the reduction in uncertainty about a student's school that comes from learning her race:

$$M = \sum_{g=1}^G p_g (E(P_n) - E(P_{n|g})). \quad (5)$$

Therefore, the M index also captures segregation as the tendency of racial groups to have different distributions across schools, or the differences in row percentages in X .

3.3. The Normalized Entropy-based Indices

It can be shown from equation (4) that $M \in [0, \log G]$. In particular, M takes its minimum value whenever the racial entropy in each school coincides with the racial entropy in the city, $E(P_{g|n}) = E(P_g)$, $n = 1, \dots, N$. This situation arises only when the racial distribution of each school equals the racial distribution of the city, in which case it is said that the city is completely integrated. The M index reaches its maximum value when the racial groups are uniformly distributed in the city and there is no ethnic mix within schools. In other words, according to the M index complete segregation requires two conditions: there must be no racial mix within schools, *and* races must be uniformly distributed in the city. For any given racial marginal distribution P_g , M attains its maximum at the city's racial entropy, $E(P_g)$. This fact suggests normalizing M by $E(P_g)$:

$$H \equiv \frac{M}{E(P_g)} = \sum_{n=1}^N p_n \left(\frac{E(P_g) - E(P_{g|n})}{E(P_g)} \right). \quad (6)$$

Therefore, the H index measures the proportional increase in expected information about race that occurs when learning about the school that the student attends. Consequently, H captures segregation as relative differences in the column percentages in city X . As with M , there is complete integration whenever the racial distribution of each school equals the racial distribution of the city. However, in contrast to M , H reaches its maximum value whenever there is no racial mix within schools, thus providing a characterization of complete segregation that is independent of the racial distribution in the city. Although H is neither *I1* nor *I2*, this characterization of complete segregation coincides with the one provided by any *I1* or *I2* index that satisfies the principle of transfers.⁷

It can be shown from equation (5) that, as a function of the school entropies by racial group, M reaches its minimum value, 0, whenever the

⁷ The *principle of transfers*, first proposed by James and Taeuber (1985) for segregation studies, states that segregation must decrease if a student of a given group moves from a school where her group's proportion is above that in the city as a whole to a school where her group's proportion is below that in the city as a whole.

school entropy is the same for all racial groups, $E(P_{n|g}) = E(P_n)$, $g = 1, \dots, G$, while it reaches its maximum value, $\log N$, when the schools are evenly distributed in the city and each racial group attends a disjoint set of schools. Thus, the notion of complete segregation as departure from row percentages for M also demands two conditions: In addition to requiring no racial mix within organizational units, schools must be uniformly distributed at the city level. For any given school distribution P_n , M attains its maximum at the schools entropy at the city level, $E(P_n)$. This fact suggests normalizing M by $E(P_n)$:

$$H^* \equiv \frac{M}{E(P_n)} = \sum_{g=1}^G p_g \left(\frac{E(P_n) - E(P_{n|g})}{E(P_n)} \right) = \frac{E(P_g)}{E(P_n)} H. \quad (7)$$

The H^* index has not been defined previously, although it is closely related to both M and H . Intuitively, it captures the proportional expected increased in the information about the school when learning about the race of a student. Consequently, in contrast to H , H^* captures segregation as differences in the row percentages in city X . As with M and H , there is complete integration whenever the racial distribution of each school equals the racial distribution of the city. As with H , it can only take values within the unit interval, and it reaches the unity whenever there is no racial mix within schools. Finally, although H^* is neither $I1$ nor $I2$, this characterization of complete segregation coincides with the one provided by any $I1$ or $I2$ index that satisfies the principle of transfers.

4. DECOMPOSABILITY PROPERTIES OF THE ENTROPY-BASED INDICES

4.1. *Decomposability Properties of the M Index*

It is easy to show that the M index satisfies both **SSD** and **SGD** in the multigroup case. First, equation (1) takes the form

$$M = M^B + \sum_{k=1}^K p_k M_k^W, \quad (8)$$

where

$$M^B = \sum_{k=1}^K p_k(E(P_g) - E(P_{g|k})) = \sum_{g=1}^G p_g(E(P_k) - E(P_{k|g}))$$

is the between-groups term that captures what we will refer to as cluster segregation, and

$$M_k^W = \sum_{n \in X^k} p_n(E(P_{g|k}) - E(P_{g|n \in X^k})) = \sum_{g=1}^G p_{g|k}(E(P_{n|n \in X^k}) - E(P_{n|g, n \in X^k}))$$

captures school segregation within cluster k . Given that the M index satisfies **SSD**, the contribution $CM_k^W = p_k M_k^W$ is consistent with all the obvious interpretations of the concept “contribution to segregation by cluster k .” Similarly, M admits the decomposition

$$M = M_B + \sum_{l=1}^L p_l M_l^W, \quad (9)$$

where

$$M_B = \sum_{n=1}^N p_n(E(P_l) - E(P_{l|n})) = \sum_{l=1}^L p_l(E(P_n) - E(P_{n|l}))$$

is the between-groups term that captures school segregation by supergroup, and

$$\begin{aligned} M_l^W &= \sum_{n=1}^N p_n(E(P_{g|g \in X_l}) - E(P_{g|n, g \in X_l})) \\ &= \sum_{g \in X_l} p_{g|g \in X_l}(E(P_{n|g \in X_l}) - E(P_{n|g, g \in X_l})) \end{aligned}$$

captures school segregation within supergroup l . Given that the M index satisfies **SGD**, the contribution $CM_l^W = p_l M_l^W$ is consistent with all the obvious interpretations of the concept “contribution to segregation by supergroup l .”

4.2. Weaker Decomposability Properties

Although the H and H^* indices violate **SSD** and **SGD**, it can be seen that they satisfy some weaker decomposability properties. First, consider any partition of the N schools into $K < N$ clusters, and recall that H can be computed by dividing the M index by the racial entropy, $E(P_g)$. On the one hand, starting from the definition of M in equation (5) and decomposition (8) we have

$$H = \frac{M^B}{E(P_g)} + \sum_{k=1}^K p_k \frac{M_k^W}{E(P_g)}.$$

Multiplying and dividing each summand of the second term by the within-group's racial entropy, $E(P_{g|k})$, and using the relation between the un-normalized and the normalized indexes, we have

$$H = H^B + \sum_{k=1}^K p_k \frac{E(P_{g|k})}{E(P_g)} H_k^W, \quad (10)$$

where H^B captures cluster segregation, and H_k^W captures school segregation within cluster k . On the other hand, starting from the definition of M in equation (4) and decomposition (8), for the H^* index we have

$$H^* = \frac{M^B}{E(P_n)} + \sum_{k=1}^K p_k \frac{M_k^W}{E(P_n)}.$$

Multiplying and dividing the between-groups term by $E(P_k)$ and each summand of the second term by $E(P_{n|k})$, we have

$$H^* = \frac{E(P_k)}{E(P_n)} H^{*B} + \sum_{k=1}^K p_k \frac{E(P_{n|k})}{E(P_n)} H_k^{*W}, \quad (11)$$

where H^{*B} captures cluster segregation and H_k^{*W} captures school segregation within cluster k .

Second, consider any partition of the G groups into $L < G$ supergroups. Starting from equations (5) and (9), for the H index we have

$$H = \frac{M_B}{E(P_g)} + \sum_{l=1}^L p_l \frac{M_l^W}{E(P_g)}.$$

Multiplying and dividing the between-groups term by $E(P_l)$ and each summand of the second term by $E(P_{g|l})$, we have

$$H = \frac{E(P_l)}{E(P_g)} H_B + \sum_{l=1}^L p_l \frac{E(P_{g|l})}{E(P_g)} H_l^W \quad (12)$$

where H_B captures school segregation by supergroup, and H_l^W captures school segregation within supergroup l .⁸ Finally, starting from equations (4) and (9), we have

$$H^* = \frac{M_B}{E(P_n)} + \sum_{l=1}^L p_l \frac{M_l^W}{E(P_n)}.$$

Multiplying and dividing each summand of the second term by $E(P_{n|l})$, we have

$$H^* = H_B^* + \sum_{l=1}^L p_l \frac{E(P_{n|l})}{E(P_n)} H_l^{*W}, \quad (13)$$

where H_B^* captures school segregation by supergroup, and H_l^{*W} captures school segregation within supergroup l .

4.3. *Ambiguities in the Interpretation of the Contributions to Segregation*

It should be noted at the outset that the contributions of the between-groups and within-groups terms expressed as a percentage of the H and the H^* indices in expressions (10)–(11) and (12)–(13) pose no problem because they coincide with those same relative contributions for the M index in expressions (8) and (9), respectively. Thus, for example, in the case of decomposition (10) we have

⁸ Equation (12) figures prominently in Reardon et al. (2000); see their equation (4).

$$\begin{aligned}
& \frac{H^B}{H} + \sum_{k=1}^K \left(p_k \frac{E(P_{g|k})}{E(P_g)} \right) \left(\frac{H_k^W}{H} \right) \\
&= \frac{M^B/E(P_g)}{M/E(P_g)} + \sum_{k=1}^K \left(p_k \frac{E(P_{g|k})}{E(P_g)} \right) \left(\frac{M_k^W/E(P_{g|k})}{M/E(P_g)} \right), \\
&= \frac{M^B}{M} + \sum_{k=1}^K p_k \frac{M_k^W}{M}.
\end{aligned}$$

Similarly, for decomposition (11) we have

$$\begin{aligned}
& \left(\frac{E(P_k)}{E(P_n)} \right) \left(\frac{H^{*B}}{H^*} \right) + \sum_{k=1}^K \left(p_k \frac{E(P_{n|k})}{E(P_n)} \right) \left(\frac{H_k^{*W}}{H^*} \right), \\
&= \left(\frac{E(P_k)}{E(P_n)} \right) \left(\frac{M^B/E(P_k)}{M/E(P_n)} \right) + \sum_{k=1}^K \left(p_k \frac{E(P_{n|k})}{E(P_n)} \right) \left(\frac{M_k^W/E(P_{n|k})}{M/E(P_n)} \right), \\
&= \frac{M^B}{M} + \sum_{k=1}^K p_k \frac{M_k^W}{M}.
\end{aligned}$$

It is important to recognize, however, that the terms in decompositions (10) and (13) admit the same interpretations as those terms in any **SSD** and **SGD** index. Let's first, define cluster k 's contribution to overall segregation as $CH_k^W = p_k \frac{E(P_{g|k})}{E(P_g)} H_k^W$. It is easy to show that CH_k^W can be interpreted both as the amount by which overall segregation falls if the segregation within cluster k is eliminated, and the amount by which overall segregation increases if segregation within cluster k is introduced starting from the position of zero segregation within each cluster. Likewise, we can define the contribution of all clusters to segregation as $CH^W = \sum_{k=1}^K CH_k^W$. It turns out that CH^W equals the reduction in segregation that would arise if the segregation within all clusters were eliminated. Finally, the interpretation of the between-groups term in decomposition (10), H^B , is subject to the same conceptual limitation pointed out earlier in Section 3.1 in relation to the decomposition of any **SSD** index. Namely, H^B can be interpreted as the level of segregation if racial differences across clusters were the only source of school segregation so that $H_k^W = 0$ for all $k = 1, \dots, K$. However, it cannot be interpreted as a decrease in segregation if racial

differences at the cluster level were eliminated. For reasons of brevity, the properties of decomposition (13) are not discussed in detail. Nevertheless, similar arguments to those provided for decomposition (10) can be used to show that the terms in decomposition (13) can be interpreted as those in the decomposition of any **SGD** index for any partition of ethnic groups into supergroups.

However, as discussed in the introduction, decompositions (11) and (12) present serious problems of interpretation. Example 1 in the next paragraph illustrates that equation (12) does not provide the H index with a decomposition that admits the same interpretation as that of any **SGD** index. It first shows that the contribution of supergroup l to overall segregation, $CH_l^W = p_l \frac{E(P_{gl})}{E(P_g)} H_l^W$, cannot generally be interpreted as the amount by which overall segregation falls if the segregation within supergroup l is eliminated. The reason is that in this case the overall racial entropy $E(P_g)$ will usually change, and this may induce changes in the weights of the contributions by other supergroups. The example also shows that the term CH_l^W cannot always be interpreted as the amount by which overall segregation increases if segregation within supergroup l is introduced starting from the position of zero segregation within each racial supergroup. Finally, it becomes clear that $CH_B = \frac{E(P)}{E(P_g)} H_B$ cannot be interpreted as the level of segregation if differences in the supergroup distributions across schools were the only source of school segregation.

Example 1. Consider two cities, X and Y , with students from three racial groups, *white*, *Asian*, and *black*, and two schools, $s1$ and $s2$. The joint frequencies of students across schools and racial groups can be summarized in two matrices:

$$\begin{array}{c}
 \begin{array}{cc}
 \begin{array}{cc}
 X = \begin{bmatrix} 7 & 38 \\ 3 & 2 \\ 20 & 5 \end{bmatrix} & \text{and} & Y = \begin{bmatrix} 7 & 28 \\ 3 & 12 \\ 20 & 5 \end{bmatrix} & \begin{bmatrix} \text{white} \\ \text{Asian} \\ \text{black} \end{bmatrix} \\
 \begin{bmatrix} s1 & s2 \end{bmatrix} & & \begin{bmatrix} s1 & s2 \end{bmatrix} \\
 \text{Schools} & & \text{Schools}
 \end{array}
 \end{array}
 \end{array}$$

Suppose that we group together white and Asian students, referring to the resulting supergroup as wa . To begin with, according to index H

school segregation within supergroup wa is zero in city Y , but positive in X ,

$$H_{wa}^W(X) = 100 * \frac{M_{wa}^W(X)}{E(P_{g|wa})} = 100 * \frac{4.41}{32.51} = 13.57.^9$$

However, the contribution of within-supergroups segregation in city X ,

$$CH_{wa}^W(X) = p_l \frac{E(P_{g|l})}{E(P_g)} H_{wa}^W(X) = 0.67 \frac{32.51}{85.32} 13.57 = 3.45,$$

is not equal to the fall in overall segregation when eliminating segregation within supergroup wa —that is, moving from city X to city Y , $H(Y) - H(X) = -7.14$. The reason is that the overall racial entropy has increased: $E(P_g(Y)) = 104.38$ versus $E(P_g(X)) = 85.32$. It is clear that $CH_{wa}^W(X) = 3.45$ does not equal the amount by which overall segregation increases if segregation within supergroup l is introduced starting from the position of zero segregation within each racial supergroup—that is, moving from city Y to city X , $H(X) - H(Y) = 7.14$. Finally, the term

$$CH_B(X) = \frac{E(P_l)}{E(P_g)} H_B(X) = \frac{63.65}{85.32} 27.12 = 20.23$$

does not equal the level of segregation if differences in the supergroup distributions across schools were the only source of school segregation, $H(Y) = \frac{0.6365}{1.0438} 27.12 = 16.54$.¹⁰

4.4. *Additional Problems of Interpretation Due to the Nature of the Weights*

All weights in decompositions (10) to (13) are not invariant to changes in the within-groups distributions, leading to several problems of

⁹ All entropy and index calculations reported hereafter are computed using natural logarithms and are multiplied by 100.

¹⁰ Note that the contributions of the between- and the within-supergroups terms expressed as a percentage of the H (H^*) indices in expressions (11) and (12) pose no interpretability problem because they coincide with those same relative contributions for the M index.

interpretation. Consider decomposition (12) for H . The nature of the weights $\frac{E(P_l)}{E(P_g)}$ and $p_l \frac{E(P_{gl})}{E(P_g)}$ leads to two problems. (1), We may have two cities with the same H_B but different contribution $CH_B = \frac{E(P_l)}{E(P_g)} H_B$ to overall segregation due to differences in the entropy ratio $\frac{E(P_l)}{E(P_g)}$ and (2), for a given joint distribution of supergroups and schools, P_{ln} , the weights $p_l \frac{E(P_{gl})}{E(P_g)}$ generally change in response to exogenous changes in the joint distribution of groups and schools within supergroups. Thus, although supergroup demographic shares, p_l , remain constant, the overall racial entropy at group level $E(P_g)$ or the racial entropy at group level in supergroup l , $E(P_{gl})$, may change. Consequently, the contribution to within-groups segregation, $CH_W = \sum_{l=1}^L CH_l^W$, may change in a direction contrary to what the terms H_l^W would indicate. Both problems are illustrated in the example that follows.

Example 2. Consider two cities, X and Y , with students from four racial groups (*white, Asian, black, and Hispanic*) and two schools ($s1$ and $s2$). The relative frequencies (expressed as a percentage) of students across schools and racial groups can be summarized in two matrices:

$$\begin{array}{c}
 \begin{array}{cc}
 \begin{array}{cc}
 \begin{array}{c} X = \begin{bmatrix} 9 & 36 \\ 3 & 2 \\ 20 & 5 \\ 20 & 5 \end{bmatrix} \\ \begin{bmatrix} s1 & s2 \end{bmatrix} \\ \text{Schools}
 \end{array}
 \end{array}
 \end{array}
 \end{array}
 \text{ and }
 \begin{array}{c}
 \begin{array}{cc}
 \begin{array}{cc}
 \begin{array}{c} Y = \begin{bmatrix} 9.05 & 35.95 \\ 2.95 & 2.05 \\ 36 & 9 \\ 4 & 1 \end{bmatrix} \\ \begin{bmatrix} s1 & s2 \end{bmatrix} \\ \text{Schools}
 \end{array}
 \end{array}
 \end{array}
 \end{array}
 \begin{array}{c}
 \text{Ethnic groups} \\
 \begin{bmatrix} \text{white} \\ \text{Asian} \\ \text{black} \\ \text{Hispanic} \end{bmatrix}
 \end{array}
 \end{array}
 \end{array}
 \end{array}
 \end{array}$$

Suppose that we group together, on the one hand, white and Asian students, referring to the resulting supergroup as wa , and, on the other hand, black and Hispanic students, referring to the resulting supergroup as bh . There are two points to note here. First, the joint distribution of supergroups and schools is the same in both cities X and Y and, consequently, so is the value for school segregation by supergroup, $H_B(X) = H_B(Y) = 24.03$. However, the contribution of between-groups segregation to overall segregation, CH_B , is larger in Y than in X ($CH_B(Y) = \frac{69.31}{101.82} 24.03 = 16.36$ versus $CH_B(X) = \frac{69.31}{120.23} 24.03 = 13.86$)

simply because the entropy ratio is larger there. Second, measured by H_l^W , supergroup *wa* experiences slightly more school segregation in X than in Y ($H_{wa}^W(X) = 10.28$ versus $H_{wa}^W(Y) = 9.74$), while supergroup *bh* has no school segregation in both cities ($H_{bh}^W(X) = H_{bh}^W(Y) = 0$). Since the difference in the shares of black and Hispanic students is much smaller in X than in Y , both the overall racial entropy and the racial entropy within supergroup *bh* are larger in X than in Y : $E(P_g(X)) = 120.23$ versus $E(P_g(Y)) = 101.82$, and $E(P_{g|bh}(X)) = 34.66$ versus $E(P_{g|bh}(Y)) = 9.48$. As a result, even though the joint frequency of supergroups and schools is the same for both cities, the weights $p_l \frac{E(P_{g|l})}{E(P_g)}$ are so much larger in city Y —the city with less segregation within supergroup *wa*—that the contribution of within-groups segregation is also larger there:

$$CH_W(Y) = 0.50 \frac{32.50}{101.82} 9.73 = 1.55 \text{ versus}$$

$$CH_W(X) = 0.50 \frac{32.51}{120.23} 10.28 = 1.39.$$

Decomposition (10) for H presents analogous problems of interpretation for the within-groups term as $CH^W = \sum_{k=1}^K p_k \frac{E(P_{g|k})}{E(P_g)} H_k^W$ may change in a direction contrary to what the terms H_k^W would indicate. Also, the decompositions (11) and (13) for H^* have similar problems of interpretation. For decomposition (11), we may have two cities with the same between-groups segregation, H^{*B} , but different contributions to overall segregation due to differences in the entropy ratio $\frac{E(P_k)}{E(P_n)}$. Finally, the contributions to within-groups segregation,

$$CH^{*W} = \sum_{k=1}^K p_k \frac{E(P_{n|k})}{E(P_n)} H_k^{*W}, \text{ and } CH_W^* = \sum_{l=1}^L p_l \frac{E(P_{n|l})}{E(P_n)} H_l^{*W},$$

may change in a direction contrary to what the terms H_k^{*W} and H_l^{*W} would indicate, respectively.¹¹

¹¹ For the sake of brevity, proofs of the statements in this paragraph using illustrative examples will be available only upon request.

TABLE 1
School Enrollment, Ethnic Mix, Entropies, and School Segregation in the United States, 1989–2005

	Number of Students (millions)			Racial Shares (%)			
	1989	2005	Change (%)	1989	2005	Change	
Minorities	8.61	12.24	42.10	34.78	48.05	13.27	
Native American	0.17	0.23	33.77	0.68	0.89	0.20	
Asian	1.03	1.40	36.11	4.15	5.49	1.34	
Black	3.99	4.53	13.70	16.10	17.80	1.70	
Hispanic	3.43	6.08	77.33	13.85	23.87	10.02	
White	16.14	13.23	−18.06	65.22	51.95	−13.27	
Total	24.76	25.47	2.87	100	100	0	
Entropies and Segregation Indexes							
	$E(P_g)$	$E(P_n)$	$\sum_{n=1}^N p_n E(P_{g n})$	$\sum_{g=1}^G p_g E(P_{n g})$	M	H	H^*
1989	101.27	1040.25	57.35	996.32	43.92	43.37	4.22
2005	119.07	1035.72	70.17	986.82	48.90	41.07	4.72
Change	17.80	−4.53	12.82	−9.50	4.98	−2.30	0.50

Notes: Ethnic shares are the percentages of students from every race/ethnic group. The terms Native American, Asian, black, and white refer to non-Hispanic members of these racial groups; Asian includes Native Hawaiians and Pacific Islanders; Native American includes American Indians and Alaska Natives (Inuit or Aleut). The term Hispanic is an ethnic rather than a racial category since Hispanic persons may belong to any race. Minorities include all categories except white.

4.5. Decomposability Properties in Practice: The M versus the H Index

It will be illustrative to see how the decomposability properties of the M and the H indices fare in practice with data about the evolution of the U.S. student population enrolled in public schools in Core-Based Statistical Areas (CBSAs)—urban clusters of 10,000 or more inhabitants, referred to in the sequel as *cities*—during the 1989–1990 and 2005–2006 academic years.¹² Table 1 clarifies two issues. First, the evolution of the

¹² Results pertain to those schools for which racial and ethnic information is available both in 1989 and in 2005. Given that a small proportion of schools did not report results in 1989, focusing on the schools that did probably gives a fairer comparison between the distributions observed in 1989 and in 2005 because it does not include those schools that did report in 2005 but failed to do so in 1989.

ethnic diversity of the student population is shown. Minorities (namely, Native Americans, blacks, Asians, and Hispanics) already represent 34.8% of the total population of 24.8 million in 1989. Since all of them grew more rapidly than whites during this period, they represent as much as 48.1% of the total population of 25.5 million in 2005. Second, the segregation levels achieved by the different entropy indices are shown. In particular, the change in the M index during this period is $\Delta M = 48.90 - 43.92 = 4.98$. Suppose that we group together Asian, black, Hispanic, and Native American students, referring to the resulting “minorities” supergroup as m . Consider now the evolution of segregation between whites versus minorities and the evolution of segregation within minorities. Since only one supergroup is considered, equation (9) simplifies to $M = M_B + p_m M_m^W$, where p_m denotes the share of minorities in the student population, M_m^W is the M index within minorities, and M_B is the M index of school segregation for whites versus all minorities combined. The observed increase in overall segregation is due primarily to the increase in M_B , which becomes $\Delta M_B = 1.83$. In addition, the share of the minorities (who are highly segregated among themselves) increases substantially to $\Delta p_m = 0.13$. Thus, in spite of the fact that school segregation within minorities is decreasing, $\Delta M_m^W = -8.25$, the contribution of segregation within minorities to overall segregation is positive, $\Delta C M_m^W = 3.15$. Consequently, $\Delta M = 1.83 + 3.15 = 4.98$.

Given equation (2), we can see that H decreases because the racial entropy is increasing ($119.07 - 101.27 = 17.80$) faster than M :

$$\begin{aligned}\Delta H &= (48.90/119.07) - (43.92/101.27), \\ &= 41.07 - 43.37 = -2.30.\end{aligned}$$

But how does H account for the trends in the minorities’ partition? Note that, with only one supergroup, decomposition (12) simplifies to

$$H = \frac{E(P_l)}{E(P_g)} H_B + p_m \frac{E(P_{g|m})}{E(P_g)} H_m^W.$$

However, interpretability of the results presented here is potentially compromised by the fact that some schools have been created while others have disappeared between 1989 and 2005. Nevertheless, results using all observations are qualitatively similar, suggesting that the selection mechanisms at work are not essential to our analysis. Results obtained using the full sample are available upon request.

The H index also finds a decrease in segregation within minorities, $\Delta H_m^W = -7.13$, and a very small increase in school segregation between whites and minorities, $\Delta H_B = 0.03$. In spite of the increasing importance of minorities in the student population, the within-minorities weight increases only slightly (from 0.36 to 0.42) as a combined result of the decrease in the racial entropy within minorities (from 105.40 in 1989 to 103.71 in 2005), together with the increase in the overall racial entropy (from 101.27 to 119.07). The small increase in the weight does not offset the large decrease in segregation within minorities, and, hence, the contribution of segregation within minorities to overall segregation is negative, $\Delta CH_m^W = -0.11$. Moreover, the contribution of between-groups segregation is also affected by the evolution of the ratio $\frac{E(P)}{E(P_g)}$. It turns out that simply because the racial entropy is growing relatively more than the supergroup entropy between whites and minorities, most of the reported decrease in the entropy index, $\Delta H = -2.30$, stems from the decrease in the contribution of the between-groups term, $\Delta CH_B = -2.19$, in spite of the reported increase in H_B .

5. INVARIANCE PROPERTIES

5.1. *The Invariance Question*

Consider for a moment the special but important case of occupational segregation by gender, and assume that segregation data in 1950 and 2000 are being compared in a given country. Several questions are often asked. First, should the measurement of occupational segregation be independent of the fact that female labor participation has greatly increased over time? Many people would agree that, as long as the male and female distributions over occupations remain constant, the degree of segregation should be the same in the two situations—that is, that an index of occupational segregation by gender should satisfy *II*. In the school segregation case with several racial groups, the question becomes whether segregation should be invariant to changes in the ethnic composition of the population as long as the distribution of each group within schools remains constant. Second, should occupational segregation be independent from the fact that agricultural and industrial occupations are much more important in 1950 than in 2000, while service occupations carry much more weight in 2000 than in 1950?

Many people would agree that, as long as the gender composition of each occupation remains constant, the degree of segregation should be the same in the two situations—that is, an index of occupational segregation should be *I2*. In the school segregation case with several racial groups, the question becomes whether segregation should be invariant to changes in the size distribution of schools as long as the racial composition of each school remains constant.

As indicated in the introduction to this paper, the three entropy-based measures M , H , and H^* violate both properties—that is, they mix up segregation changes with changes in the marginal distributions in segregation comparisons over time or across space. However, Mora and Ruiz-Castillo (2009) present two decompositions of the M index in pairwise comparisons over time or across space that isolate the effects of the changes in the marginal distributions. In the first place, to identify an *I1* term in a decomposition of a pairwise comparison, the differences in the M index can be written as

$$\Delta M = \Delta \text{Net}(I1) + \Delta M(P_g) + \Delta E(P_n), \quad (14)$$

where $\Delta E(P_n)$ is the change in the school entropy, $\Delta M(P_g)$ isolates changes in M due to changes in the racial marginal distribution, P_g , while $\Delta \text{Net}(I1)$ is an *I1* term in the sense that it equals zero as long as $P_{n|g}$ remains constant. The term $\Delta \text{Net}(I1)$ is referred to in the discussion that follows as changes in net segregation viewed as differences in rows. In the second place, to identify an *I2* term in a decomposition of a pairwise comparison, the differences in the M index can be written as

$$\Delta M = \Delta \text{Net}(I2) + \Delta M(P_n) + \Delta E(P_g), \quad (15)$$

where $\Delta M(P_n)$ isolates changes in M due to changes in P_n , $\Delta E(P_g)$ is the change in the racial entropy, and $\Delta \text{Net}(I2)$ is an *I2* term in the sense that it equals zero as long as $P_{g|n}$ remains constant. In the discussion that follows, the term $\Delta \text{Net}(I2)$ is referred to as changes in net segregation viewed as differences in columns.

Decompositions (14) and (15) are not available for the H and H^* indexes. However, it is sometimes argued that since normalization makes complete segregation as defined in H independent of P_g , then the notion of segregation captured by H “is independent of the population’s diversity” (e.g., see Reardon et al. 2000:354). Clearly, H is neither

I1 nor *I2*, but to what extent does H reduce the invariance problems in M ? Taking into account equation (14) and the linear approximation to changes in H , $\Delta H \simeq \frac{1}{E(P_g)}(\Delta M - \Delta E(P_g))$, it is obvious that as long as $\Delta M(P_g) \simeq 0$ and $E(P_g) \simeq 1$, then $\Delta H \simeq \Delta \text{Net}$ (*I2*). However, it will presently be seen that changes in H can be a very inadequate approximation to isolate *I2* changes in $P_{g|n}$. First, by means of a numerical example it will be shown that changes in H (and also changes in H^*) may be unduly influenced by changes in P_g and in P_n when the racial and school entropies do not change. Second, in the case of the evolution of the U.S. student population enrolled in public schools, it will be seen how a large increase in the racial entropy coupled with a relatively smaller change in the school marginal distribution leads both to H greatly undervaluing the reductions in net segregation as differences in columns and H^* missing the reductions in net segregation as differences in rows.

5.2. Changes in the Marginal Distributions Without Changes in the Entropies

The next example illustrates how neither H nor H^* correct for the lack of invariance in M if the marginal distributions of schools and races change but the entropies do not.

Example 3. Consider two cities, X and Y , with students from three racial groups, *white*, *black*, and *Hispanic*, and three schools, $s1$, $s2$, and $s3$. The joint absolute frequencies of students across schools and racial groups are summarized in two matrices:

$$\begin{array}{c}
 X = \begin{bmatrix} 30 & 10 & 5 \\ 5 & 15 & 5 \\ 5 & 10 & 15 \\ [s1 & s2 & s3] \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 10 & 10 & 10 \\ 5 & 15 & 25 \\ 10 & 10 & 5 \\ [s1 & s2 & s3] \end{bmatrix} \begin{array}{c} \text{Ethnic groups} \\ \begin{bmatrix} \text{white} \\ \text{black} \\ \text{Hispanic} \end{bmatrix} \end{array} \\
 \text{Schools} \qquad \qquad \qquad \text{Schools}
 \end{array}$$

City X is predominantly white, while city Y is predominantly black. Hispanics are the second largest group in X and the smallest group

in Y . However, racial entropies in both cities (multiplied by 100) are the same: $E(P_g(X)) = E(P_g(Y)) = 106.71$. School 1 is the largest and school 3 the smallest in city X , while the order is reversed in city Y . However, these changes in the school marginal distribution do not affect the school entropy (multiplied by 100): $E(P_n(X)) = E(P_n(Y)) = 108.05$. Moreover, both the school entropy and the racial entropy are close to 1. Consequently, changes in H and H^* are very similar to changes in M : $M(X) - M(Y) = 6.56$ versus $H(X) - H(Y) = 6.15$ versus $H^*(X) - H^*(Y) = 6.07$. However, according to decomposition (14), net segregation as differences in rows is lower in X than in Y , $\Delta\text{Net}(II) = -7.98$, and the change in the racial distribution increases M in X , $\Delta M(P_g) = 17.19$. Similarly, according to decomposition (15), net segregation as deviations in columns is lower in X than in Y , $\Delta\text{Net}(I2) = -5.98$, and the change in the school distribution increases segregation in X , $\Delta M(P_n) = 12.54$. Hence, neither H nor H^* correct for the lack of invariance in M if the marginal distributions of schools and races change but the entropies do not.

5.3. *The Effects of an Increase in the Racial Entropy: Invariance Properties in Practice*

The case of the evolution of the U.S. student population enrolled in public schools already studied in Section 3.2 is reconsidered here to evaluate whether, in practice, changes in either the H or the H^* index can be seen as reasonable approximations of $I2$ or II terms, respectively. In Section 3.2 it was reported that during the 1989–2005 period the M index increased by 4.98, the H index decreased by -2.30 because the racial entropy increased relatively more than M , and the H^* index slightly increased by 0.50 because the school entropy decreased. However, in equation (15) the change in the M index due to the change in the racial entropy is 17.80, while the change due to the change in the marginal distribution of schools is -0.59 . Therefore, the variation in net segregation independent of these effects is

$$\Delta\text{Net}(I2) = 4.98 - (-0.59) - 17.80 = -12.23.$$

Hence, the change in the normalized entropy index H greatly undervalues the improvement in net segregation as differences in columns.

In contrast, the change in the M index in equation (13) due to the change in the schools' entropy is -4.53 , while the change due to the change in the marginal distribution of racial groups is 10.63 . Therefore, the change in net segregation independent of these effects is

$$\Delta \text{Net}(II) = 4.98 - (-4.53) - 10.63 = -1.11.$$

Hence, the change in the normalized entropy index H^* misses the improvement in net segregation as differences in rows.

6. THE NORMALIZATION ISSUE

Clearly, it is convenient for any index to be normalized in the sense that it reaches a maximum value for a particular notion of complete segregation and a minimum value for a particular notion of complete integration. Most researchers would identify the absence of segregation with the situation where organizational units have the same racial composition or, equivalently, where demographic groups have the same distribution across organizational units. Similarly, most researchers would accept that demographic groups are completely segregated whenever they do not mix at all within organizational units. A segregation index is said to be *normalized in the unit interval*—or to possess the *NOR* property—if it takes value 0 whenever there is no segregation and it takes value 1 whenever it reaches complete segregation as defined above.

It has been shown that while H and H^* satisfy *NOR*, the M index does not because it requires an additional condition to reach maximum segregation. However, there are conceptual reasons to defend the notion of complete segregation implicit in M . Both H and H^* rank all cities with no racial mixing within schools as equally segregated, while M assigns a higher segregation level to cities in which there is less initial expected information about a student's racial group. Following an example for another purpose in Frankel and Volij (forthcoming), consider city A with three schools and three racial groups and city B with two schools and two racial groups, such that

$$A = \begin{bmatrix} 50 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 50 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}.$$

Given each city's marginal distributions, segregation is at a maximum in both cities according to the three indexes. Both H and H^* assign to each city a segregation value of 1. However, learning a student's school (racial group) in A conveys more information about a student's race (school) than in B . Consequently, segregation in A is larger than in B according to the M index: $M(A) = 1.10$ and $M(B) = 0.69$. Consider now a third completely segregated city C :

$$C = \begin{bmatrix} 99 & 0 \\ 0 & 1 \end{bmatrix}.$$

Both H and H^* assign again to C a segregation value of 1.¹³ However, since there is much less uncertainty about a student's racial group (school) in C than in either A or B , segregation in C according to M is much smaller than before: $M(C) = 0.06$.

As Clotfelter (1979) pointed out, a critical problem with segregation indices that satisfy *NOR* is that they fail to capture well changes in interracial contact. Compare the effect of merging the two schools in city C , yielding the one-school city represented by column vector $[99 \ 1]'$, with the effect of merging the two schools in B , yielding the one-school city represented by $[50 \ 50]'$. The first merger has a very small effect on the interracial exposure of the average student, while the second one has a much larger effect: Each student switches from a completely segregated school to one that is completely integrated. The M index reflects this difference, falling by 0.06 in C versus 0.69 in B . In contrast, H and H^* miss the difference because the segregation value they both assign decreases by 1 in the two cases.

Furthermore, as has been indicated in the introduction, Frankel and Volij (forthcoming) establish the incompatibility of *NOR* and

¹³ As a matter of fact, any $I1$ or $I2$ segregation index that satisfies the principle of transfers and is bounded above by 1 would also assign to the three cities A , B , and C a maximum segregation value of 1 in this example. However, as already stated, H and H^* violate the two invariance properties $I1$ or $I2$, proving that both $I1$ and $I2$ are independent properties from *NOR*.

decomposability properties *SSD* and *SGD*, providing an argument in empirical studies for avoiding indexes that satisfy *NOR*.

Finally, it should be noted that all segregation indices that are bounded above can be weakly normalized, in the sense that they can be expressed as proportions of maximum segregation, by simply dividing them by their maximum values. In particular, the *M* index reaches its maximum at the smallest value between $\log(G)$ and $\log(N)$ because, as a measure of differences in the rows in city *X*, it cannot be larger than $\log(N)$, and, as a measure of differences in the columns in city *X*, it cannot be larger than $\log(G)$. Given that in most empirical applications $\log(G) < \log(N)$, normalizing *M* in this weak sense is simply equivalent to computing the logarithm in base *G*. The resulting measure can be interpreted as the proportion of maximum differences in columns. However, this exercise is not useful for two reasons. First, the most robust feature of the index—namely, the ranking it induces—is still the same and captures both differences in rows and differences in columns. Second, although the resulting index takes values in the unit interval, it still does not satisfy *NOR*.

7. CONCLUSIONS

This paper borrows from the income inequality literature the methodological criterion that one way to select an adequate segregation measure is to study which basic and subsidiary but useful properties different indices satisfy. The importance of doing this is discussed by one of the leading advocates of this approach: “If this search is not undertaken, there is a tendency to continue using those measures that have been popular in the past. The index is then chosen by default, or historical accident, rather than by any assessment of its merits” (Shorrocks 1988:433).¹⁴ We have discussed three types of subsidiary properties

¹⁴ Grusky and Charles (1998:497) complain that this situation has indeed been prevalent in the history of research on occupational segregation by gender: “For all its faddishness, the concept of path dependency proves useful in understanding the history of sex segregation research, and not merely because the index of dissimilarity (hereafter, *D*) has shaped and defined the methodology of segregation analysis over the last 25 years. It is perhaps more important that *D* has been so dominant during this period that it undermined all independent conceptual development. Indeed, segregation scholars have effectively assumed that sex segregation is simply whatever *D* measures.”

as they apply to three entropy-based segregation indices, M , H , and H^* .

First, it is often convenient to have segregation measures with the subsidiary property of additive decomposability. In a decomposition context, consider the notion of contribution to overall segregation by a subgroup k , or by all subgroups together in a certain partition, or consider the question of how much segregation can be attributed to a given discrete variable. As in the income inequality or the economic poverty literature, it is not always possible that all intuitive interpretations of these questions coincide under a certain decomposability property. As shown in this paper, for the first time in the literature these questions receive the more unambiguous answers that are possible in a segregation context under the decomposability properties *SSD* and *SGD* that are only satisfied by the M index. The H and the H^* indices satisfy some weaker decomposition properties. However, numerical examples and actual data have been used to establish that the dependence of the weights in these decompositions on *both* demographic information about the marginal distributions *and* school and racial entropies pose serious problems of interpretation, specially in the decomposition of the H index for partitions of groups into supergroups, and the decomposition of the H^* index for partitions of schools into clusters.

Second, the invariance properties that require a segregation measure to be independent from changes in the relative importance of demographic groups or organizational units have also greatly concerned many authors in the segregation field. The M index is not invariant in this sense but changes in overall segregation according to the M index can be decomposed in two complementary ways to isolate terms that capture changes in net segregation independent of variations in the marginal distributions of schools and racial groups. No such decompositions are available to the H and the H^* indices. When such demographic changes are important, as we have shown to be the case in an example in Section 5.2 and when assessing the change in school segregation in the U.S. during 1989–2005, this is a serious limitation.

Finally, many authors have insisted on the convenience of a third subsidiary property—namely, normalization. This can be easily achieved in our case by dividing the M index into the appropriate population entropy. If the racial entropy is chosen, then the H index is obtained. Similarly, if the entropy of the schools is chosen, then the H^* index is obtained. However, the cost of either normalization is very

high indeed. On the one hand, at a conceptual or intuitive level, it can be argued that neither the H nor the H^* index captures changes in inter-racial or inter-group exposure well. On the other hand, all normalized indices, including the H and the H^* indices, violate the strong decomposability properties *SSD* and *SGD* with the consequences already analyzed.

In conclusion, applied researchers have available three segregation indices based on the entropy notion first advocated by Theil and Finizza (1971): the M index on the one hand, and the H and H^* indices on the other. However, the advantages of the M index are inescapable. In the first place, Frankel and Volij (forthcoming) have formally characterized the ranking induced by the M index in terms of eight ordinal axioms—a result that allows us to know exactly which value judgments are invoked when using this ranking rather than the ones induced by the remaining entropy-based indices for which no such characterization result is available.¹⁵ But beyond this convenient situation, we select which index to use in practice by also taking into account its cardinal properties. In this respect, this paper has shown that when decomposability properties are desired in the empirical work there is much to be gained by focusing exclusively on the un-normalized M index. In addition, when invariance properties are also thought to be useful, it has been seen that applied researchers would do better using the M index and its invariant decompositions rather than using either H or H^* . Finally, the significance of the segregation differences and levels can only be studied under an alternative hypothesis if the measure is explicitly embedded in a statistical framework. Researchers with these considerations in mind can exploit the statistical properties established in Mora and Ruiz-Castillo (2010) for the M index. No comparable statistical framework has yet been provided for the H and H^* indices.

¹⁵ Few segregation indices have been similarly characterized. In the two groups case, Chakravarty and Silber (1992) characterize an index of absolute segregation, while Chakravarty and Silber (2007) axiomatically derive a class of numerical indices of relative segregation that parallel the multidimensional Atkinson inequality indices. Two members of that class are monotonically related to the square root index, independently characterized by Hutchens (2004), and the M index. In the multigroup case, Frankel and Volij (2010) provide an ordinal characterization of an Atkinson index.

REFERENCES

- Bourguignon, Francois. 1979. "Decomposable Income Inequality Measures." *Econometrica* 47:901–20.
- Chakravarty, Satya R., and Jacques Silber. 1992. "Employment Segregation Indices: An Axiomatic Characterization." Pp. 912–20 in *Models and Measurement of Welfare and Inequality*, edited by W. Eichhorn. New York: Springer-Verlag.
- . 2007. "A Generalized Index of Employment Segregation." *Mathematical Social Sciences* 53:185–95.
- Clotfelter, Charles T. 1979. "Alternative Measures of School Desegregation: A Methodological Note." *Land Economics* 54:373–80.
- Fisher, Mary J. 2003. "The Relative Importance of Income and Race in Determining Residential Outcomes in U.S. Urban Areas, 1970–2000." *Urban Affairs Review* 38:669–96.
- Fisher, Claude S., Gretchen Stockmayer, Jon Stiles, and Michael Hout. 2004. "Distinguishing the Geographic Levels and Social Dimension of U.S. Metropolitan Segregation, 1960–2000." *Demography* 41:37–59.
- Flückiger, Yves, and Jacques Silber. 1999. *The Measurement of Segregation in the Labor Force*. Heidelberg: Germany: Physica-Verlag.
- Foster, James E. 1983. "An Axiomatic Characterization of the Theil Measure of Income Inequality." *Journal of Economic Theory* 65:105–21.
- Frankel, David M., and Oscar Volij. (Forthcoming). "Measuring School Segregation." *Journal of Economic Theory*.
- Fuchs, Victor R. 1975. "A Note on Sex Segregation in Professional Occupations." *Explorations In Economic Research* 2:105–11.
- Grusky, David B., and Maria Charles. 1998. "The Past, Present, and Future of Sex Segregation Methodology." *Demography* 35:497–504.
- Herranz, Neus, Ricardo Mora, and Javier Ruiz-Castillo. 2005. "An Algorithm to Reduce the Occupational Space in Gender Segregation Studies." *Journal of Applied Econometrics* 20:25–37.
- Hutchens, Robert M. 2004. "One Measure of Segregation." *International Economic Review* 45:555–78.
- Iceland, John. 2002. "Beyond Black and White." Presented at the American Sociological Association Meetings, August 16–19, Chicago, Illinois.
- James, David R., and Karl E. Taeuber. 1985. "Measures of Segregation." Pp. 1–32 in *Sociological Methodology*, vol. 15, edited by N. B. Tuma. San Francisco, CA: Jossey-Bass.
- Massey, Douglas, and Nancy Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67:281–315.
- Miller, Vincent P., and John M. Quigley. 1990. "Segregation by Racial and Demographic Group: Evidence from the San Francisco Bay Area." *Urban Studies* 27:3–21.
- Mora, Ricardo, and Javier Ruiz-Castillo. 2003. "Additively Decomposable Segregation Indexes. The Case of Gender Segregation by Occupations in Spain." *Journal of Economic Inequality* 1:147–79.

- . 2004. "Gender Segregation by Occupations in the Public and the Private Sectors. The Case of Spain In 1977 and 1992." *Investigaciones Económicas* 28:399–428.
- . 2009. "The Invariance Properties of the Mutual Information Index of Multigroup Segregation." Pp. 3–53 in *Research on Economic Inequality*, Vol. 17, *Occupational and Residential Segregation*, edited by Y. Flückiger, J. Silber, and S. Reardon. Bingley, UK: Emerald Books.
- . 2010. "A Kullback-Leibler Measure of Conditional Segregation." Working Paper 10-15, Universidad Carlos III de Madrid.
- Reardon, Sean, and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." Pp. 33–67 in *Sociological Methodology*, vol. 32, edited by Ross M. Stolzenberg. Boston, MA: Blackwell Publishing.
- Reardon, Sean, John T. Yun and Tamela McNulty. 2000. "The Changing Structure of School Segregation: Measurement and Evidence of Multiracial Metropolitan Area School Segregation, 1989–1999." *Demography* 37:351–64.
- Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures." *Econometrica* 48:613–25.
- . 1984. "Inequality Decomposition by Population Subgroups." *Econometrica* 52:1369–85.
- . 1988. "Aggregation Issues in Inequality Measurement." Pp. 429–51 in *Measurement in Economics: Theory and Applications of Economic Indices*, edited by W. Eichhorn. Heidelberg, Germany: Physica.
- Theil, Henry. 1967. *Economics and Information Theory*. Amsterdam: Netherlands: North Holland.
- . 1971. *Principles of Econometrics*. New York: Wiley.
- . 1972. *Statistical Decomposition Analysis*. Amsterdam, Netherlands: North Holland.
- Theil, Henry, and Anthony J. Finizza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Information Concepts." *Journal of Mathematical Sociology* 1:187–94.

University. He received his PhD in sociology from Carnegie Mellon and his undergraduate degree in computer science from Princeton University. His areas of research include statistical and dynamic methods for social network analysis, numerical optimization, quantitative methodology, and agent-based modeling.

Ricardo Mora is an associate professor in the Department of Economics at Universidad Carlos III de Madrid (Spain). Mora develops and applies empirical methods in topics related to labor economics, agricultural economics, and public economics. In addition to his contributions to the measurement of segregation, he has recently published research on agricultural productivity, wage differentials, and the measurement of quality of care in hospitals.

Meike Morren is a doctoral candidate in the Department of Methodology and Statistics at Tilburg University. Her research focuses on the cross-cultural comparability of attitudes and integrates cognitive interviewing and latent variable modeling to obtain a deeper understanding about how response styles may confound cross-cultural differences.

Javier Ruiz-Castillo is a professor in the Department of Economics at Universidad Carlos III de Madrid (Spain). His research is related to welfare economics, public economics, and microeconomics. In addition to his contributions to the measurement of segregation, he has recently published research on the measurement of scientific performance, the welfare analysis of tax reform, and the imputation of rental values to owner-occupied housing.

Elizabeth Savoca is a professor in the Economics Department at Smith College. Her research includes many topics related to health policy: the impact of tort reforms on the resolution of medical malpractice disputes, the effect of professional regulations on the quality of health care, and the antecedents and consequences of mental illness in adults. She has published extensively on econometric issues related to misclassification in survey measures of psychiatric health.

Jeroen Vermunt is a professor in the Department of Methodology and Statistics at Tilburg University. His research is on methodology of