Original Article

Sociological Methods & Research
1-38

DOI: 10.1177/0049124121986204
journals.sagepub.com/home/smr
SAGE

A Method for Studying Differences in Segregation Across Time and Space

Benjamin Elbers[1]

Abstract

An important topic in the study of segregation are comparisons across space and time. This article extends current approaches in segregation measurement by presenting a five-term decomposition procedure that can be used to understand more clearly why segregation has changed or differs between two comparison points. Two of the five terms account for differences in segregation that are due to the differing marginal distributions (e.g., the gender and occupational distributions), while one term accounts for differences in segregation due the different structure of segregation (what might be termed "pure" segregation). The decomposition thus presents a solution to the problem of margin dependency, frequently discussed in the segregation literature. Finally, two terms account for the appearance or disappearance of units when analyzing change over time. The method can be further extended to attribute structural changes to individual units, which makes it possible, for instance, to quantify the effect of each occupation on changing gender* segregation. The practical advantages of the decomposition are illustrated by two examples: a study of changing occupational gender segregation in the United States and a study of changing residential segregation in Brooklyn, New York.

[1] Department of Sociology, Columbia University, New York, NY, USA

Corresponding Author:
Benjamin Elbers, Department of Sociology, Columbia University, 606 West 122nd Street, 5th Floor, Suite 501, New York, NY 10027, USA.
Email: be2239@columbia.edu

Keywords

segregation, occupational segregation, residential segregation, school segregation, decomposition

Studies of segregation are concerned with a variety of substantive problems. Social scientists are interested in residential racial segregation, in the racial or class-based segregation of schools and workplaces, or the gender segregation of occupations. More generally, any study of the association between two categorical variables can be regarded as a segregation problem. Segregation is usually studied by applying a segregation index to a contingency table, which provides a one-number summary of the association between, for instance, gender and occupations.

Often, the interest in the study of a segregation problem lies not only in describing segregation at one point in time or in one place but in comparing levels of segregation over time, across countries or cities, or between population groups. For instance, in the school segregation literature, there is a debate about the resegregation of schools along racial lines (Reardon and Owens 2014). The workplace segregation literature documented a decrease in within-workplace racial segregation levels but a decrease in between-workplace segregation (Ferguson and Koning 2018). The gender-occupational literature is interested not only in comparing segregation over time within a single country but also across regional or national economies (e.g., Charles and Grusky 2004). When comparing across time, the message of segregation studies is often that segregation has either increased or decreased, but the deeper causes for these differences often remain unclear. The contribution of this article is to provide a general and practical method for the study of change or difference in segregation (hereafter abbreviated as "change" when this doesn't create confusion). The method developed here brings practical advantages to many segregation problems, and proposes a solution to the long-standing problem of margin dependency:

1. Among the practical advantages, the method allows an analysis of where differences in segregation originate. For instance, we might ask whether declining occupational gender segregation arises from manual or professional occupations or whether the declines in segregation are associated with changes in the educational composition of certain occupations. In school or residential segregation, it would be

of interest to know which schools or neighborhoods contribute most to changes in segregation. This is especially relevant when considering different types of schools such as charter or private schools. Relatedly, one may study the association between gentrification and segregation at the neighborhood level.

2. The method also allows for a change in the number of units under study. This problem arises naturally in the study of school segregation: When comparing school segregation across two points in time, some schools will have closed and new ones will have opened. The problem may also occur with occupational segregation: Over time, some occupations will become obsolete and vanish, while new occupations appear. The method developed here allows the researcher to quantify the effect of these "appearing" and "disappearing" units on the total change in segregation. While this seems a natural question, it has only received scant attention in the segregation literature. An exception is Ferguson and Koning (2018) who studied the effect of firm turnover on workplace segregation.
3. Finally, the method provides a solution to the problem of margin dependency. Taking again occupational gender segregation as an example, it is intuitively clear that some of the declines in gender segregation of recent decades may be due to compositional changes. Deindustrialization has led to declines in factory jobs and a decrease in the share of manual and routine occupations, which have often been almost entirely male (Weeden 2004). If these occupations are still as segregated as they were before, and only their relative share has declined, this will register as a decrease in (most) segregation indices. Thus, it would be desirable to distinguish between these changes, which are referred to as marginal changes (because the change is reflected in the marginal row or column sums of the contingency table), from changes in "pure segregation." A major part of the article will elaborate on this distinction and on the exact meaning of "pure segregation."

The methodological literature on segregation indices has engaged mostly with point (3), the margin dependency of segregation indices, while the useful innovations described in points (1) and (2) have received almost no attention. The method described in this paper proposes a solution to the margin dependency problem that can be summarized as follows: Margin dependency is desirable in the cross-section to characterize the "average" level of segregation an individual experiences but is problematic when comparing levels of segregation across time or space. The solution, as first proposed by Karmel and

Maclachlan (1988), is to decompose the difference into terms that distinguish changes that are introduced because of changes in the marginal distributions alone from changes in "pure segregation." The latter will be called "structural change" throughout the article. Combining this idea with the desired properties of (1) and (2), we arrive at a five-term decomposition:

$$S(t_2) - S(t_1) = \Delta_{\text{appearing}} + \Delta_{\text{disappearing}} + \Delta_{\text{marginal-units}} + \Delta_{\text{marginal-groups}} + \sum_{u \in t_1 \cap t_2} \Delta_{\text{i,structural}},$$

where $S(\cdot)$ refers to the value of the segregation index at different points in time. The equation then says that we decompose the difference in segregation between two time points (or across population groups, places) into two terms that account for the appearance and disappearance of units under study (think school openings and school closures), two terms that account for compositional changes (the marginal distributions), both in terms of units (say, schools) and groups (say, racial groups). The last term is a summation that extends over those units that are present at both time points and describes the change in structural (or "pure") segregation that arises from each unit.

Thus, the decomposition opens up new avenues of research for scholars working on segregation problems. Its primary advantage is that it allows for a much more precise statement about the nature of change: We can pinpoint whether the segregation change is due to a change in the population of units, due to marginal change, or due to structural change. We can further drill down to study whether the structural change is concentrated in a certain set of units that are of special interest (say, charter schools). It should also be noted that the total change in segregation could be zero, but that some of the components are nonzero. In this case, some positive components would be offset by negative components. The decomposition of change could thus reveal previously obscured patterns such as an increase in "pure segregation" that is offset by declines due to marginal changes. (Such offsetting patterns are explored in the examples at the end of the article.)

The remainder of the article is organized as follows: In the next section, the issue of margin dependency and the possible solutions that have been presented in the literature are discussed. None of these solutions are deemed satisfactory. It is then argued that the only index that can fully achieve the desired five-term decomposition is the *M* index. This index, extensively discussed by Mora and Ruiz-Castillo (2003, 2009, 2011), is not as widely used as the closely related *H* index but has many desirable properties. Next, the decomposition procedure is introduced. Finally, the practical advantages of the method are shown through two examples: changing occupational gender segregation in the United States and changing residential segregation in Brooklyn, New York.

All procedures discussed in this article have been implemented in an R package. *This includes the calculation of the M and H indices, as well as the decomposition procedure. Standard errors can be bootstrapped for both the index calculation and the decomposition procedure. Examples on how to use the package are given in Online Appendix C (which can be found at http://smr.sagepub.com/supplemental/).

The Problems and Benefits of Margin Dependency

To make the following more concrete, consider U organizational units, such as schools or occupations, and a number of population groups, G, such as racial groups or genders. For an occupational segregation problem, the number of workers in each occupation–gender combination can be cross-classified in a $U \times G$ contingency table. A segregation index $S(\cdot)$ is a function that summarizes the $U \times G$ contingency table to a single number. Without loss of generality, occupational gender segregation will be used as an example for the remainder of the article.

Margin dependency refers to the property of some segregation indices that proportional changes in the marginal distributions of the contingency table lead to a change in the index value. To illustrate, consider a simplified economy with three occupations and two genders. At time point 1, there are 55 men and 45 women distributed across occupations in a way that the first occupation is integrated, while the other two are rather segregated. This matrix is shown at the left-hand sides of the arrows, with men in the first and women in the second column:

$$t_1 : \begin{bmatrix} 25 & 25 \\ 28 & 2 \\ 2 & 18 \end{bmatrix} \rightarrow t_2 : \begin{bmatrix} 20 & 20 \\ 28 & 2 \\ 4 & 36 \end{bmatrix},$$

$$t_1 : \begin{bmatrix} 25 & 25 \\ 28 & 2 \\ 2 & 18 \end{bmatrix} \rightarrow t_2^* : \begin{bmatrix} 25 & 50 \\ 28 & 4 \\ 2 & 36 \end{bmatrix}.$$

Consider then two alternative scenarios. In the first scenario (top), the size of the first, integrated occupation decreases by 20 percent, and the third occupation (which is very segregated) doubles. Note that it is not possible under these transformations to keep the overall gender proportion constant without changing the internal proportion of the remaining occupation. In the

Table 1. Margin Dependency of Different Indices.

		Unit (e.g., Occupation)	
		Margin Dependent	Margin Free
Group (e.g., Gender)	Margin dependent	*M*, *H*, *V*	SSD
	Margin free	*D*	*A*

Source: Adapted from Charles and Grusky (1995:934).

second scenario (bottom), the size of the female labor force doubles, with the numbers for men unchanged. An index that changes its value under the first transformation is called unit-margin-dependent, while an index that changes its value under the second transformation is called group-margin-dependent. A margin-free index, by definition, does not change under either of these processes. An overview of prominent indices is displayed in Table 1.

The entropy-based, information-theoretic indices *M* and *H* are margin-dependent for both groups and units. This is also true for the variance ratio index *V* (also known as separation or eta-squared index). Other indices, such as the index of dissimilarity *D*, are only margin-dependent in terms of the unit distribution. The size-standardized index of dissimilarity SSD is group-margin-dependent only,[1] and only the log-linear index *A* is margin-free in both dimensions.

The reader might be surprised to find the *H* index among the group-dependent indices. The margin dependency of the *H* index is often not explicitly considered in empirical studies although this fact is known at least since James and Taeuber (1985). For instance, An and Gamoran (2009:20) write that they "use a measure [the *H*] that is insensitive to changes in the U.S. school population, thereby concentrating solely on racial imbalance." This, however, is not entirely true. While the *H* index involves a term that partly accounts for changes in group marginals, the standardization is not complete (for a formal proof, see Mora and Ruiz-Castillo 2011). We thus emphasize here that the *H* is margin-dependent in both directions.

Often, margin dependency is considered problematic, and the segregation literature has devoted considerable effort to solving this problem. The problem stems from the assumption that marginal changes often reflect processes that are thought to be unrelated to the deeper, structural causes of segregation. For instance, deindustrialization (changing occupation marginals) or a rising share of female employment (changing group marginals) should only lead to changes in segregation if the structure of segregation changed. If the changes are in the margins only, arguments about "rising" or "decreasing" segregation may not be warranted.

The major solutions to the problem are discussed in turn:

1. A series of papers by Charles and Grusky (Charles 1992; Charles and Grusky 1995; Grusky and Charles 1998) introduced the *A* index. The *A* index is based on the insight that a measure of segregation that is invariant to row or column transformations needs to be based on odds ratios. For instance, two local odds ratios are sufficient to describe the association structure of a 3×2 table (as in the example above). If we let n_{ij} denote the number of workers in the *i*th row and *j*th column, the two odds ratios are $\theta_{1,1} = \frac{n_{1,1}n_{2,2}}{n_{1,2}n_{2,1}}$ and $\theta_{2,1} = \frac{n_{2,1}n_{3,2}}{n_{2,2}n_{3,1}}$ (Agresti 2013:54). It is easy to verify that these odds ratios are identical for all three matrices t_1, t_2, and t_2^*, which is to say that the association structure between occupations and gender does not change from t_1 to t_2 or from t_1 to t_2^*. This is the same argument that is made in favor of log-linear modeling in the study of social mobility.

Essentially, the *A* index calculates the odds ratio of male and female employment within each occupation and is then summarized by weighting all occupation-specific ratios equally. The resulting index measures only the level of association as captured by the odds ratios and is not influenced by changes in the marginal distribution of either occupations or genders. Note that the index achieves its unit-margin-independence by simply weighting all occupations equally. The index is thus more a characterization of the segregation of the average occupation and not a measure of average segregation at the individual level. Especially if the sizes of occupations differ greatly, the index is problematic (see also the exchange between Watts [1998] and Grusky and Charles [1998]). The index thus seems even less applicable when school or residential segregation is studied.

Another way to phrase this problem is that the *A* index conflicts with the criterion of organizational equivalence. Organizational equivalence implies that when two occupations with the same level of segregation are combined, segregation should be unchanged (James and Taeuber 1985). This criterion is not fulfilled when occupations are weighted equally and the segregation level of the other, uncombined occupations differ from the two occupations that are combined. This shows that the discussion about the merits of margin-free versus margin-dependent indices cannot be resolved because the two indices pursue goals that are not compatible.

2. Karmel and Maclachlan (1988) propose a decomposition that is very similar to the one developed in this article. Their approach is based on creating counterfactual contingency tables that account only for the effects of marginal and structural changes, respectively. This is done using iterative proportional fitting (IPF), which will be explained below. The counterfactual tables can then be used to disentangle marginal from structural changes. A downside of their approach is that the decomposition contains an interaction effect between the two marginal dimensions, which is hard to interpret. They also do not address the problem of appearing and disappearing units. The largest disadvantage of their method is the choice of index, which they call I_p, and which is not decomposable in terms of units or groups.
3. Mora and Ruiz-Castillo (2009) presented two formulas that supposedly quantify structural and compositional change between two M indices. With a slightly adapted notation, the difference between two M indices, defined by the matrices t_1 and t_2, is decomposed as follows:

$$\begin{aligned} M(t_2) - M(t_1) &= \Delta\mathrm{N}(\Pi^u) + \Delta\mathrm{G}^u + \Delta\mathrm{U}(\Pi^u) \\ &= \Delta\mathrm{N}(\Pi_g) + \Delta\mathrm{U}_g + \Delta\mathrm{G}(\Pi_g), \end{aligned} \tag{1}$$

where the ΔU and ΔG capture changes in the marginals of unit and group proportions, respectively, and ΔN captures "composition–invariant" changes, which, importantly, are not the same as structural changes defined through the change in odds ratios. As the authors themselves write, the interpretation of these terms hinges on crucial assumptions that are rarely met in practice (Watts 2015; Mora and Ruiz-Castillo 2009:47-50). For reasons of brevity, these problems are not explicated fully here. Instead, an especially problematic aspect of these decompositions is highlighted, and that is that there are two possible answers for each of the three components, which might provide conflicting interpretations. The decompositions on the first and the second line will only in exceptional circumstances give the same numerical results. This is easily seen by applying equation (1) to the difference between t_1 and t_2 from the example above:

$$\begin{aligned} M(t_2) - M(t_1) &= 0 + 0.00376 + 0.0267 \\ &= -0.0209 + 0.0479 + 0.00346 = 0.03. \end{aligned}$$

The first decomposition implies that structural change is zero and further suggests that the marginal change in the occupational distribution is largely responsible for the increase in segregation, which aligns with our expectations. However, the second line gives a contradictory answer, implying that structural segregation decreased (−0.0209). Furthermore, the size of the marginal components is not the same in the two decompositions. Even if the assumptions that underlie these decompositions were justified in practice (which is questionable), the fact that the two decompositions give two possibly contradictory answers is unsatisfactory and poses practical problems of interpretation. The issue here is that these decompositions are not based on the notion that only the odds ratios are invariant under row and column transformations. Finally, their decomposition also does not address the problem of appearing and disappearing units, which means that only the common subset of units can be decomposed.

The method developed in this article is based on the idea that margin dependency (especially in terms of units) is a desirable property in the cross-section. Consistent with the idea that we want to measure average segregation at the individual level, it is reasonable to argue that a segregation index should be higher when more people work in segregated occupations. If occupations are weighted equally, this is not the case. At the same time, we would also like to distinguish changes that are purely due to composition (marginal changes) from changes in pure segregation (structural changes). To illustrate this point, consider that two processes occur at the same time: The occupations that are more segregated grow at the expense of less-segregated occupations, while at the same time segregation *within* each occupation declines. The overall change in segregation will be positive if the first process leads to a greater change than the second process. If attention is only paid to the total difference, the conclusion will be that segregation has become "worse" (which is a warranted statement, at least for the average worker). However, the statement is also imprecise because the segregation of each individual occupation has in fact decreased. The decomposition of change into the two components thus allows the researcher to pinpoint more clearly the sources of segregation change. Importantly, the prevalence and direction of the two trends may call for different policy responses.

Thus, the article advocates for an approach that uses a margin-dependent index in the cross-section, which is then decomposed when we compare over time or across places. The proposed solution combines and expands the approaches (1)–(3) discussed above. Charles and Grusky provide the key insight that any structural changes are reflected in the odds ratios and that these are the only measures of association that are invariant under marginal transformations. Karmel and Maclachlan use IPF to arrive at counterfactual tables. Finally, Mora and Ruiz-Castillo's contributions highlighted the advantages of the entropy-based index M, which will be adopted below.

The Choice of Index

Recently, the H has become increasingly popular for the study of racial segregation, which is most likely due to two distinct advantages. First, the H allows for attractive decompositions. Second, the H allows for a natural treatment of the multigroup case, which has become increasingly important for the study of racial segregation in the United States and is a natural requirement in other segregation problems. In their comprehensive overview of multigroup segregation indices, Reardon and Firebaugh (2002) conclude "that the information theory index H is the most conceptually and mathematically satisfactory index" (p. 33).

In a recent series of papers, Mora and Ruiz-Castillo (2003, 2009, 2011) pointed to an alternative but closely related index, which they called the *Mutual Information Index (M)*. Both the M and H were introduced by Theil (Theil 1967, 1971, 1972; Theil and Finizza 1971). Mora and Ruiz-Castillo, as well as Frankel and Volij (2011), outlined some of the advantages of the M over the H. Importantly, Mora and Ruiz-Castillo (2011) showed that the decomposition of an H index into between- and within-group terms (for instance, white/nonwhite) may be ambiguous, and they thus recommend the adoption of the M if such decompositions are desired.[2]

To define H and M, assume that we observe the gender composition of U occupations. Define n_{gu} as the number of workers with gender g in occupation u and the total number of workers as n. From this contingency table, define $p_{\cdot u} = \sum_{g=1}^{G} n_{gu}/n$ and $p_{g\cdot} = \sum_{u=1}^{U} n_{gu}/n$ as the marginal probabilities of occupations and gender, respectively. The joint probability of being in occupation u and gender g is $p_{gu} = n_{gu}/n$. We also write $p_{g|u} = p_{gu}/p_{\cdot u}$ as the conditional probability of having gender g given occupation u (and $p_{u|g}$ likewise).

The M index quantifies how strongly each occupation's gender distribution deviates from the overall (or expected) gender distribution. This yields a "local" segregation score for each occupation, called L_u. The occupation scores are then weighted by the size of the occupation, $p_{\cdot u}$. To measure the deviation, the logarithm of the ratio between conditional and marginal probabilities is used. As Theil (1972) has shown, the logarithm allows for the attractive decomposition properties. Thus,

$$M = \sum_u p_{\cdot u} L_u = \sum_u p_{\cdot u} \left(\sum_g p_{g|u} \log \frac{p_{g|u}}{p_{g\cdot}} \right). \tag{2}$$

Because the M is symmetric, it can also defined by summing proportion-weighted scores for each gender, that is,

$$M = \sum_g p_{g\cdot} L_g = \sum_g p_{g\cdot} \left(\sum_u p_{u|g} \log \frac{p_{u|g}}{p_{\cdot u}} \right). \tag{3}$$

The simple expressions for the M, that is,

$$M = \sum_u p_{\cdot u} L_u = \sum_g p_{g\cdot} L_g,$$

show that the M is symmetric (i.e., the meaning of groups and units can be exchanged) and that the M is margin-dependent in both directions. From the standpoint of decomposing changes in segregation, this is an attractive property.

The M can also be motivated from an information-theoretic perspective, which is helpful to understand its basic properties. First, define the entropy $E(\cdot)$ of a distribution as

$$E(\mathbf{p}) = -\sum_i p_i \log p_i,$$

where $\mathbf{p}$ is a vector of probabilities that sums to 1. Entropy is a nonnegative measure of *expected information* or *uncertainty* (Theil 1972). Consider two events that occur with probabilities .99 and .01. The expected information of the next observation from this distribution is close to zero, that is, $E([.99, .01]) = .06$, as we were virtually certain that the first event would occur. However, for two events that will occur with a probability of ½ each, the expected information is large, that is, $E([.5, .5]) = \log 2 \approx .69$. The entropy is maximized at $\log n$ when the probability of each event is $1/n$, where n is the number of events. Intuitively, the entropy is minimized at zero when it is certain which event will occur.

To define M from this perspective, we ask: How much more information does the overall distribution provide compared with the gender distribution of a specific occupation? Formally, this is the difference in entropies at the occupation level, weighted by the occupation's proportion:

$$M = \sum_u p_{\cdot u} \left[E(\mathbf{p}_{g\cdot}) - E(\mathbf{p}_{g|u}) \right], \tag{4}$$

where $\mathbf{p}_*$ refers to the relevant vector of probabilities. Due to the symmetry of the M, this expression can also be formulated from the gender perspective:

$$M = \sum_g p_{g\cdot}\left[E(\mathbf{p}_{\cdot u}) - E(\mathbf{p}_{u|g})\right]. \tag{5}$$

It follows that M is minimized at zero when the gender distribution of each occupation is identical to the overall gender distribution. M is maximized at $\min(\{\log U, \log G\})$. To see this, note that equation (4) is maximized when the entropy $E(\mathbf{p}_{g\cdot})$ is maximized, and the entropy $E(\mathbf{p}_{g|u})$ for each occupation is minimized. This is the case when each gender has the same overall proportion and when each occupation is either completely male or completely female.

It may seem odd that a segregation index can only be maximized when all groups are the same size, but it is in line with information-theoretic principles. This point will become clearer with an example. Consider two labor markets A and B with 200 workers each, and only three occupations. The labor markets differ in their gender distributions. Labor market A has 100 women and 100 men, while B has 20 women and 180 men. The workers are distributed as follows, with the occupations indexed by the rows of the matrix:

$$\text{A}: \begin{bmatrix} \text{women} & \text{men} \\ 100 & 0 \\ 0 & 50 \\ 0 & 50 \end{bmatrix} \qquad \text{B}: \begin{bmatrix} \text{women} & \text{men} \\ 20 & 0 \\ 0 & 90 \\ 0 & 90 \end{bmatrix}$$

In both labor markets, all three occupations are completely segregated in the sense that there is no mixing within occupations. For these matrices, $M(A) = 0.69$ and $M(B) = 0.33$. The M index thus finds that segregation in A is twice as high as in B. This suggests to standardize the M index by the gender entropy, which gives the H index:

$$H = \frac{M}{E(\mathbf{p}_{g\cdot})}.$$

For the two cities, it follows that $H(A) = H(B) = 1$. The H is attractive because it is standardized between zero and one,[3] which facilitates comparisons between two cities with differing gender distributions. Nonetheless, there is an argument to be made for the M index. While the H index sees the amount of segregation as equal between the two cities, the M takes into account that it is much “harder” in A to achieve complete segregation than it is in B: Given that in B 90 percent of the workers are men, it is less surprising to find an all-men occupation in B than it is in A.

The Decomposition of Change

Generating Counterfactuals Through IPF

Instead of attempting the margin-free measurement of segregation at each point in time, the approach outlined here follows the idea that *changes* in segregation indices can be decomposed into marginal and structural changes (Watts 1998; Mora and Ruiz-Castillo 2009; Watts 2015). This method was proposed by Theil himself (Theil 1972:131) and was extended by Karmel and Maclachlan (1988) in the context of occupational gender segregation. Karmel and Maclachlan used another segregation index, but the approach is applicable whenever a margin-free comparison of two contingency tables is desired.

The basic idea is to adjust the contingency table from time point t_1 forward so that only marginal changes between the two time points are taken into account. Consider a labor market with men and women distributed across three occupations. We observe the labor market at two points in time. Between these two time points, the number of men has grown and the number of women has declined. At the same time, occupations have changed in size, with especially strong declines in the third occupation. The question is: If there are changes in segregation, how much of these changes can be attributed to changes in the distribution of gender and occupation marginals alone, and how much of the change can be attributed to changes in the odds ratios?

At the two time points, the workers are distributed across occupations as follows:

$$t_1 : \begin{bmatrix} \text{women} & \text{men} \\ 20 & 100 \\ 180 & 50 \\ 600 & 50 \end{bmatrix} \quad \text{t}_2 \text{:} \begin{bmatrix} \text{women} & \text{men} \\ 10 & 170 \\ 80 & 60 \\ 240 & 40 \end{bmatrix}$$

Both the M and the H register large changes in segregation: The M increases by over 80 percent between t_1 and t_2, while the H increases by 33 percent. To identify how much of this change is due to marginal changes, the matrix at t_1 is transformed to have the same margins as t_2, while leaving the association structure (i.e., the odds ratios) intact. This can be achieved using IPF: First, the cells of t_1 are scaled to achieve the overall *gender* marginal distribution of t_2. The adjusted cell counts are then scaled to achieve the marginal *occupation* distribution of t_2. This process is repeated until the margins of the adjusted table are within a small percentage of t_2. The first steps of the procedure are shown here:

$$
\begin{aligned}
&\begin{bmatrix} 20 & 100 \\ 180 & 50 \\ 600 & 50 \end{bmatrix} && \Rightarrow \begin{bmatrix} 20 \times 330/800 & 100 \times 270/200 \\ 180 \times 330/800 & 50 \times 270/200 \\ 600 \times 330/800 & 50 \times 270/200 \end{bmatrix} \\
&= \begin{bmatrix} 8.3 & 135 \\ 74.3 & 67.5 \\ 248 & 67.5 \end{bmatrix} && \Rightarrow \begin{bmatrix} 8.3 \times 180/144.4 & 135 \times 180/144.3 \\ 74.3 \times 140/141.8 & 67.5 \times 140/141.8 \\ 248 \times 280/315.5 & 67.5 \times 280/315.5 \end{bmatrix} \\
&= \begin{bmatrix} 10.3 & 168.4 \\ 73.4 & 66.6 \\ 220.1 & 59.9 \end{bmatrix} && \Rightarrow \begin{bmatrix} 10.3 \times 330/303.8 & 168.4 \times 270/294.9 \\ 73.4 \times 330/303.8 & 66.6 \times 270/294.9 \\ 220.1 \times 330/303.8 & 59.9 \times 270/294.9 \end{bmatrix} \\
&= \begin{bmatrix} 11.2 & 154.2 \\ 79.7 & 61 \\ 239.1 & 54.8 \end{bmatrix} && \Rightarrow \begin{bmatrix} 11.2 \times 180/165.4 & 154.2 \times 180/165.4 \\ 79.7 \times 140/140.8 & 61 \times 140/140.8 \\ 239.1 \times 280/293.9 & 54.8 \times 280/293.9 \end{bmatrix} \\
&= \begin{bmatrix} 12.2 & 167.8 \\ 79.2 & 60.7 \\ 227.8 & 52.2 \end{bmatrix} && \Rightarrow \text{...(10 steps omitted)} \\
&= \begin{bmatrix} 13.7 & 166 \\ 83.5 & 56.5 \\ 233 & 47.3 \end{bmatrix} = t'_1
\end{aligned}
$$

The transformations at rows 1 and 3 adjust the gender marginal, while the transformations at rows 2 and 4 adjust the occupation marginals. It is unimportant whether the procedure starts with the group or the unit marginals; it will always converge (for details on IPF, see Deming and Stephan 1940; Agresti 2013).[4] After four steps, both margins are already within 3-4 percent of the desired marginals. After 14 steps, the procedure yields the matrix shown in the last row, where the marginals are within 0.1 percent of the desired marginals. The resulting matrix t'_1 is a counterfactual version of the t_1 matrix, where only the marginals changed in the direction empirically observed in t_2, but the odds ratios are the same as in t_1. This allows a decomposition of overall change in segregation levels as follows:

$$
\begin{aligned}
M(t_2) - M(t_1) &= \overbrace{M(t_2) - M(t'_1)}^{\text{structural}} + \overbrace{M(t'_1) - M(t_1)}^{\text{marginal}} \qquad (6) \\
&= (0.273 - 0.238) + (0.238 - 0.150) \\
&= 0.035 + 0.088 = 0.123.
\end{aligned}
$$

The "marginal" component quantifies how much we would expect segregation to change given that the marginals changed toward those of t_2. The "structural" component quantifies any additional amount of segregation that is unexplained by marginal changes. To understand the behavior of the decomposition, it is useful to consider the two extreme cases of "structural change only" and "marginal change only." Considering t_1, it is possible to construct an alternative matrix that redistributes the workers across occupations in such a way that the marginals will stay the same (e.g., by distributing 50 workers from occupation 1 to the other two occupations, and moving the same number of women to occupation 1). A decomposition of these two matrices will find that marginal change is zero because the IPF procedure converges immediately without changing any cell counts. Thus, the marginal term of equation (6) would compare identical matrices, and the difference would be zero—as desired. Similarly, it is also possible to construct a matrix where simply the number of, say, women doubled. In this case, the IPF procedure scales the margins in exactly this way, which means that the structural term of equation (6) compares identical matrices, and we again obtain the desired result.

One criticism that can be leveled against this decomposition is that the choice of t_1 as the baseline is somewhat arbitrary, especially if the matrices are not compared over time, but across space or, say, across birth cohorts. The results are similar but not identical when we instead choose t_2 as the baseline and apply the IPF procedure to this matrix:

$$\begin{aligned} M(t_2) - M(t_1) &= \overbrace{M(t'_2) - M(t_1)}^{\text{structural}} + \overbrace{M(t_2) - M(t'_2)}^{\text{marginal}} \\ &= 0.026 + 0.097 = 0.123. \end{aligned}$$

In decomposition analysis, this is known as the path dependency problem (Kitagawa 1955; Fortin, Lemieux, and Firpo 2011), where the results of the decomposition are dependent on the order in which elements are eliminated. As proposed by Shorrocks (2013), the solution to this problem is the Shapley decomposition, which considers all possible ways in which an element can be eliminated. In this case, the decomposition results in a simple averaging of the two scenarios (Deutsch, Flückiger, and Silber 2009):

$$\begin{aligned} M(t_2) - M(t_1) = & \overbrace{\frac{1}{2}(M(t_2) - M(t'_2)) + \frac{1}{2}(M(t'_1) - M(t_1)}^{\Delta_{\text{marginal}}} \\ & \underbrace{+\frac{1}{2}(M(t_2) - M(t'_1)) + \frac{1}{2}(M(t'_2) - M(t_1))}_{\Delta_{\text{structural}}}. \end{aligned} \tag{7}$$

For the example, this is

$$M(t_2) - M(t_1) = \underbrace{\frac{1}{2}(0.097 + 0.088)}_{\Delta_{\text{marginal}}} + \underbrace{\frac{1}{2}(0.035 + 0.026)}_{\Delta_{\text{structural}}}$$
$$= 0.092 + 0.031 = 0.123.$$

From this decomposition, we conclude that marginal changes are responsible for about three quarters of the overall change in the *M*, while structural changes account for only a quarter of the increase. Compared to segregation indices that focus on structure only (i.e., odds ratios), the procedure introduced here quantifies the effects of both marginal and structural changes. It will be argued below that marginal changes are often an important part of segregation processes and that it is therefore not always desirable to "purge" the influence of the marginal distributions.

This aggregate view of segregation differences can be further decomposed. The key property that is exploited here is that in the marginal component, the odds ratios are the same, and that in the structural component, the marginal distributions of units and groups are the same.

Decomposing Marginal Changes Further

The marginal change can be further subdivided into two components: One component quantifies the contribution of changing unit marginals and one quantifies the contribution of changing group marginals. Karmel and Maclachlan proposed a simpler decomposition that includes an interaction term, but the Shapley decomposition can be used to quantify the contributions of either margins without an interaction term. A full proof of this strategy is provided by Deutsch et al. (2009), and we will present here the intuitive understanding of this decomposition. Again, we consider all the ways in which either marginal component can be eliminated. For this, we need to consider all possible combinations between unit marginals, group marginals, and odds ratios from both t_1 and t_2. As a shorthand notation, we will write $M(U; G; O)$ to identify the *M* that is calculated based on the unit (row) marginals from *U*, the group (column) marginals from *G*, and the odds ratios from *O*. For instance, $M(t_1) = M(t_1; t_1; t_1)$ and $M(t'_1) = M(t_2; t_2; t_1)$. Given all possible combinations, there are eight unique matrices, including the two unaltered ones. This decomposition thus requires six distinct IPF procedures. For instance, to arrive at $M(t_1; t_2; t_1)$, the matrix t_1 has to be adjusted toward the column marginals of t_2 while retaining its original t_1 row marginals. The decomposition then relies on averaging all possible elimination strategies. To quantify the effect of marginal change in the rows, there are four possible elimination strategies:

$$\Delta_{\text{marginal}-\text{units}} = \frac{1}{4}(M(t_2;t_1;t_1) - M(t_1;t_1;t_1)) + \frac{1}{4}(M(t_2;t_2;t_1) - M(t_1;t_2;t_1)) + \frac{1}{4}(M(t_2;t_2;t_2) - M(t_1;t_2;t_2)) + \frac{1}{4}(M(t_2;t_1;t_2) - M(t_1;t_1;t_2)). \quad (8)$$

Note that within each subtraction, only the row margins are changed, with the other two factors held constant. Similarly, for the columns:

$$\Delta_{\text{marginal}-\text{groups}} = \frac{1}{4}(M(t_1;t_2;t_1) - M(t_1;t_1;t_1)) + \frac{1}{4}(M(t_2;t_2;t_1) - M(t_2;t_1;t_1)) + \frac{1}{4}(M(t_2;t_2;t_2) - M(t_2;t_1;t_2)) + \frac{1}{4}(M(t_1;t_2;t_2) - M(t_1;t_1;t_2)). \quad (9)$$

Simple algebra shows that $\Delta_{\text{marginal}-\text{units}} + \Delta_{\text{marginal}-\text{groups}} = \Delta_{\text{marginal}}$. Applying this decomposition to the example above, we get:

$$\begin{aligned}\Delta_{\text{marginal}} &= \Delta_{\text{units}} + \Delta_{\text{groups}} \\ &= 0.082 + 0.01 = 0.092.\end{aligned}$$

Among the changes in the marginals, the shift in the unit marginals was much more important for the increase in segregation than the shifting gender distribution, despite the large changes.

Decomposing Structural Changes Further

Usually, structural change is of greater interest than marginal change. The term for the structural component admits two straightforward decompositions based on local segregation scores. These decompositions were not exploited by Karmel and Maclachlan (1988) or others because their index did not admit disaggregation by local segregation scores. The key property that these decompositions exploit is that $p^{t_2}_{\cdot u} = p^{t'_1}_{\cdot u}$, $p^{t_2}_{\cdot g} = p^{t'_1}_{\cdot g}$, $p^{t_1}_{\cdot u} = p^{t'_2}_{\cdot u}$, and $p^{t_1}_{\cdot g} = p^{t'_2}_{\cdot g}$, that is, the equivalence of the margins. We can thus write:

Table 2. Decomposition of Structural Changes Into Contributions of Each Occupation.

Occupation	Proportion		Observed		Counterfactual		Weighted Difference
u	$p_{\cdot u}^{t_1}$	$p_{\cdot u}^{t_2}$	$L_u(t_1)$	$L_u(t_2)$	$L_u(t'_2)$	$L_u(t'_1)$	$\Delta_{u,\text{structural}}$
1	.12	.300	.928	.573	1.056	.515	.016
2	.23	.233	.001	.001	0.003	.004	.000
3	.65	.467	.059	.216	0.075	.177	.014

$$\begin{aligned}\Delta_{\text{structural}} &= \frac{1}{2}(M(t_2) - M(t'_1)) + \frac{1}{2}(M(t'_2) - M(t_1)) \\ &= \sum_{u=1}^{U}\left(\frac{1}{2}p_{\cdot u}^{t_2}[L_u(t_2) - L_u(t'_1)] + \frac{1}{2}p_{\cdot u}^{t_1}[L_u(t'_2) - L_u(t_1)]\right) \\ &= \sum_{u=1}^{U}\Delta_{u,\text{structural}},\end{aligned} \tag{10}$$

where $L_u(X)$ refers to the local segregation score for unit u in matrix X. The difference in structural segregation can thus be attributed solely to differences in the conditional probabilities, holding the marginals constant. Clearly, this decomposition is only possible because the M can be expressed as the weighted average of local scores. In the example, the decomposition results in three terms, one for each occupation. Table 2 shows the results for the detailed structural decomposition. Occupation one and three are responsible for the increase in structural segregation, while in occupation 2, local segregation is low and almost unchanged. In more realistic settings with a greater number of units, the local segregation scores could now also be grouped by occupational major group or another characteristic (e.g., wage levels of occupations), if individual occupations are not of much interest. The sources of an increase or decrease in structural segregation, net of any marginal confounding, can thus be precisely understood.

Appearance and Disappearance of Units

Until now, we assumed that at both points in time, all units and groups have nonzero counts. However, this assumption is often not met in practice. In the case of school segregation, schools may have closed down and new schools may have opened. In the case of occupational segregation, some occupations may have vanished and new occupations have become established. Capitalizing on the decomposition properties of the M, the approach used here can be extended to account for the effects of adding or removing units and groups.

Assume the simple case that in a labor market of five occupations, two occupations become obsolete:

$$t_1 : \begin{bmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \\ 5 & 15 \\ 15 & 5 \end{bmatrix} \rightarrow t_2 : \begin{bmatrix} 8 & 23 \\ 23 & 8 \\ 19 & 19 \end{bmatrix}.$$

In this scenario, the workers from the vanished occupations were distributed across the remaining occupations, so that there are still 50 men and women each. Between t_1 and t_2, the M declines from 0.105 to 0.076. Is this purely an effect of the workers being redistributed? Or were the occupations that vanished more segregated than the occupations that remained.

To answer this question, define the set $S = \{1, 2, 3\}$ for the three remaining occupations, and $D = \{4, 5\}$ for the occupations that vanish. The sets S and D define “super-units” that are composed of individual units, and the share p_D is the proportion of workers in set D at t_1. The goal is to decompose $M(t_1)$ into the contribution of the occupations that vanish and those that continue to exist, which can be done using the general form of the between-within decomposition of M (Mora and Ruiz-Castillo 2011). Total segregation thus equals the between-super-unit M plus the weighted M within the two matrices defined by the two super-units, that is,

$$M(t_1) = \underbrace{M\begin{pmatrix} 30 & 30 \\ 20 & 20 \end{pmatrix}}_{\text{between vanished/remaining}} + \underbrace{p_D M\begin{pmatrix} 5 & 15 \\ 15 & 5 \end{pmatrix}}_{\text{within vanished}} + \underbrace{(1 - p_D) M\begin{pmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \end{pmatrix}}_{\text{within remaining}}.$$

Then solve for the last M term, which we call M^*:

$$M^*(t_1) = M\begin{pmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \end{pmatrix} = \frac{1}{1 - p_D}\left[M(t_1) - M\begin{pmatrix} 30 & 30 \\ 20 & 20 \end{pmatrix} - p_D M\begin{pmatrix} 5 & 15 \\ 15 & 5 \end{pmatrix}\right]$$
$$= 0.087 = \frac{1}{0.6}[0.105 - 0 - 0.4 \times 0.131]$$

This expression summarizes the mechanical effect of dropping occupations on the M index. To arrive at the "reduced M" on the left-hand side, we subtract from M all the sources of segregation that are due to the vanished occupations only, which consists of a "between" and a "within" term. The between term summarizes how strongly the gender composition of the vanished occupations deviates from the remaining occupations, in total, while the within term summarizes how much segregation there is *within* the vanished occupations. The division by $1 - p_D$ has the effect of scaling the other occupations' proportions upward.[5]

$M^*(t_1)$ will be larger than $M(t_1)$ when the occupations that vanish were less segregated compared to the remaining occupations and will be smaller in the opposite case. In this case, removing occupations 4 and 5 from t_1 reduces the M from $M(t_1) = 0.105$ to $M^*(t_1) = 0.087$. The "reduced M" can now be compared to the situation at t_2 using the regular IPF method. The approach outlined here thus amounts simply to a comparison of only those units that overlap across time points. However, an advantage of the M, which neither the H nor other indices have, is that there is an intuitive interpretation for the "missing" units.

Applying the decomposition to the example above gives the following:

$$M(t_2) - M(t_1) = \Delta_{\text{removals}} + \Delta_{\text{marginal}} + \Delta_{\text{structural}}$$
$$0.076 - 0.105 = -0.017 + -0.006 + -0.006 = -0.029.$$

In total, about 60 percent of the decline in segregation can be attributed to the effect of removing occupations 4 and 5. The remaining decline is equally due to changes in the marginals and to structural changes.

For simplicity, the example was only concerned with the removal of units, but additional units, such as newly arising occupations, can be handled in exactly the same way.

Summary of Decomposition Approach

The full, five-term decomposition of change between two segregation indices is thus:

$$\begin{aligned} M(t_2) - M(t_1) = \Delta_{\text{appearing}} + \Delta_{\text{disappearing}} + \Delta_{\text{marginal-units}} \\ + \Delta_{\text{marginal-groups}} + \sum_{u \in t_1 \cap t_2} \Delta_{u,\text{structural}}. \end{aligned} \quad (11)$$

For most segregation problems, equation (11) is the minimum that is required to robustly understand changes in segregation because the possible sources of change may point in opposite directions. Large changes in the marginals may hide worsening segregation at the structural level or improvements in structural segregation might be overwhelmed by changes in the marginals.

Often, it is also of interest to compare *several* points in time or across space, and not just two. In this case, one point can be set as the reference point, with the decomposition then comparing all other points to the focal point in time. In a time series of occupational segregation, the first or the last point are obvious candidates for the reference point, while in a ranking of occupational segregation by cities the city with the median occupational segregation could be a good candidate.

Note also that this procedure can be used to decompose any M index. Because the cross-sectional decomposition of an M index again yield M indices, their change can also be studied over time. For instance, when studying occupational segregation, one might be interested in the change not only in the total M but also for the partial M indices that define segregation within major occupational groups. (This will be done in the example below.) The total M admits to the following decomposition, assuming K major groups:

$$M = M_{\text{between}} + \sum_{k=1}^{K} p_k M_k, \tag{12}$$

where M_{between} refers to the gender segregation between the occupational major groups, p_k is the proportion of major group k such that $\sum_k p_k = 1$, and M_k is the segregation within major group k. When change is observed over time, the $k + 1$ M indices defined in this decomposition can then be studied using the procedure outlined here.

Example 1: Occupational Segregation

To consider the practical value of the above, I study occupational gender segregation in the United States between 1990 and 2016. IPUMS provides harmonized occupational codings based on the 1990 Census occupational codes for this period (Ruggles et al. 2018). The sample has been selected to comprise the employed, civilian population aged 16–66 with nonmissing occupations. The occupational codes for 1990 were grouped into nine major groups (see Table 3).

When comparing occupations over time, two problems arise. First, the degree to which fine-grained occupations are recorded changes over time, and this is often a problem induced by the harmonization efforts. For instance, "sociologists" are not coded separately in 2000–2016 but are available as a separate code in 1990. Second, occupations may vanish or new occupations

Table 3. Descriptive Statistics.

	1990	2000	2010	2016
Sample size (in 1,000)	5,917	6,542	1,443	1,441
A. Number of occupations				
Number of occupations	369	336	330	319
Appearing occupations		0	0	0
Disappearing occupations		33	6	11
B. Labor force participation (%)				
Female	46	47	48	48
C. Distribution of occupational major groups (%)				
Managerial	12	12	13	14
Professional	13	16	17	18
Technical	4	4	4	4
Sales	12	11	11	11
Administrative	16	16	14	13
Service	13	14	17	17
Farming, forestry	2	2	2	2
Production, craft	11	11	9	9
Operators, laborers	16	14	12	12
D. Female labor force by major groups (%)				
Managerial	43	44	45	46
Professional	54	57	59	59
Technical	46	48	49	48
Sales	49	50	51	51
Administrative	78	74	72	70
Service	57	59	60	59
Farming, forestry	17	18	17	19
Production, craft	8	10	10	11
Operators, laborers	27	25	20	20

may appear. "Stenographers," for instance, are no longer coded in later years, and this is probably because they no longer exist as a recognizable occupation. In many cases, it is hard to distinguish whether the problem is one of harmonization or one of disappearing occupations. For the purpose of this example, we will make the simplifying assumption that the harmonized occupations that are coded in each year represent recognizable, established occupations.

Descriptive Statistics and Total Segregation

Table 3 contains descriptive statistics by year. Panel A shows the number of unique occupations that are available in each year, along with the number of categories that appear and disappear in each year. Panels B and C document a well-known pattern of occupational change, both in terms of female labor force participation and in terms of a changing occupational distribution. Panel D shows that there is considerable heterogeneity in terms of female labor force participation across occupational groups and a heterogeneous pattern of change. In most occupational groups, female labor force participation increased, while in the administrative and operators/laborers major groups, the share of women declined.

We calculate the M and the H for the total labor force, as well as separately for each major occupational group. This is based on the decomposition of the M into between and within-cluster terms, as in equation (12). In this case, the between-group term measures the segregation that is induced by the major occupational groups alone, while the within terms measure the segregation of detailed occupations within each major group. Because the number of observations are in the millions, bootstrapped standard errors are negligible ($<.0005$) and therefore not shown.

The results are shown in Figure 1. Overall gender segregation, shown in the top panel, declined by 15 percent from 1990 to 2016 for the H and the M.[6] In 1990, the H was at 31 percent and declined to 26 percent by 2016. The between term also declined, which means that major occupational groupings are becoming less informative about gender composition over time. However, major occupational groupings account for a large amount of overall gender segregation (45 percent of total segregation in 1990 and 42 percent in 2016).

While overall segregation declined, the within terms reveal some heterogeneity. In most major groups, gender segregation declined. In others, notably farming and forestry as well as production and craft occupations, gender segregation increased strongly. This heterogeneity suggests that it is worthwhile to study major groups separately.

Decomposition of Change

Many segregation analyses would stop at this point. Using the decomposition properties of the M, as well as the decomposition of change developed in this article, we can go further and explore the patterns in more detail. To simplify the analysis of change, we focus on the changes between 1990 and 2016, without considering the intermediate years. Because no new occupations appear over time in this example, the total difference of any M term is thus decomposed into four components: the effect of those occupations that are removed, the effect of the changing occupation marginal distribution, the effect of the changing gender marginal distribution, and the total structural

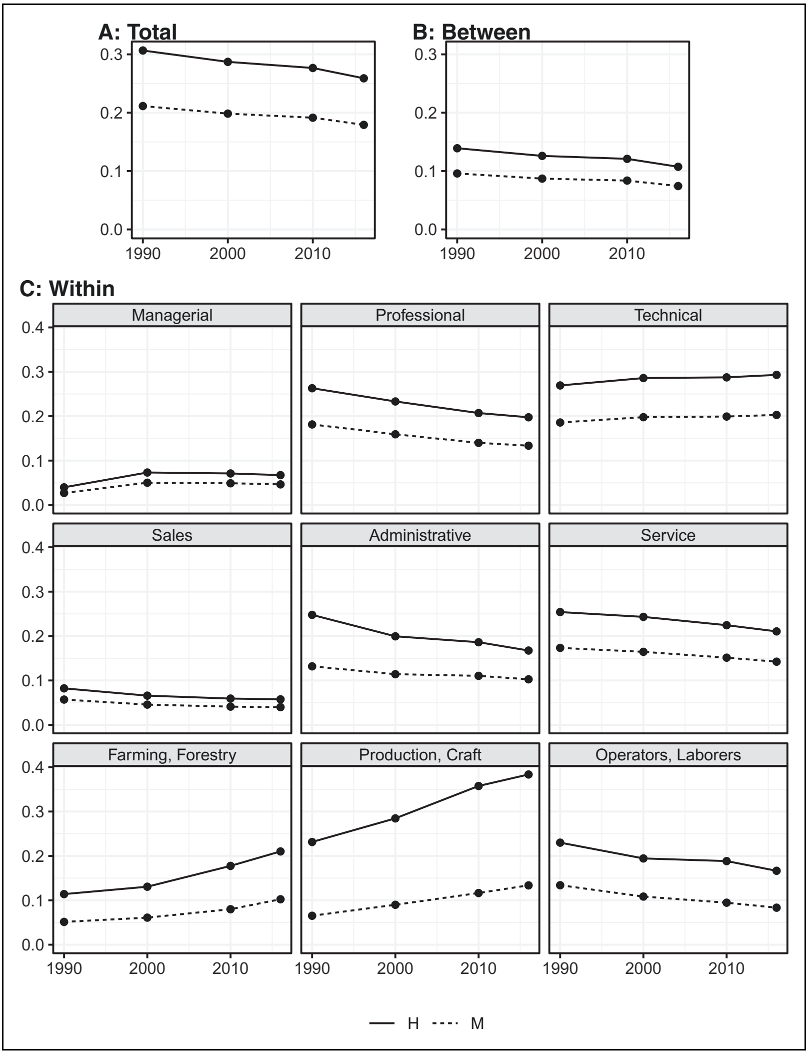

Figure 1. Occupational gender segregation, 1990–2016.
Panel A shows total segregation by gender and detailed occupations. Panel B shows segregation between gender and major occupational groups. Panel C shows within-major-group gender segregation by detailed occupations.

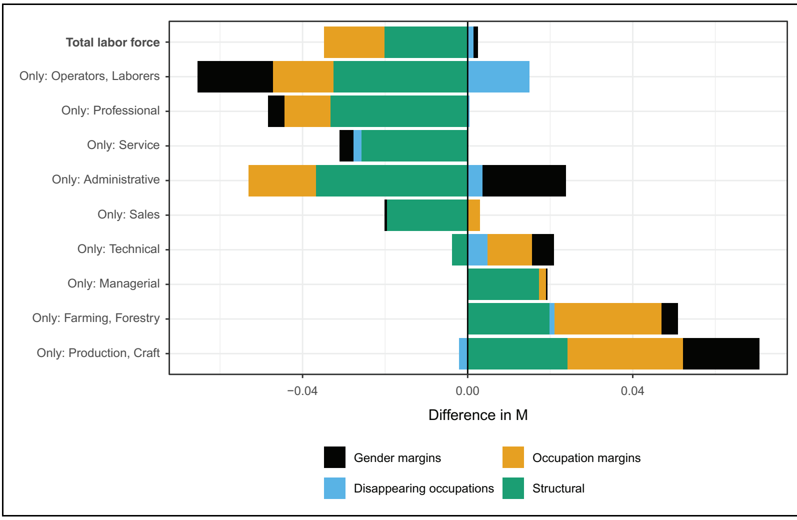

Figure 2. Decomposition of change.

component. Figure 2 shows the results graphically (without the between term), while Online Appendix B (which can be found at http://smr.sagepub.com/supplemental/) contains the full decomposition table. Again, the standard errors obtained through bootstrapping are negligible.

For the total *M*, the decline can be attributed to the changing occupational structure—that is, the labor force has shifted toward occupations that are less segregated—and, for the most part, to structural decrease. The decline in structural segregation accounts for 62 percent of the total decline in segregation. Most analysts of occupational segregation would consider this a positive development: Segregation decline is mostly due to declines in structural segregation, and the shift toward less segregated occupations has contributed even further to the decline. If all of the decline were due to the changing occupational margins only, we would still find that the average worker experiences less gender segregation. However, we could not conclude that the association of certain occupations with certain genders has lessened.

Segregation declined in five out of the nine major groups, and the share of the structural component was high in all five groups (between 65% and 117%). Within the major group of operators and laborers, the occupations that disappeared were relatively less segregated than the ones that remain, which increased segregation. However, the large marginal and structural components offset this small increase.

Segregation increased for four major groups. Except for the managerial group, structural increase plays less of a role for these groups. For technical occupations, structural change was in fact negative, but the marginal changes, especially the effect of the changing occupational distribution, led to an increase in segregation. In farming and forestry and production and craft occupations, structural segregation increased, but the changes in the marginals had a larger effect on the increase in segregation than the structural change. For the managerial occupations, the increase in segregation is almost entirely due to a structural increase in segregation, which is worrisome. Overall, a rough pattern emerges: For those occupational major groups where segregation declined, it declined in large part because of a structural decrease in segregation. When segregation increased, it increased mostly because of changes in the marginal distributions—with the notable exception of managerial occupations.

The increasing labor force participation of women accounts for only a minor part of the overall segregation difference: Around 3 percent of the total change is explained by changing gender marginals. One might wonder why the sign of these effects does not correspond to the changing patterns of female labor force participation from Table 3. Shouldn't major groups in which women are rare show a decrease in segregation if the number of women increases? For instance, the female share of production and craft workers has increased from 8 percent to 11 percent, but this led to an expected *increase* in segregation. To understand why this is the case, consider the example of carpenters. In 1990, this occupation was 98.2 percent male, while the male share in the major group was 92.2 percent. This leads to a local segregation score for carpenters (within the major group) of $0.982 \cdot \log\left(\frac{0.982}{0.922}\right) + 0.018 \cdot \log\left(\frac{0.018}{0.078}\right) = 0.036$. In 2016, the share of male workers in the major group is 88.9 percent, which represents a reduction in the share of men of about 4 percent and an increase in the share of women of about 42 percent. After proportionally increasing the number of women and reducing the number of men, the *expected* share of carpenters that are men is now 97.4 percent. (To simplify, we only consider the forward adjustment here.) This leads to a counterfactual local segregation score for carpenters of $0.974 \cdot \log\left(\frac{0.974}{0.889}\right) + 0.026 \cdot \log\left(\frac{0.026}{0.111}\right) = 0.051$. This score is higher than before, although the number of women has increased. In this case, the expected effect of proportionally increasing the share of women within each occupation increases segregation because it emphasizes existing patterns of segregation even more. The effect of the changing patterns of female labor force participation thus depends on the existing association structure between

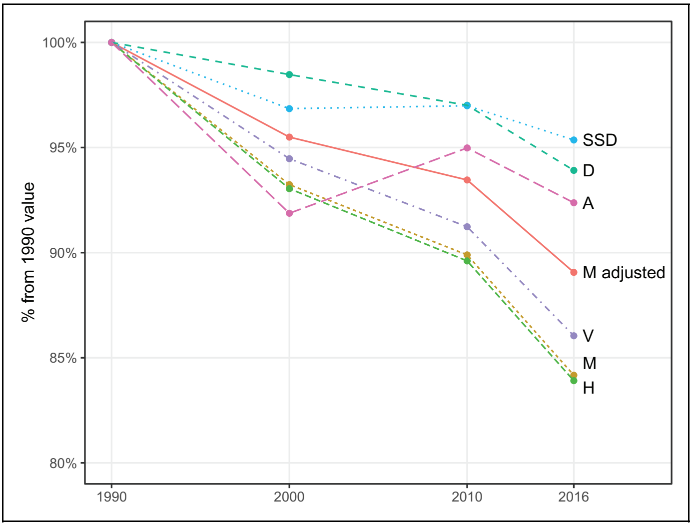

Figure 3. Comparison of the margins-adjusted *M* index with alternative indices.

occupations and gender. This shows that the marginal effects have to be interpreted as *expected* changes in segregation when the odds ratios are held constant.

Comparison With Other Indices

IPF makes it possible to create a time series of adjusted *M* indices that is not confounded by marginal changes. To do this, I choose 1990 as the reference year and adjust the other years (2000, 2010, 2016) toward the marginals of the year 1990. Alongside with the adjusted *M* index, I also calculate the observed *M* and *H* indices, and the three other indices discussed above (see Online Appendix A for formulae, which can be found at http://smr.sagepub.com/supplemental/).

The results for the five indices are shown in Figure 3. To ease comparison across the indices, the absolute numbers are transformed to be percentages of the 1990 values. First, it should be noted that all indices register a decline in segregation (although the *A* and SSD indices increase between 2000 and 2010). The structural decline, as calculated by the adjusted *M*, amounts to 10 percentage points of the 1990 value. The observed *M* and *H* "overstate" the decline, similarly to the *V* index. As seen in Figure 2, this is because the change in the occupational margins contributed to the decrease in segregation. Although the H is standardized, it gives essentially the same answer as the M. This is because the H is standardized by the gender distribution, which, however, had only a slight effect on segregation change. The effect of the occupation-margin dependency of the H is thus clearly visible here. The other indices underestimate structural change compared to the adjusted M. The differences between the margin-free A and adjusted M are due to the different occupational weights. The A weights each occupation equally, which makes is susceptible to extreme values for small occupations that arise from sampling variability (Watts 1998), which is a possible explanation for its more erratic movement compared to the other indices.

The adjusted M index has a clear interpretation and a clear advantage: It quantifies the amount of segregation that is purely due to changes in the odds ratios, net of any changes in the marginal distributions. It should be emphasized that the adjusted M is not a new segregation index, but just a regular M index, calculated on tables with identical margins. The main advantages of the decomposition will not be in the construction of an adjusted time series, as in Figure 3, but in the ability to more precisely pinpoint where the changes in segregation originate.

Example 2: Residential Segregation

A second, short example illustrates the advantages of decomposing structural segregation. These results make use of the Longitudinal Tract Database (Logan, Xu, and Stults 2014), which provides racial group counts for consistent Census tract boundaries. We just look at one example: The change in multigroup segregation in the borough of Brooklyn, New York City, from 2000 to 2010. Four racial groups are considered: Non-Hispanic whites, non-Hispanic blacks, Hispanics, and Asians.

Table 4 shows estimates of segregation by Census tracts in Brooklyn in 2000 and 2010, as well as the decomposition. The H declined from 0.437 to 0.398, which represents a decrease in segregation of about 9 percent. The difference in M values is then decomposed into the usual five terms. The main finding of this decomposition is that the decline in segregation is almost entirely due to structural change.

As a next step, the structural term is decomposed further to explore whether the declines in segregation are spatially clustered. We could use the terms $\Delta_{u,\text{structural}}$, as introduced in equation (10), but these terms are weighted by tract proportion. To show changes at the scale of the M index, we define instead the term ΔL_u, which is just the average change in local segregation scores, net of marginal changes:

Table 4. Decomposition of Change.

	Estimate
Index scores	
H in 2000	.437
H in 2010	.398
M in 2000	.552
M in 2010	.517
Difference in *M*	−.035
Difference decomposition	
Additions	.000
Removals	.000
Racial group margins	.000
Tract margins	.003
Structural	−.038

$$\Delta L_u = \frac{1}{2}[L_u(t_2) - L_u(t'_1)] + \frac{1}{2}[L_u(t'_2) - L_u(t_1)].$$

(This is simply equation [10], with the weights $p^{t_1}_{\cdot u}$ and $p^{t_2}_{\cdot u}$ dropped.) Recall that the local segregation scores are measuring how strongly each tract's racial group distribution deviates from Brooklyn's overall racial group distribution. If a tract has exactly the same racial group distribution as Brookyln, its local segregation score will be 0; if a tract's racial group distribution deviates from Brooklyn's racial group distribution, local segregation for that tract will be > 0.

Figure 4 shows a map of Brooklyn, with the tracts shaded according to the value of ΔL_u, that is, the expected difference in local segregation when the margins are held constant. As Table 4 has shown, the average structural decline in segregation was ≈ -0.04. Thus, if all tracts were affected in the same way by structural segregation, we would expect ΔL_u to be -0.04 for all neighborhoods. The map shows that this is clearly not the case. Instead, declines in structural segregation have been much more pronounced in some neighborhoods of central Brooklyn, such as Clinton Hill, Williamsburg, or Bedford-Stuyvesant, which are shaded in dark blue. In some eastern parts of Brooklyn (Canarsie and East New York), as well as southwest of Prospect Park (the area of Sunset Park), structural segregation has increased, often quite strongly. Note that these values can be interpreted at the scale of the *M*. Thus, an increase in structural segregation of 0.2 for the whole of Brooklyn would mean an increase in segregation of about 36 percent, given the

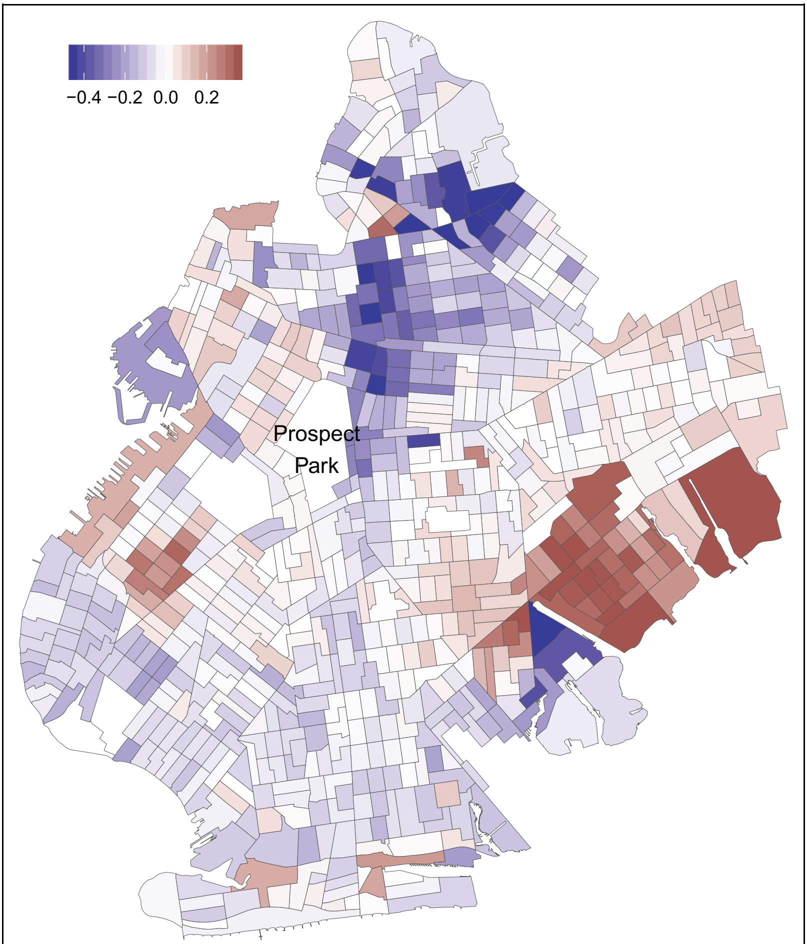

Figure 4. Tract-level differences in local segregation change ΔL_u, net of marginal changes.

baseline value of $M = .552$ in 2010. This shows that the differences we observe across tracts are quite substantial.

This analysis could now be continued in various ways. As the index was calculated as a multigroup index, a further analysis might be interested in racial group differences. Another approach is to correlate the changes in structural segregation with tract-level measures, such as income or racial composition. It seems that segregation declined most strongly in gentrifying neighborhoods, while segregation has increased especially in the eastern, disadvantaged neighborhoods.

Limitations

The major limitation of the M index is that it is not standardized between zero and one. This clearly is a disadvantage. However, as has been pointed out throughout the paper, the full decomposition of change is only possible with the M index, as (a) it is decomposable into a weighted average of local segregation scores, (b) "vanished" and "new" units have a clear interpretation, and (c) the symmetry of the decomposition requires that the index is neither standardized in terms of groups nor in terms of units (if that were the case, the respective marginal component would be underestimated). In practice, one might therefore prefer to use the H index to establish the absolute level of segregation, and report all M changes in terms of percentages. This has been done throughout the examples.

A limitation of the decomposition method is its relative complexity (certainly compared to a computation of a time series of segregation indices). This can be remedied through the use of the R package. Even with large tables and bootstrapping, the computation of the decomposition will be fast.

In the segregation literature, there has been some concern about segregation indices that are calculated on the basis of small unit sizes or small group proportions. For instance, Winship (1977) derived expressions for the expectation of the index of dissimilarity for a city with two racial groups. With 10 households per block and varying proportions of the racial groups, the expected value of the D under a random housing pattern will range from 0.246 to 0.387. This represents serious bias. For the M and the H, the expected values[7] for the same situation range from approx. 0.053 to 0.058, and from 0.076 to 0.178, respectively, which is an improvement (see also Fossett 2017:257-279). The reason for this improvement can be seen when the M is expressed in terms of the individual table cells. In this formulation, the observed value in each cell, p_{gu}, is compared to the expected value under independence (by multiplying the two marginal probabilities, $p_{g\cdot}$ and $p_{\cdot u}$):[8]

$$M = \sum_u \sum_g p_{gu} \ln \left(\frac{p_{gu}}{p_{g\cdot} p_{\cdot u}} \right). \tag{13}$$

Clearly, if p_{gu} is especially small, the logged ratio may be overly large. However, the expression is then weighted by p_{gu}, which leads to a relative decrease of the influence of the large ratio.

More generally, if one is concerned that in the problem at hand there may be zero segregation, and/or one deals with small group proportions or small unit sizes, one can take two steps to help remedy this problem: First, one can resort to the tool kit of classical statistics, such as Fisher's exact test or a chi-squared test. If these tests do not reject the null hypothesis of zero association between groups and units, then one can also conclude that there is no segregation. Second, one can use the observed marginal distributions to simulate random contingency tables, and compute the average segregation score for these tables. If the average simulated segregation score is > 0, the observed segregation score should be interpreted with caution. As a remedy, one could then combine units to arrive at a smaller contingency table. To check segregation bias for the H and M easily, the procedure has been implemented in the R package.

Finally, while the literature has devoted considerable effort to "purge" the influence of marginal differences from segregation indices, it should be noted that differences in the marginal distributions may often be the relevant social fact compared to differences in structural segregation. Consider the comparison of a labor market over time. At the beginning of the period, men comprise 80 percent of the labor market, and at the end of the period the labor market has a balanced gender composition. One could compare the absolute level of segregation over time, but it seems that in such an extreme case, the relevant difference lies in the starkly different demographic profiles. When there are large changes in the marginal distributions (as in this case), it is also questionable whether the marginal and structural changes can be interpreted as (causally) independent. The IPF method would adjust the majority gender distribution at the first time point toward the balanced situation at the second time point, assuming that the marginal changes did not affect structural change (or vice versa). The IPF method would still successfully calculate the contributions of marginal and structural changes to this trend. However, if it were true that marginal changes causally produced all structural change, then the contributions from the (unknown) true causal model would be different. The deeper point here is that changing marginal distributions can be an important part of segregative processes and that the summaries provided by a standardized segregation index should be used with caution when the margins are very different. An important empirical question to address in future work is how marginal changes and structural changes interact.

Conclusion

The article presented a general method to decompose *changes* in segregation levels. It has been shown that the difference between two *M* indices can be decomposed into marginal and structural changes, as well as into terms that account for the appearance and disappearance of units. Parts of the method are, in principle, applicable to any segregation index. However, the advantages of the *M* index became apparent when considering changing sets of units under study and when a further decomposition of structural terms is desired. The decomposition of the structural term into the contribution of individual units is especially useful, as it may reveal important heterogeneity in segregation change among the set of units. The change in structural segregation allows a more precise testing of hypotheses about the causes and effects of changing levels of segregation at the unit level, and this change will be net of any influence of the marginal distributions. The benefits of this approach have been illustrated in the two examples.

The method described here can be applied to a variety of problems. Given that the *M* is a multigroup index, no measures have to be taken to account for segregation problems with more than two groups. Thus, the decomposition can be applied to school and residential racial segregation (as in example 2) where the analysis extends beyond just the majority–minority group dichotomy. Other examples where the method might be usefully applied are workplace racial and gender segregation. As firms shut down and new firms are founded, these studies typically have to account for a changing distribution of units. Furthermore, it is likely that there are firm differences in the propensity to segregate by race and/or gender (e.g., large and small firms).

The examples in this article are focused on comparisons over time, but the method applies equally to comparisons across space. One useful application would be for comparisons of occupational segregation across countries or cities. One might suspect that observed differences in occupational segregation between cities are often due to marginal changes. For instance, a city with a large production sector will likely have higher gender segregation than a city with an employment profile skewed toward service occupations. One might suspect that the differences in segregation are just a consequence of the differences in the marginal occupational distribution. If one compares many cities or countries with each other, it will often not be feasible to compute all pairwise comparisons. Similar to regression analysis, one could instead choose one city as the “reference category” and compute comparisons against this reference city. Alternatively, one could also pool all of the city-specific data sets to capture overall segregation and then compute differences of each city compared to the overall average.

The *M* index and the decomposition can be applied to a much wider array of problems than are usually considered in segregation analysis. The *M* index, as any segregation index, is a measure of statistical association between two categorical variables and could thus be usefully applied to variables other than gender, occupation, racial groups, schools, or firms. For instance, the study of social mobility relates the parental class distribution to the class distribution of the children. It could prove insightful to apply entropy-based indices to this problem as well, as it would allow researchers to make statements about which classes contribute the most toward increases and decreases in social mobility.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research in this article has been supported by National Science Foundation grant SES-1423828. This research was supported by a grant from the American Educational Research Association which receives funds for its AERA Grants Program from the National Science Foundation under NSF award NSF-DRL #1749275. Opinions reflect those of the author and do not necessarily reflect those AERA or NSF.

ORCID iD

Benjamin Elbers https://orcid.org/0000-0001-5392-3448

Supplemental Material

The supplemental material for this article is available online.

Notes

1. The unit-margin dependency of the *D* has long been recognized. This led to the development of the size-standardized index of dissimilarity SSD, which, while not margin-dependent on the unit distribution, reintroduces a dependency on the group distribution.
2. Mora and Ruiz-Castillo (2011:161) identify a small number of papers that prefer the *M* index over the *H*. Beyond those, DiPrete et al. (2017) and Forster and Bol (2018) have used the *M* index in the context of school-to-work linkages.

3. Three caveats apply: First, the standardization is only limited to the range from zero to one when $U \geq G$, which is the case in most segregation problems. Alternatively, Mora and Ruiz-Castillo (2011) also define the H^* index. This index is defined by standardizing the M by the unit distribution entropy, that is, $H^* = \frac{M}{E(\mathbf{p}_{\cdot u})}$. This index is maximized when $E(\mathbf{p}_{u|g}) = 0$ for all groups, which is only true when all members of each group are concentrated at exactly one unit. This, of course, is not possible with two groups and more than two units. The H^* index is thus only appropriate when $G \geq U$, which for practical segregation problems is usually not the case. Second, the maximal value can only be reached if there are more subjects than there are units. As Carrington and Troske (1998: 239) write, "in a sample with 10 black workers and 20 firms, for example, evenness is unobtainable because it is impossible for each firm to get half a black worker." In many practical segregation problems, this is usually not a problem as there are more subjects than units. Third, the standardization only works when the size of the smallest group is larger than the smallest unit. For instance, consider a labor market of 200 women and 700 men distributed across three occupations of size 300. Even if the occupations are maximally segregated (i.e., two all-men occupations, and one occupation with 200 women and 100 men), the indices reach their maxima at $M = 0.32$ and $H = 0.6$. Whether such marginal constraints matter in practice depends on the concrete application.
4. The iterative proportional fitting procedure requires positive counts in each cell, which in practice may not always be in case. The canonical solution here is to replace zero counts with a very small number, for example, 0.0001.
5. This can be clearly seen by assuming that we drop one unsegregated occupation only (occupation 3 from t_1). Then, the expression simplifies to $M^* = \frac{M(t_1)}{1-p_D}$. This shows that the mechanical consequence on M when an unsegregated occupation vanishes depends only on the size of the occupation, p_D.
6. Because the M is sensitive to the number of categories, one might suspect that the higher gender segregation in 1990 is an artifact of measurement. The normalization of the H index corrects for the changing number of categories, as shown above. Either way, in this case the variation in the number of occupations is too small to matter: if we restrict the calculation to the 317 occupations that are available at all five points in time, the M and H values are within 1 percent of the values presented in Figure 1.
7. These values have been simulated using the "mutual_expected" function of the R package.
8. This equation also shows that the M is a "rescaled likelihood ratio test" (Card, Heining, and Kline 2013:983) and provides a natural way to relate the M index to other approaches that are concerned with the study of association in contingency tables.

References

Agresti, Alan. 2013. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: Wiley.

An, Brian P. and Adam Gamoran. 2009. "Trends in School Racial Composition in the Era of Unitary Status." Pp. 19-47 in *From the Courtroom to the Classroom*, edited by Claire Smrekar and Ellen Goldring. Cambridge, MA: Harvard Education Press.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics* 128 (3): 967-1015.

Carrington, William J. and Kenneth R. Troske. 1998. "Interfirm Segregation and the Black/White Wage Gap." *Journal of Labor Economics* 16 (2): 231-60.

Charles, Maria. 1992. "Cross-national Variation in Occupational Sex Segregation." *American Sociological Review* 57 (4): 483-502.

Charles, Maria and David B. Grusky. 1995. "Models for Describing the Underlying Structure of Sex Segregation." *American Journal of Sociology* 100 (4): 931-71.

Charles, Maria and David B. Grusky. 2004. *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Stanford, CA: Stanford University Press.

Deming, W. Edwards and Frederick F. Stephan. 1940. "On a Least Squares Adjustment of a Sampled Frequency Table." *Annals of Mathematical Statistics* 11 (4): 427-44.

Deutsch, Joseph, Yves Flückiger, and Jacques Silber. 2009. "Analyzing Changes in Occupational Segregation: The Case of Switzerland (1970– 2000)." Pp. 171-202 in "Occupational and Residential Segregation," Vol. 17 of *Research on Economic Inequality*, edited by Yves Flückiger, Sean F. Reardon, and Jacques Silber. Bingley, UK: Emerald.

DiPrete, Thomas A., Thijs Bol, Christina Ciocca Eller, and Herman G. van de Werfhorst. 2017. "School-to-work Linkages in the United States, Germany, and France." *American Journal of Sociology* 122 (6): 1869-1938.

Ferguson, John-Paul and Rembrand Koning. 2018. "Firm Turnover and the Return of Racial Establishment Segregation." *American Sociological Review* 83 (3): 445-74.

Forster, Andrea G. and Thijs Bol. 2018. "Vocational Education and Employment over the Life Course Using a New Measure of Occupational Specificity." *Social Science Research* 70:176-97.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." Pp. 1-102 in *Handbook of Labor Economics*, edited by Orley Ashenfelter and Card David. Vol. 4. Amsterdam: Elsevier.

Fossett, Mark. 2017. *New Methods for Measuring and Analyzing Segregation, Vol. 42 of the Springer Series on Demographic Methods and Population Analysis*. Cham, Switzerland: Springer.

Frankel, David M. and Oscar Volij. 2011. "Measuring School Segregation." *Journal of Economic Theory* 146 (1): 1-38.

Grusky, David B. and Maria Charles. 1998. "The Past, Present, and Future of Sex Segregation Methodology." *Demography* 35 (4): 497-504.

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

Karmel, T. and M. Maclachlan. 1988. "Occupational Sex Segregation—Increasing or Decreasing?" *Economic Record* 64 (3): 187-95.

Kitagawa, Evelyn M. 1955. "Components of a Difference between Two Rates." *Journal of the American Statistical Association* 50 (272): 1168.

Logan, John R., Zengwang Xu, and Brian J. Stults. 2014. "Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database." *The Professional Geographer* 66 (3): 412-20.

Mora, Ricardo and Javier Ruiz-Castillo. 2003. "Additively Decomposable Segregation Indexes. The Case of Gender Segregation by Occupations and Human Capital Levels in Spain." *The Journal of Economic Inequality* 1 (2): 147-79.

Mora, Ricardo and Javier Ruiz-Castillo. 2009. "The Invariance Properties of the Mutual Information Index of Multigroup Segregation." Pp. 33-53 in *Research on Economic Inequality*, Vol. 17, edited by Yves Flückiger, Sean F. Reardon, and Jacques Silber. Bingley, UK: Emerald Group.

Mora, Ricardo and Javier Ruiz-Castillo. 2011. "Entropy-based Segregation Indices." *Sociological Methodology* 41 (1): 159-94.

Reardon, Sean F. and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." *Sociological Methodology* 32 (1): 33-67.

Reardon, Sean F. and Ann Owens. 2014. "60 Years after Brown: Trends and Consequences of School Segregation." *Annual Review of Sociology* 40 (1): 199-218.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2018. "IPUMS USA: Version 8.0." doi:10.18128/D010.V8.0.

Shorrocks, Anthony F. 2013. "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value." *The Journal of Economic Inequality* 11 (1): 99-126.

Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam, the Netherlands: North Holland.

Theil, Henri. 1971. *Principles of Econometrics*. New York: Wiley.

Theil, Henri. 1972. *Statistical Decomposition Analysis*. Amsterdam, the Netherlands: North Holland.

Theil, Henri and Anthony J. Finizza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts." *Journal of Mathematical Sociology* 1 (2): 187-93.

Watts, Martin. 1998. "The Analysis of Sex Segregation: When Is Index Measurement Not Index Measurement?" *Demography* 35 (4): 505-08.

Watts, Martin. 2015. "The Use and Abuse of Entropy-based Segregation Indices." *6th Meeting of the Society for the Study of Economic Inequality*. http://www.ecineq.org/ecineq_lux15/FILESx2015/CR2/p217.pdf

Weeden, Kim A. 1998. "Revisiting Occupational Sex Segregation in the United States, 1910-1990: Results from a Log-linear Approach." *Demography* 35 (4): 475-87.

Weeden, Kim A. 2004. "Profiles of Change: Sex Segregation in the United States, 1910-2000." Pp. 131-78 in *Occupational Ghettos: The Worldwide Segregation of Women and Men*, edited by Maria Charles and David B. Grusky. Stanford, CA: Stanford University Press.

Winship, Christopher. 1977. "A Revaluation of Indexes of Residential Segregation." *Social Forces* 55 (4): 1058-66.

Author Biography

Benjamin Elbers is a PhD candidate in sociology at Columbia University. His research interests include segregation methodology, education, and labor markets. Recent articles have appeared in *Management Science and Social Forces.*