

# The Divergence Index: A Decomposable Measure of Segregation and Inequality\*

Elizabeth Roberto<sup>†</sup>

December 2, 2016

## Abstract

Decomposition analysis is a critical tool for understanding the social and spatial dimensions of inequality, segregation, and diversity. In this paper, I propose a new measure – the Divergence Index – to address the need for a decomposable measure of segregation. Although the Information Theory Index has been used to decompose segregation within and between communities, I argue that it measures relative diversity not segregation. I demonstrate the importance of this conceptual distinction with two empirical analyses: I decompose segregation and relative homogeneity in the Detroit metropolitan area, and I analyze the relationship between the indexes in the 100 largest U.S. cities. I show that it is problematic to interpret the Information Theory Index as a measure of segregation, especially when analyzing local-level results or any decomposition of overall results. Segregation and diversity are important aspects of residential differentiation, and it is critical that we study each concept as the structure and stratification of the U.S. population becomes more complex.

**Keywords:** segregation, inequality, diversity, measurement, entropy, decomposition

---

\*Thank you to Richard Breen, Scott Page, and Russell Golman for their valuable feedback. This research was supported by: the James S. McDonnell Foundation Postdoctoral Fellowship Award in Studying Complex Systems, and the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center.

<sup>†</sup>Department of Sociology, Princeton University, 108 Wallace Hall, Princeton, NJ 08544

Inequality, segregation, and diversity are complex phenomena that occur across multiple levels of social and spatial organization. In the United States, the racial and ethnic diversity of the national population has grown in recent decades, but there is considerable variation in the level of diversity and pace of change across communities and regions (Frey 2015; Hall, Tach, and Lee 2016; Lichter, Parisi, and Taquino 2015). Over the same period, income inequality has increased, and there are wide and persistent income gaps between ethnoracial groups (Bloome 2014; Burkhauser and Larrimore 2014).

Growing ethnoracial diversity provides new opportunities for intergroup contact, but recent declines in racial residential segregation at the neighborhood-level have been offset by increases in segregation between municipalities (Lichter et al. 2015). Income inequality is implicated as both a cause and consequence of segregation, but the extent to which socioeconomic advancement is associated with spatial integration varies by race group (Iceland and Wilkes 2006; Lichter, Parisi, and Taquino 2012).

Decomposable measures of inequality, segregation, and diversity are critical tools for understanding the social and spatial dynamics of these complex phenomenon. For example, they allow us to examine how much of the total income inequality in the U.S. occurs among individuals within particular groups (e.g. ethnoracial or educational) and how much occurs between the groups. Such an analysis allows to assess the extent to which group membership is a determinant of inequality (Breen and Chung 2015). Entropy-based measures have long been a staple of decomposition studies. Theil (1967, 1972; 1971) introduced the concept of entropy to the social sciences as a measure of population diversity (see also: Reardon and Firebaugh 2002; White 1986) and income inequality.

Despite having decomposable measures of both inequality and diversity, we lack a decomposable measure of segregation. The Dissimilarity Index (Duncan and Duncan 1955; Jahn, Schmid, and Schrag 1947; Taeuber and Taeuber 1965) is the most widely used measure of residential segregation, but it can not be decomposed into the segregation occurring within and between groups or places (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Theil 1972). The Information Theory Index (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Theil and Finizza 1971; White 1986) has become the gold standard for decomposition studies of segregation (Bischoff 2008; Farrell 2008; Fischer 2008; Fischer et al. 2004; Parisi, Lichter, and Taquino 2011). However, I argue that it is misleading to interpret the Information Theory Index as a measure of segregation – it measures the diversity of local areas relative to the region’s overall diversity, rather than measuring the difference between the local and overall proportions of each group.

The aim of this paper is to improve upon existing indexes by proposing a new decomposable measure of segregation and inequality: the Divergence Index. The Divergence Index summarizes the difference between two distributions. To measure racial residential segregation, the index measures how surprising the racial composition of local areas is given the overall racial composition of the region. The index equals zero, indicating no segregation, when there is no difference between the local and overall race proportions. Higher values of the index indicate greater divergence and more segregation. The Divergence Index can be decomposed into the segregation or inequality occurring within and between groups or spatial units, and it can be calculated for continuous and discrete distributions as well as for joint distributions, such as income by race. By creating an alternative measure, I provide a distinct lens, which enables richer, deeper, more accurate understandings of segregation and inequality.

I begin by comparing the concepts of segregation, inequality, and diversity, and noting the key distinctions between them. I then provide a brief review of popular measures of each concept – the Theil Index, Information Theory Index, and Dissimilarity Index. Next, I introduce my proposed

measure – the Divergence Index – and describe its unique features. Finally, I demonstrate the conceptual distinction between segregation and diversity with two empirical analyses. I decompose racial residential segregation and relative homogeneity between the city and suburbs in the Detroit metropolitan area, and show that interpreting the Information Theory Index as a measure of segregation leads us to opposite conclusions compared to the Divergence Index. I then analyze the empirical relationship between the two indexes in the 100 largest U.S. cities, and find a weak correlation between local-level results, providing further evidence that the indexes are measuring different concepts.

## Inequality, Segregation, and Diversity

Social inequality and segregation are tightly coupled concepts. Inequality refers to the uneven distribution of resources, opportunities, or outcomes across a population (e.g. individuals or groups). Segregation refers to the uneven distribution of the population across separate or distinct places, occupations, or institutions. Hence inequality and segregation both involve the uneven distribution of some quantity across units.

All measures of inequality and segregation have an implied or explicit comparative reference that defines equality or evenness (Coulter 1989), such as the uniform distribution of income across individuals, or the random distribution of individuals across neighborhoods. Measures evaluate the degree of inequality or segregation for a given distribution by measuring it against the comparative reference. For example, when a small portion of the population holds a large share of all income, the income distribution is unequal. When the racial composition of neighborhoods differs widely across a city, racial segregation is high.

In contrast to inequality and segregation, the concept of diversity describes the variety of “types” or groups in the population (Page 2007, 2011). Diversity indexes measure the number of groups and in what proportion they are represented. Diversity can also be measured in relative terms by comparing the diversity of one population or context to another, such as the racial diversity of neighborhoods compared to a city’s overall racial diversity.

What distinguishes the concept of diversity from segregation and inequality is its indifference to the specific groups that are over- or under-represented in a population. Diversity measures are only concerned with the variety or relative quantity of groups, whereas inequality and segregation measures are concerned with which groups (or which parts of a distribution) are over- and under-represented.

Measures of diversity can not distinguish between a setting in which the proportion of a minority group and a majority group match their proportions in the overall population, and one in which the proportions of the minority and majority groups are swapped. This is a characteristic of diversity (and relative diversity) measures, but it becomes problematic when a diversity index is used to measure segregation. As noted by Abascal and Baldassarri (2015), diversity indexes “flatten fundamentally hierarchical relations between groups. . . . As an analytic concept, ‘diversity’ (i.e. ‘heterogeneity’) not only sidesteps issues of material and symbolic inequalities, it masks the distinction between in-group and out-group contact” (p. 755).

For example, consider measuring gender segregation across academic majors at a university – the student population of the university is 75% women and 25% men, and engineering majors are 25% women and 75% men. The relative proportion of men and women differs in the engineering major and the overall student population, but both have a 3 to 1 mix of genders. If we interpret relative diversity as a measure of segregation, we would conclude that the engineering major is not

segregated because it has the same level of gender diversity as the university.

However, the gender proportions within the engineering major are surprising given the university context. Men are over-represented and women are under-represented relative to their overall proportions at the university. Rather than comparing gender *diversity*, we can measure segregation as the *difference* between the actual proportion of each gender in the engineering major and the overall student population. A major that has the same gender distribution as the university is not segregated. Given the striking difference between the gender proportions of the engineering major and the university, we would conclude that the major is segregated. As demonstrated in this example, measuring the concepts of diversity and segregation can lead us to opposite conclusions.

In the next section of the paper, I describe three common measures of diversity, inequality, and segregation: the Theil Index, Information Theory Index, and Dissimilarity Index. I then introduce the Divergence Index and compare it to the existing measures. I have restricted the discussion of existing measures to widely-used indexes that summarize or compare whole distributions. This excludes measures that target specific points of comparison within a distribution, such as a ratio of values for the 90th and 10th percentiles (Breen and Salazar 2011). With the exception of the Dissimilarity Index, all of the indexes are entropy-based measures.

## Existing Measures of Inequality, Segregation, and Diversity

### Entropy and the Theil Index

Entropy is commonly used in physics and information theory to measure the randomness of a system or the information content of a message (Coulter 1989; Cover and Thomas 2006; Shannon 1948; Theil 1967). Theil (1967, 1972; 1971) introduced the concept of entropy to the social sciences as a measure of population diversity (see also Reardon and Firebaugh 2002; White 1986) and income inequality.

Entropy is the amount of information needed to describe a probability distribution. If two outcomes are equally likely, there is high uncertainty about what the outcome will be and high entropy. If one outcome has a higher probability, there is less uncertainty about what the outcome will be and lower entropy.<sup>1</sup> Entropy measures the probability of an outcome ( $m$ ) occurring, weighted by its probability of occurrence ( $\pi_m$ ).<sup>2</sup> The entropy of each outcome ( $m$ ) is  $E_m = \log \frac{1}{\pi_m}$ . Weighting

---

<sup>1</sup>Entropy can be thought of as the uncertainty associated with the value of a random draw from a probability distribution. If an outcome has a probability of 100%, the entropy of the distribution is 0 – there is no uncertainty. If there are two equally likely outcomes, such as with a fair coin toss, the entropy of each outcome ( $E_m$ ) is 1 and the average uncertainty ( $E$ ) is 1, its maximum value. In other words, when two outcomes are equally likely, we have maximum uncertainty about what the outcome will be.

<sup>2</sup>The entropy equations can be defined using logarithms to any base. The base of the logarithm defines the units of the index (Shannon 1948; Theil 1972). Log base 2 ( $\log_2$ ) is typically used in information theory, which gives results in units of binary bits of information. It is common for inequality measures to use the natural logarithm ( $\ln$ ), which has the mathematical constant ( $e$ ) as its base.

each outcome by the probability of its occurrence, the overall entropy is:<sup>3</sup>

$$E = \sum_{m=1}^M \pi_m \log \frac{1}{\pi_m}$$

Interpreted as a measure of diversity,  $m$  indexes the groups (e.g. race or income group) in a population and the “probability of an outcome” is the proportion of each group. If all individuals in a population are associated with the same group, there is no diversity in the population. There is no uncertainty about a randomly selected individual’s group membership, and entropy is equal to 0. On the other hand, if individuals are evenly distributed among two or more mutually exclusive groups, there is maximum diversity in the population, and entropy is equal to 1.<sup>4</sup>

The properties of entropy have been well documented (e.g. Cover and Thomas 2006; Shannon 1948; Theil 1967), and are summarized in Table A5. It can be calculated for any number of groups, and it has known upper and lower bounds with substantive interpretations. Importantly, entropy and entropy-based measures are decomposable (see Appendix B for equations). However, entropy can only be calculated for discrete distributions, such as the proportion of each race group, but not continuous distributions, like the distribution of income.

Theil (1972; 1971) derived several indexes using the logic of entropy, such as his measure of income inequality – the Theil Index (Theil 1967). The Theil Index has many desirable properties, which are summarized in Table A5. In contrast to standard entropy indexes, the Theil Index can be calculated for continuous distributions. The Theil Index is written as:

$$I = \frac{1}{N} \sum_{i=1}^N \frac{x_i}{\bar{x}} \log \frac{x_i}{\bar{x}}$$

where  $x_i$  is the income of individual earners or groups of earners, and  $\bar{x}$  is the average income. When all incomes are equal (i.e. all individuals earn the mean income), there is no inequality and  $I$  is 0. The index measures the difference between the observed distribution and a single value, the mean. It is a special case of the generalized entropy class of measures, which also includes mean log deviation and half the coefficient of variation (Cowell 1980b, 1980a; Cowell and Kuga 1981; Shorrocks 1980, 1984).<sup>5</sup>

## The Information Theory Index

Theil also developed the Information Theory Index, another entropy-based measure. He used it to study racial segregation in Chicago public schools (Theil and Finizza 1971). The index has also been proposed as a measure of residential segregation (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; White 1986), and it has become the gold standard for decomposition studies of segregation (Bischoff 2008; Farrell 2008; Fischer 2008; Fischer et al. 2004; Parisi et al. 2011).

---

<sup>3</sup>Following standard usage, I define  $0 \log 0 = 0$ , because  $\lim_{x \rightarrow 0} (x \log x) = 0$ .

<sup>4</sup>For example, if there are equal numbers of men and women in a population, then the next person you meet is just as likely to be a man as a woman. There is maximum uncertainty because each group is equally probable, which indicates high entropy and high diversity. But in settings where there is a small minority of women, there is less uncertainty about what the gender will be of the next person you meet, and therefore there is low entropy and low diversity.

<sup>5</sup>It is also approximately equivalent to Atkinson’s inequality measure when the value of the weights in the social welfare function is close to 0 (Schwartz and Winship 1980).

However, I argue that the Information Theory Index measures relative homogeneity, and it is misleading to interpret it as a measure of segregation. It compares the diversity of local areas to the overall diversity of a region (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; White 1986), rather than measuring the difference between the local and overall proportions of each group.

For a single area ( $i$ ), the index measures the extent to which the area’s entropy ( $E_i$ ) is reduced below the region’s entropy ( $E$ ), standardized by dividing by the region’s entropy (Theil and Finizza 1971):

$$H_i = \frac{E - E_i}{E}$$

Or, equivalently, it is one minus the ratio of local diversity to overall diversity (Reardon and Firebaugh 2002).

$$H_i = 1 - \frac{E_i}{E}$$

The region’s index score is the weighted average of  $H_i$  across all local areas:

$$H = 1 - \sum_{i=1}^N \frac{\tau_i E_i}{TE} = 1 - \frac{\bar{E}_i}{E} \quad \text{or} \quad H = \sum_{i=1}^N \frac{\tau_i}{T} H_i$$

where  $T$  is the overall population count, and  $\tau_i$  is the population count for area  $i$ .  $H$  represents the relative reduction in the average entropy of components ( $\bar{E}_i$ ) below the maximum attainable entropy ( $E$ ) (Theil and Finizza 1971). Or, equivalently, it is one minus the ratio of average local diversity to overall diversity (Reardon and Firebaugh 2002).

The Information Theory Index typically ranges between 0 and 1. A value of 1 indicates that there is no diversity in local areas. A value of 0 indicates that all local areas are as diverse as the region. The minimum value can be less than 0, and Reardon and O’Sullivan (2004) interpret negative values of the index as indicating “hyper-integration,” which occurs when localities are more diverse, on average, than the region as a whole. In other words, groups are more equally represented in local areas than in the overall population. Additional properties of the Information Theory Index are summarized in Table A5.

## The Dissimilarity Index

The Dissimilarity Index (Duncan and Duncan 1955; Jahn et al. 1947; Taeuber and Taeuber 1965) is the most popular measure of residential segregation. It is also used to measure inequality, known as mean relative deviation (Reardon and Firebaugh 2002). As a segregation index, it measures the deviation of each location’s population composition from the overall population composition. Or, equivalently, it measures how evenly the population of each group is distributed across a region.

Unlike the previously discussed indexes, the Dissimilarity Index is not an entropy-based measure. The index is calculated as the absolute difference between the proportion of groups  $A$  and  $B$  in the  $i^{\text{th}}$  location, summed over all locations and divided by 2:

$$DI = \frac{1}{2} \sum_{i=1}^N \left| \frac{\tau_{iA}}{T_A} - \frac{\tau_{iB}}{T_B} \right|$$

where  $\tau_{iA}$  is group  $A$ ’s population count in location  $i$  and  $T_A$  is the total population of group  $A$ , and

likewise for group  $B$ .<sup>6</sup> If group A and B are distributed across locations in the same proportions, then there is no segregation. Segregation is measured as the extent to which the spatial distribution of group B deviates from group A.<sup>7</sup>

One of the appeals of the Dissimilarity Index is its straight forward interpretation. It is the proportion of one group that would have to move to another location to equalize the distribution of groups across locations (Duncan and Duncan 1955; Massey and Denton 1988). The moves must be from locations where the group is overrepresented to locations where the group is underrepresented (White 1986).

Despite its ease of calculation and interpretation, the Dissimilarity Index has a number of notable limitations, which have been well documented (Cortese, Falk, and Cohen 1976; Falk, Cortese, and Cohen 1978; Fossett and South 1983; Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Theil 1972; Winship 1978). I summarize the properties of the index and its limitations in Table A5. Given the focus of this paper, a chief drawback of the index is that it is not additively decomposable (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Theil 1972) – total segregation can not be decomposed into the segregation occurring within and between groups or spatial units. In the following section, I introduce my proposed measure – the Divergence Index – which addresses the limitations of the Dissimilarity Index.

## The Divergence Index

I developed the Divergence Index to address the need for a decomposable measure of segregation. The index is based on relative entropy, an information theoretic measure of the difference between two probability distributions (Cover and Thomas 2006). Relative entropy, also known as Kullback–Leibler (KL) divergence (Kullback 1987), shares many properties with entropy, but instead of characterizing a single distribution, it compares one distribution to another. The index can be used to measure inequality as well as segregation.

The Divergence Index measures the difference between a distribution,  $P$ , and another empirical, theoretical, or normative distribution,  $Q$ .<sup>8</sup> The index represents the divergence of a model ( $Q$ ) from reality ( $P$ ). It can be interpreted as a measure of surprise: How surprising are the observations ( $P$ ), given the expected value ( $Q$ )? Or, how surprising is an empirical distribution ( $P$ ), given a theoretical distribution ( $Q$ )?

---

<sup>6</sup>It can also be calculated as a weighted mean by weighing the absolute deviation for each component by its population size (White 1986), or rescaled by dividing by the maximum possible value of the index given the overall proportion of each group (Zoloth 1976).

<sup>7</sup>Although the index is typically used to measure segregation for two mutually exclusive groups, it can be rewritten to measure the segregation of multiple groups:

$$DI = \sum_{m=1}^M \sum_{i=1}^N \frac{\tau_j}{2TI} |\pi_{im} - \pi_m|$$

where  $\pi_{im}$  is group  $m$ ’s proportion of the population in location  $i$ ,  $\pi_m$  is group  $m$ ’s proportion of the overall population, and  $I$  is Simpson’s Interaction Index defined as  $\sum_{m=1}^M \pi_m (1 - \pi_m)$  (Morgan 1975; Reardon and Firebaugh 2002; Sakoda 1981).

<sup>8</sup>The index measures the entropy of  $P$  relative to  $Q$ , or the relative entropy of  $P$  with respect to  $Q$ .

For discrete probability distributions  $P$  and  $Q$ , the divergence of  $Q$  from  $P$  is defined as:<sup>9</sup>

$$D(P \parallel Q) = \sum_{m=1}^M P_m \log \frac{P_m}{Q_m}$$

The  $Q$  distribution defines the standard against which segregation or inequality is measured. It should represent the expected state of equality or evenness in the  $P$  distribution.  $Q$  can be theoretically determined or empirically derived. For example, it can be a standard probability distribution (e.g. a normal or uniform distribution), a prior state of the  $P$  distribution, or the mean of the observed data ( $P$ ). The index has known upper and lower bounds with substantive interpretations. The minimum value is 0, indicating no difference between  $P$  and  $Q$ . The maximum value can be less than or greater than 1.<sup>10</sup>

The Divergence Index is a non-symmetric measure of the dissimilarity between the two distributions (Bavaud 2009).<sup>11</sup> The divergence of  $Q$  from  $P$  does not necessarily equal the divergence of  $P$  from  $Q$ .<sup>12</sup> The asymmetry is an intentional feature of the measure. As Bavaud (2009) states, “the asymmetry of the relative entropy does not constitute a defect, but perfectly matches the asymmetry between data and models” (p. 57).

One of the unique features of the Divergence Index is that it can be calculated for either discrete distributions (relative entropy) or continuous distributions (differential relative entropy) (Cover and Thomas 2006). The desirable properties of both relative entropy and differential relative entropy have been well documented (e.g. Bavaud 2009; Cover and Thomas 2006). Many follow directly from the properties of entropy, while others depend on how the reference distribution is specified. (Table A5 summarizes the properties of the Divergence Index.) Like entropy, relative entropy is additively decomposable. For example, we can aggregate individuals into groups and calculate the inequality occurring within each group and between the groups. The sum of the within- and between-group components of inequality is equal to overall inequality.

Several other inequality measures have been derived from relative entropy and KL divergence. The Theil Index, described earlier, is a special case of relative entropy, which measure the difference between a single distribution and a summary statistic for that distribution – the mean. The theoretical state of equality is one in which everyone’s income is equal to the mean. The Theil Index belongs to the generalized entropy class of measures, which also includes mean log deviation, half the coefficient of variation, and the Atkinson Index (Breen and Salazar 2011; Cowell 1980b, 1980a; Cowell and Kuga 1981; Shorrocks 1980, 1984).<sup>13</sup> The Divergence Index can likewise be used to compare a distribution to a single value (see Appendix C), but also provides the flexibility to holistically compare two distributions.

---

<sup>9</sup>Following standard usage, I define  $0 \log 0 = 0$ , because  $\lim_{x \rightarrow 0} (x \log x) = 0$ .

<sup>10</sup>The Divergence Index can be standardized to have a range of 0 to 1 by dividing by its maximum value for a given population. However, standardizing the index transforms it from an absolute to a relative measure of inequality and segregation, and negates several of its desirable properties, including aggregation equivalence and independence. (See A5.)

<sup>11</sup>In contrast, entropy ( $E$ ) is symmetric in  $P(x)$  and  $1 - P(x)$ .

<sup>12</sup>It is possible to calculate a symmetric version of the index as the sum of  $D(P \parallel Q)$  and  $D(Q \parallel P)$ , but such an index does not capture the concepts of segregation and inequality that motivate this paper.

<sup>13</sup>The Theil Index is approximately equivalent to Atkinson’s inequality index with weights that are close to 0 in its social welfare function (Schwartz and Winship 1980).



The use of relative entropy and KL divergence was incorporated into the “relative distribution” method for measuring inequality (Handcock and Morris 1999). The relative distribution method compares distributions rather than summarizing their individual shapes, as with the Theil Index. The method also includes the median relative polarization index, which summarizes changes in the relative distribution. Relative distribution measures have been reviewed in detail elsewhere (Handcock and Morris 1998, 1999; Hao and Naiman 2010; Liao 2002), and have been used to analyze specific distributional shifts in income.

Recently, Bloome (2014) used KL divergence as a summary measure of racial disparity by comparing the distribution of income for white and black households. Sasson (2016) used divergence to study educational disparities in adult mortality. In the economics literature, divergence is used to study industrial localization and agglomeration (e.g. Mori, Nishikimi, and Smith 2005). More generally, divergence underlies popular statistical methods of model selection, including the Akaike Information Criterion (AIC) (Akaike 1974). For the remainder of the paper, I will focus on using the Divergence Index to measure residential segregation.

## Measuring Segregation with the Divergence Index

To study residential segregation, the Divergence Index measures the difference between the overall proportion of each group in the region (e.g. a city or metropolitan area) and the proportion of each group in local areas within the region. The overall proportion of each group in the region is the reference distribution ( $Q$ ), which represents the expected local proportion of each group if there is no segregation. The index asks: how surprising is the composition of local areas given the overall population of the region? If there is no difference between the local proportions of each group and the overall proportions, then there is no segregation in the region. More divergence between the overall and local proportions indicates more segregation.

Like the Dissimilarity Index, the Divergence Index measures how evenly the population of each group is distributed across locations in the region. However, the Dissimilarity Index follows a linear function and treats any deviation as equally surprising – the degree of segregation is directly proportional to the size of the deviation. In contrast, the Divergence Index follows a likelihood function and treats large deviations from evenness as much more surprising (i.e. segregated) than small deviations.<sup>14</sup>

The Divergence Index for location  $i$  is:

$$D_i = \sum_{m=1}^M \pi_{im} \log \frac{\pi_{im}}{\pi_m}$$

where  $\pi_{im}$  is group  $m$ 's proportion of the population in location  $i$ , and  $\pi_m$  is group  $m$ 's proportion of the overall population. If a location has the same composition as the overall population, then  $D_i = 0$ , indicating no segregation. To measure segregation spatially, we would replace  $\pi_{im}$  with  $\tilde{\pi}_{rim}$ , which is group  $m$ 's proportion of the spatially weighted population within a given distance  $r$  of location  $i$  (for examples, see Roberto 2015).

Overall segregation in the region is the population-weighted average of the divergence for all

---

<sup>14</sup>The greater the divergence of  $Q$  from  $P$ , the lower the probability of observing the local proportions ( $P$ ) if there is no segregation in the region ( $Q$ ).

locations:

$$D = \sum_{i=1}^N \frac{\tau_i}{T} D_i$$

where  $T$  is the overall population count, and  $\tau_i$  is the population count for location  $i$ . If all locations have the same composition as the overall population, then  $D = 0$ , indicating no segregation in the region.

The Divergence Index is additively decomposable, meaning that we can aggregate residential locations into districts and calculate the segregation occurring within and between the districts in a region. The sum of the within and between components of segregation is equal to overall segregation for the region. For example, to measure residential segregation for districts within a city, we rewrite the Divergence Index as the sum of between-district segregation and the average within-district segregation. The average within-district segregation for district  $j$  is:

$$D_j = \sum_{i \in S_j} \frac{\tau_i}{T_j} \sum_{m=1}^M \pi_{im} \log \frac{\pi_{im}}{\pi_{jm}}$$

where  $S_j$  is the set of locations in district  $j$ . The reference distribution,  $\pi_{jm}$ , is the population composition of district  $j$ , which is calculated as the population-weighted average of the group proportions for all localities ( $i$ ) within the district:  $\pi_{jm} = \sum_{i \in S_j} \frac{\tau_i}{T_j} \pi_{im}$ , where  $T_j$  is the population count for district  $j$ . The between-district segregation is:

$$D_0 = \sum_{j=1}^J \frac{T_j}{T} \sum_{m=1}^M \pi_{jm} \log \frac{\pi_{jm}}{\pi_m}$$

Total segregation is the sum of the between-district segregation ( $D_0$ ) and the average within-district segregation ( $D_j$ ):

$$D = D_0 + \sum_j \frac{T_j}{T} D_j$$

## Comparing the Divergence Index and Information Theory Index

The Divergence Index and Information Theory Index share many desirable properties, particularly their decomposability. However, the indexes measure different concepts and should not be used interchangeably. The Divergence Index measures segregation and inequality, while the Information Theory Index measures relative diversity. Each concept is interesting in itself and important to study, especially as the structure and stratification of the U.S. population becomes more complex. The concept of diversity concerns the variety or relative quantity of groups in a population. It is indifferent to the core concern of segregation – the degree to which specific groups are over- or under-represented in the local population.<sup>15</sup>

If the set of conditions that I outline in the next section are satisfied, it is possible to derive an

---

<sup>15</sup> Measures of diversity can not distinguish between a setting in which the proportion of a minority group and a majority group match their proportions in the overall population, and one in which the proportions of the minority and majority groups are swapped.

equivalence between the Divergence Index and Information Theory Index at the aggregate level of a city or region. However, no such equivalence exists at the local-level for locations or districts within a city or region.

## Equivalence between the Overall Indexes

The Information Theory Index,  $H$ , measures the ratio of local diversity to overall diversity. Whereas the Divergence Index,  $D$ , measures the difference between the local and overall group proportions. It is possible to derive an equivalence between  $H$  and  $D$ , but only if overall entropy ( $E$ ) is nonnegative and greater than or equal to the average local entropy ( $\bar{E}_i$ ):  $0 \leq E \geq \bar{E}_i$ .

If both conditions hold then, we can derive the equivalence between  $H$  and  $D$  by first rewriting the equation for  $D$  as:  $D = E - \bar{E}_i$  (Theil and Finizza 1971). Recall that we can write the equation for  $H$  as:  $H = \frac{E - \bar{E}_i}{E}$ . From this, we can derive the equivalence as:

$$H = \frac{D}{E} \quad \text{and} \quad D = HE$$

$H$  is equivalent to  $D$  standardized by  $E$ , or the ratio of  $D$  to  $E$ . Next, I describe the conditions that lead  $E$  to be negative or less than the average local entropy – if either occurs, then the equivalence provided above does not apply.

**Overall Entropy is Negative** The entropy of a discrete distribution is always nonnegative, however Cover and Thomas (2006:244) show that the entropy of a continuous distribution (called “differential entropy”) can be negative. For example, the differential entropy of a uniform distribution  $U(0, a)$  is negative for  $0 < a < 1$ . This occurs because the density of the distribution is  $\frac{1}{a}$  from 0 to  $a$ , and

$$E = - \int_0^a \frac{1}{a} \log \frac{1}{a} dx = \log a$$

Because  $a < 1$ , therefore  $\log a < 0$ . In contrast, both relative entropy and differential relative entropy (the discrete and continuous versions of the Divergence Index) are always nonnegative (Cover and Thomas 2006).

**Average Local Entropy is Greater than Overall Entropy** Theil and Finizza (1971) assumed that the population of schools were mutually exclusive in their study of racial school segregation in Chicago, IL, and they concluded that the average entropy of schools in a district ( $\bar{E}_i$ ) cannot be greater than the entropy of the district ( $E$ ). In other words, they concluded that the schools within a district cannot be more diverse, on average, than the district as a whole. Although this was a reasonable assumption for their specific case, it does not generalize to all contexts.

I find that average local entropy ( $\bar{E}_i$ ) can be greater than overall entropy ( $E$ ) if three conditions hold: if the overall population is not maximally diverse (i.e. at least one group is over- or under-represented), if any subunits have more diversity than the overall population (e.g. if there are local areas where groups are more equally represented than in the overall population), and if the subunits are not mutually exclusive.

The first two conditions are quite common when measuring segregation. The third condition –

non-exclusive subunits – arises when measuring segregation spatially. Spatial segregation measures, including the spatial version of the Divergence Index provided above, include a proximity-weighted contribution from nearby areas in each location’s population. This creates overlapping local environments or *ego-centric neighborhoods* (Lee et al. 2008; Reardon et al. 2009, 2008), which are not mutually exclusive. Non-exclusive subunits are also common in social network analysis, such as studying students with overlapping friendship networks.

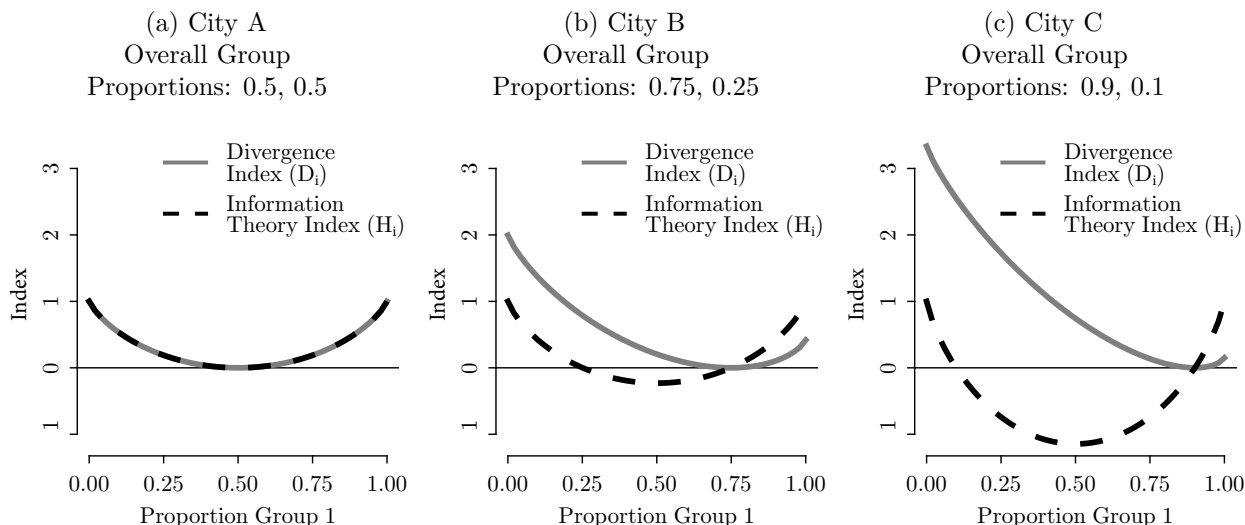
When the three conditions listed above occur, then average local entropy ( $\bar{E}_i$ ) can be greater than overall entropy ( $E$ ), and  $E$  can not be used to derive the equivalence between the Information Theory Index and the Divergence Index. Moreover, when  $\bar{E}_i$  is greater than  $E$ , then the Information Theory Index will be negative.<sup>16</sup>

## Comparing the Local Indexes

To illustrate the similarities and differences between the Divergence Index and Information Theory Index, Figure 1 compares the functional form of local results for three hypothetical cities. For the sake of the illustration, the two conditions listed above are both satisfied – overall entropy in the cities is positive, and average local entropy is not greater than overall entropy – and an equivalence exists between the city-level results, though not the local results.

Each city is divided into mutually exclusive local areas, and there are two groups in the cities’ populations. The proportion of each group varies across cities: 50-50 in city A, 75-25 in city B, and 90-10 in city C. The horizontal axes in Figure 1 show the proportion of group 1 in the local areas within each city. The vertical axes show the index score for local areas within the city across the full range of possible values for the local proportions of group 1. The solid lines plot the local index values for  $D_i$  and the dashed lines plot the local index values for  $H_i$ .

Figure 1: Comparing Local Values of the Divergence Index and Information Theory Index in Three Hypothetical Cities



<sup>16</sup>It is possible to observe nonnegative values of  $H$  when  $E$  is negative, but only if  $\bar{E}_i$  is also negative. It is also possible for  $H$  to be greater than 1, but only when measured for a continuous distribution and when either (but not both)  $E$  or  $\bar{E}_i$  is negative.

The minimum and maximum values of  $D_i$  and  $H_i$  vary across the three hypothetical cities in Figure 1. Local values of the Divergence Index,  $D_i$ , take their minimum value, which is always 0, when the local population composition is the same as the overall composition of the city.  $D_i$  reaches its maximum value when a city’s minority group is 100% of the local population. In a city where two groups are equally represented, like city A, it is just as surprising to observe a location where 100% of the residents are in group 1 as a location where 100% of the residents are in group 2. However, when there is a large majority group, as in cities B and C, it is more surprising to observe a location where all residents are in the minority group than a location where all residents are in the majority group. Further, it is more surprising to observe a location where all residents are in the minority group in a city C with a 10% minority population than in a city B with a 25% minority population. This is demonstrated in Figure 1 by comparing the local value of the Divergence Index in cities A, B, and C when the local proportion of the majority group (group 1) is 0.

Local values of the Information Theory Index,  $H_i$ , reach their maximum value when *any group* is 100% of the local population, regardless of the city’s population composition.  $H_i$  equals 0 when local diversity is the same as the city’s diversity, regardless of whether any group is over- or under-represented in the local population. For example, Figure 1b shows that  $H_i = 0$  when the proportion of group 1 in the local population is either 0.25 or 0.75, even though the proportion of group 1 in the city is 0.75.

$H_i$  takes its minimum value, which is typically less than zero, when a local area has an even mix of groups, regardless of the city’s diversity. The minimum value of  $H_i$  is a decreasing function of the city’s overall diversity. (Recall that  $H_i$  is 1 minus the ratio of local diversity to overall diversity.) Given the same level of local diversity, the value of  $H_i$  will be lower in a city with a less overall diversity than in a city with more overall diversity. This is demonstrated in Figure 1 by comparing across cities. The inflection point, or minimum value, of the function for  $H_i$  is 0 in city A where there is an even mix of groups, slightly negative in city B, and even more negative in city C, which is the least diverse city with a 90-10 mix of groups.

If local areas are marginally more diverse, on average, than the overall population, then  $H$  will be negative.<sup>17</sup> Reardon and O’Sullivan (2004) interpret negative values of  $H$  as indicating “hyper-integration” – each group is more equally represented in local areas, on average, than in the overall population.\* In contrast,  $D$  and  $D_i$  are never negative (Cover and Thomas 2006).

The results for the indexes are the *same* when there is an even mix of groups in the city population, as in city A (Figure 1a). If the proportion of each group in local areas is the same as the city proportions, then both indexes equal zero. If all local areas are monoracial, such that each group is either 100% or 0% of the local population, then both city-level indexes reach their maximum value. If the proportion of each group varies across local areas, then the measures would each find some degree of segregation or relative homogeneity. Moreover, the results for both indexes will be the same in the rare case that the overall population is maximally diverse.

The difference between the indexes is greatest when there is a small minority group in the population. At the extreme, if there is only one group present in the city and all local areas are monoracial,  $D$  and  $H$  give *opposite* results.  $H$  would show that the city is maximally homogenous (all  $H_i = 1$  and  $H = 1$ ) because there is no diversity in either the local areas or the city.<sup>18</sup> In

---

<sup>17</sup>Negative values of  $H$  occur when  $\bar{E}_i$  is greater than  $E$ . (Recall that  $H = 1 - \frac{\bar{E}_i}{E}$ .) In a previous section, I explained the conditions under which this occurs.

<sup>18</sup>Technically,  $H$  is undefined if there is only one group in the population, because  $H = 1 - \frac{0}{0}$ . If there are two groups in

contrast,  $D$  would find that the city is not at all segregated (all  $D_i = 0$  and  $D = 0$ ), because there is no difference between the composition of local areas and the city as a whole – each local area is a microcosm of the city.

The Divergence Index and Information Theory Index measure different concepts. The Information Theory Index measures *how diverse* the local and overall populations are, whereas the Divergence Index measures *how different* they are. The Information Theory Index is 1 minus the ratio of local diversity to overall diversity, and equals 0 when all local areas have the same level of diversity as the overall population. In contrast, the Divergence Index measures the difference between the local population composition and the overall population composition, and equals 0, indicating no segregation, when there is no difference between the local and the overall population compositions.

## Decomposing Segregation and Diversity in the Detroit Metro Area

Decomposition analysis is an ideal strategy for comparing how the segregation within and between different units or geographic areas contributes to overall segregation. Several studies have used the Information Theory Index to decompose segregation within and between communities, municipalities, or school districts (Bischoff 2008; Farrell 2008; Fischer 2008; Fischer et al. 2004; Parisi et al. 2011). However, I argue that such results should be interpreted in terms of relative homogeneity not segregation. To demonstrate the importance of this distinction, I use the Divergence Index and Information Theory Index to analyze racial residential segregation and relative homogeneity in the Detroit, MI metropolitan area.

The Detroit metro area is commonly cited as one of the most racially segregated places in the U.S. A large majority of the city’s residents are black (82%), while the surrounding area’s population is predominantly white (see Table D1).<sup>19</sup> I use population data from the 2010 decennial census aggregated at the level of census tracts (U.S. Census Bureau 2011)<sup>20</sup>, and compare white-black segregation and relative homogeneity results for the city of Detroit and the Detroit metro area. I then decompose overall segregation in the metro area into the segregation within and between the city of Detroit and the remainder of the metro area (the “suburbs”). I repeat the same decomposition for relative homogeneity and compare the results.

The city of Detroit is less diverse than the metro area, with overall entropy scores of 0.42 and 0.81, respectively. This contrast is transparent from Table D2, showing that the proportion of white and black residents is closer to parity in the metro area than in the city. Greater overall diversity provides the opportunity for greater local diversity in census tracts as well. However, despite the metro area’s greater overall diversity, the average local entropy scores for the city and metro are quite similar: 0.29 and 0.33. Compared to the metro area, census tracts in the city have levels of diversity that are, on average, more similar to the city’s overall diversity. This is reflected in the Information Theory Index scores of 0.32 for the city, compared to 0.59 for the metro area.

---

the population, the limit of  $H$  as the minority group’s population count approaches 0 (and  $E$  and  $\bar{E}_i$  approach 0) is 1.

<sup>19</sup>I use census data for mutually exclusive race categories, combined with Hispanic or Latino ethnicity. The Hispanic category includes all individuals who identified Hispanic or Latino as their ethnicity, along with any category of race. All other categories of race refer to individuals who identified as Not Hispanic or Latino.

<sup>20</sup>Census tracts are geographic units defined by the Census Bureau. They have an average population of 4,000 individuals and are intended to approximate neighborhoods. Most studies of residential segregation use census tract data.

Table 1: White-Black Segregation and Diversity in Detroit and the Metro Area

	Detroit	Metro Area
Overall Entropy ( $E$ )	0.42	0.81
Average Local Entropy ( $\bar{E}_i$ )	0.29	0.33
Information Theory Index ( $H$ )	0.32	0.59
Divergence Index ( $D$ )	0.14	0.48

Figure 2: White-Black Segregation and Diversity Between Detroit and the Suburbs

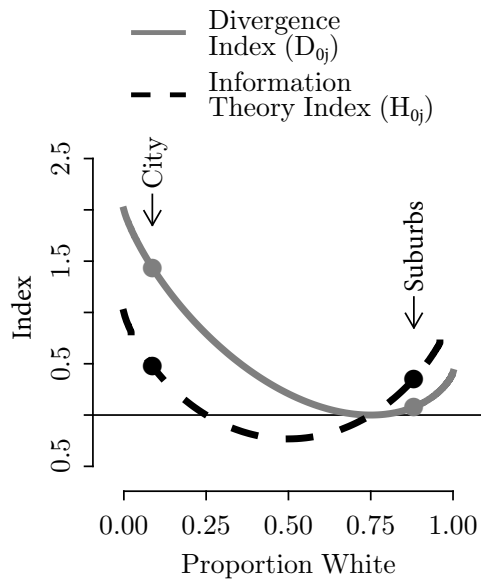


Table 2: Decomposition of White-Black Segregation and Diversity in the Detroit Metro Area (Proportion of Overall Index Score)

	Divergence Index	Information Theory Index
Overall Segregation	1.00	1.00
Between-Subareas	0.63	0.63
Detroit	0.50	0.13
Suburbs	0.14	0.50
Within-Subareas	0.37	0.37
Detroit	0.05	0.05
Suburbs	0.32	0.32

White-black segregation in the city of Detroit is low, 0.14, as measured with the Divergence Index. (See Table 1.) In contrast, segregation in the metro area is moderately high, 0.48. The city’s low level of segregation indicates that there is little difference, on average, between the composition of census tracts and the city’s overall population. The city’s population is predominantly black, and so is the local population of most tracts. In contrast, the higher segregation in the metro area indicates that the composition of census tracts differs greatly from the overall composition of the metro area.

To better understand the regional dynamics of segregation, I decompose overall segregation in the metro area into the segregation occurring between Detroit and the suburbs, and the segregation occurring among the tracts within each these subareas.<sup>21</sup> The between-subarea component of segregation measures how surprising the racial composition of each subarea is given the metro area’s overall racial composition. The within-subarea component of segregation measures how surprising the racial composition of tracts within each subarea is given the subarea’s overall racial composition. Total segregation for the metro area is the sum of the between-subarea segregation and the average within-subarea segregation. The total is equal to measuring segregation for all tracts in the metro area. In the same fashion, I decompose relative homogeneity into between- and within-subarea components with the Information Theory Index.

Table 2 reports the results for the subarea decomposition of the Divergence Index and the Information Theory Index. The table shows the proportion of the metro area’s index scores attributable to the between and within-subarea components. The decomposition of the Divergence Index shows that about two-thirds of the metro area’s segregation occurs between Detroit and the suburbs. Segregation among the tracts within each of the subareas accounts for the balance (37%). The decomposition of the Information Theory Index shows the same pattern.

The decomposition reveals that the largest differences in both population composition and diversity occur at the regional level – *between* Detroit and the suburbs. There is comparatively less difference at the local level – among the tracts *within* each subarea. However, if we take a closer look at the components of the between-subarea decomposition in Table 2, there is a stark difference in the two sets of results. Results for the Divergence Index show that Detroit contributes more to the between-subarea score than the suburbs, while results for the Information Theory Index show that Detroit contributes less than the suburbs.

Figure 2 shows the raw between-subarea index scores. The horizontal axis shows the proportion white, and the vertical axis shows the index scores of the subareas. The solid line shows the functional form of segregation measured with Divergence Index, and the dashed line shows relative homogeneity measured with the Information Theory Index. The points in each figure indicate the raw index score for each subarea – the city of Detroit and the suburbs. The raw scores are the values of each index prior to applying the weights for each subarea’s share of the metro population. In contrast, Table 2 reports the proportion of the total between-subarea score attributable to each subarea after weighting each subarea’s raw index scores but its share of the metro population (0.17 for Detroit and 0.83 for the suburbs). Figure 2 shows the pronounced difference between the between-subarea index scores for the city and suburbs measured with the Divergence Index, but not with the Information Theory Index, which is nearly the same for both the city and suburbs.

The between-subarea Divergence Index compares the difference between the subarea proportions and overall metro area proportions. The proportion white is 0.75 in the metro area, compared to 0.09

---

<sup>21</sup>Note that it is not possible to use the Dissimilarity Index for this decomposition because it is not additively decomposable.



in Detroit and 0.88 in the suburbs. (See Table D2.) From the perspective of the Divergence Index, 0.09 is a very surprising local proportion, more so than 0.88, given that the overall proportion white is 0.75. Therefore, there is greater divergence between the population compositions of Detroit and the metro area than between the suburbs and the metro area, and greater divergence indicates higher segregation. Detroit’s between-subarea segregation score is sufficiently higher than the suburbs that even after weighting each subarea’s score by its share of the metro population (0.17 for Detroit and 0.83 for the suburbs) Detroit’s contribution to between-subarea segregation is still larger than the suburbs.

Results for the Information Theory Index show an opposite trend: Detroit contributes less to overall segregation than the suburbs. White residents are over-represented in the suburbs and black residents are over-represented in Detroit, relative their metro proportions. But the city and suburban populations both have about the same level of diversity, and each has less diversity than the overall metro population. The Information Theory Index is concerned only with the mix of groups in each subarea relative to the metro, not the specific group proportions. Detroit contributes less than the suburbs not because their index scores differ, but because their scores are weighted differently when calculating the total between-subarea score – by their share of the metro population, which is much smaller for Detroit than the suburbs.<sup>22</sup>

This analysis demonstrates that it is problematic to interpret the decomposition results for the Information Theory Index as segregation. Within Detroit, census tracts are largely representative of the overall racial composition of the city. But there are stark differences in the racial composition of the city compared to the rest of the metro area. Detroit has a large majority of black residents whereas the suburban population is predominately white. It seems apparent that Detroit is segregated within the metropolitan context, but if we interpret the Information Theory Index as a measure of segregation, it would lead us to the opposite conclusion.

## Comparing Segregation and Diversity in U.S. Cities

In this section, I further demonstrate the distinction between measuring segregation and diversity by analyzing the empirical relationship between the Divergence Index and Information Theory Index in the 100 largest U.S. cities. I measure segregation and relative diversity for 4 combinations of ethnoracial groups<sup>23</sup> – white-black, white-Hispanic, white-black-Hispanic, and white-black-Hispanic-Asian – using tract-level data from the U.S. decennial census (U.S. Census Bureau 2011). I measure the city-level correlation between  $D$  and  $H$  and the local-level correlation between  $D_i$  and  $H_i$ . A weak correlation would provide evidence that the two indexes measure different concepts.

At the city-level, there is a strong correlation between  $D$  and  $H$ , ranging from 0.98 for white-black and white-Hispanic results, and 0.94 for white-black-Hispanic-Asian results (see Table 3). At the local-level, the correlation between  $D_i$  and  $H_i$  for tracts within each city is much weaker – the mean correlation across cities ranges from 0.10 for white-Hispanic results, and 0.39 for white-black-Hispanic-Asian results. Repeating the same analysis with block-level census data yields similar results (see Table 4).

---

<sup>22</sup>The Detroit population accounts for 17% of the metro area population. Detroit’s share is the same whether we include only the white and black population or the entire population.

<sup>23</sup>I use census data for mutually exclusive race categories, combined with Hispanic or Latino ethnicity. The Hispanic category includes all individuals who identified Hispanic or Latino as their ethnicity, along with any category of race. All other categories of race refer to individuals who identified as Not Hispanic or Latino.

Table 3: Correlation between the Divergence Index and Information Theory Index for Census Tracts within the 100 Largest U.S. Cities

	White-Black	White-Hispanic	White-Black-Hispanic	White-Black-Hispanic-Asian
Correlation of City-Level Results	0.98	0.98	0.96	0.94
Average Correlation of Tract-Level Results	0.22	0.10	0.38	0.39

Table 4: Correlation between the Divergence Index and Information Theory Index for Census Blocks within the 100 Largest U.S. Cities

	White-Black	White-Hispanic	White-Black-Hispanic	White-Black-Hispanic-Asian
Correlation of City-Level Results	0.96	0.92	0.92	0.88
Average Correlation of Block-Level Results	0.31	0.26	0.42	0.40

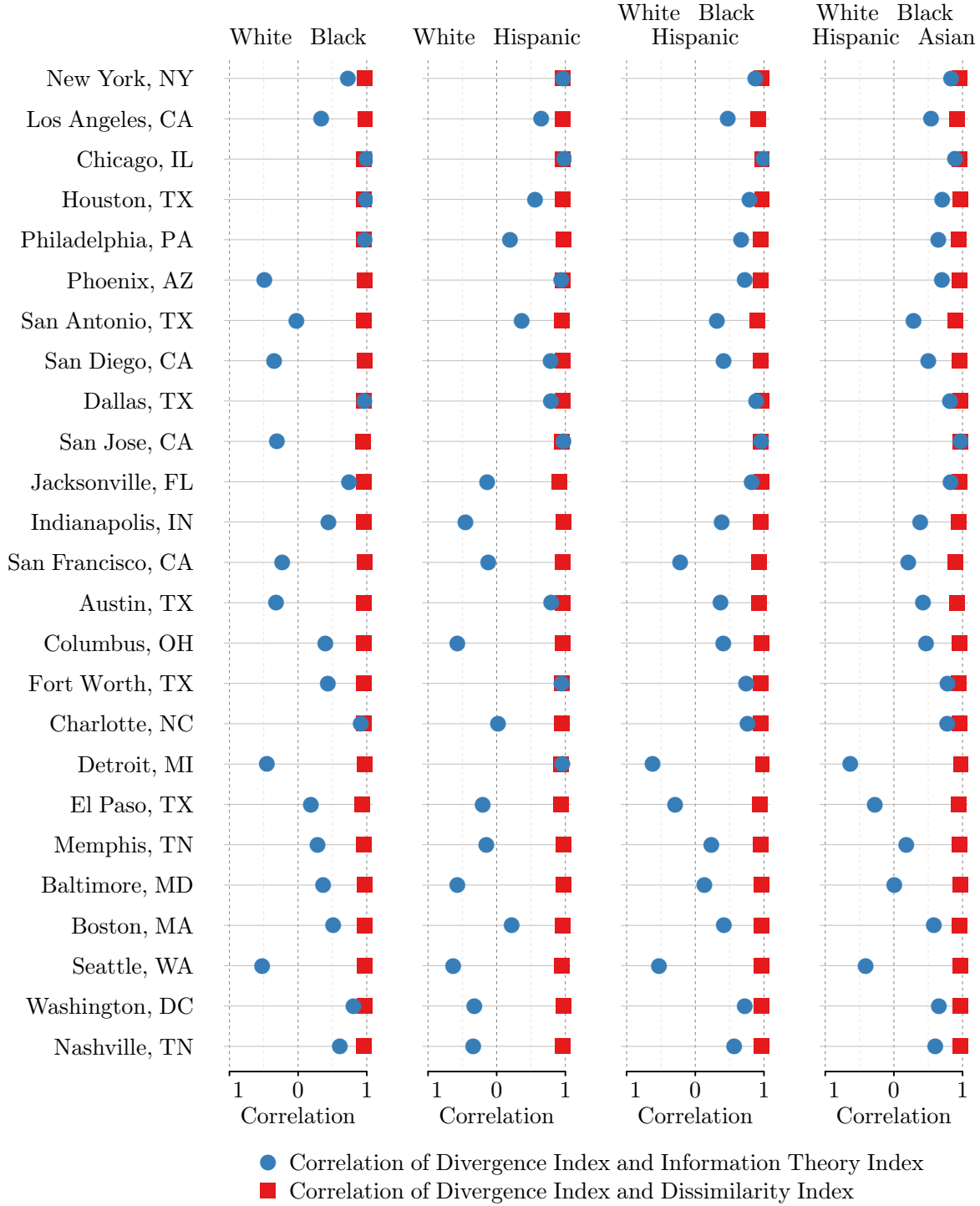
Figure 3 displays the tract-level correlation for each city for the Divergence Index and Information Theory Index as blue circles. For comparison, the correlation between the Divergence Index and Dissimilarity Index is displayed as red rectangles. The correlations between results for each combination of ethnoracial groups are shown in separate panels. Figure 3 includes the 25 largest U.S. cities, and Figure E1 in Appendix E shows the same information for all 100 cities.

The correlation between the Divergence Index and Information Theory Index ranges between -0.7 and 1 for all sets of ethnoracial groups at the block, tract, and city levels. In some situations, the two indexes yield similar results, but more often than not, their results lead to different conclusions.<sup>24</sup> In contrast, the correlation between the Divergence Index and Dissimilarity Index is consistently between 0.8 and 1. There are differences between the results for the two indexes, largely attributable to their different mathematical basis<sup>25</sup>, but the consistently strong correlation between their results is evidence that they are measuring the same concept.

<sup>24</sup>More than half of the correlations measured at the block, tract, and city level are between -0.5 and 0.5.

<sup>25</sup>As described earlier in the paper, the Dissimilarity Index follows a linear function and treats any deviation as equally surprising – the degree of segregation is directly proportional to the size of the deviation. In contrast, the Divergence Index follows a likelihood function and treats large deviations from evenness as much more surprising (i.e. segregated) than small deviations.

Figure 3: Correlation between Results for Census Tracts within the 25 Largest U.S. Cities



## Conclusion

Decomposition analysis is a critical tool for examining the social and spatial dimensions of diversity, segregation, and inequality. In this paper, I proposed a new measure – the Divergence Index – which addresses the need for a decomposable measure of segregation. Although previous studies have used the Information Theory Index to decompose segregation within and between communities, municipalities, or school districts (e.g., Bischoff 2008; Farrell 2008; Fischer 2008; Fischer et al. 2004; Parisi et al. 2011), I have shown that it measures relative diversity not segregation.

I illustrated the importance of the conceptual distinction between segregation and diversity by decomposing racial residential segregation and relative homogeneity between the city and suburbs in the Detroit metropolitan area. I found that census tracts within Detroit are largely representative of the overall racial composition of the city. But there are stark differences in the racial composition of the city compared to the rest of the metro area. Detroit has a large majority of black residents whereas the suburban population is predominately white. It seems apparent that Detroit is segregated within the regional context, but if we interpret the Information Theory Index as a measure of segregation, it would lead us to the opposite conclusion.

I further demonstrated the difference between the Divergence Index and Information Theory Index by analyzing the empirical relationship between the two indexes in the 100 largest U.S. cities. The correlation between overall results for the Divergence Index and Information Theory Index tend to be quite strong. However, the correlation of local results is much weaker, and can be near 0. The weak correlation between the local-level indexes provides further evidence that they are measuring different concepts.

Although the Divergence Index and Information Theory Index share many desirable properties, they measure different concepts and should not be used interchangeably. Segregation and relative diversity are both important aspects of residential differentiation, and it is important to study each concept, especially as the structure and stratification of the U.S. population becomes more complex. However, it is problematic to interpret the Information Theory Index as a measure of segregation, especially when analyzing local-level results or any decomposition of the overall results. By creating an alternative measure, I provide a distinct lens, which enables richer, deeper, more accurate understandings of segregation and inequality.

# Appendix A

## Desirable Properties of Measures

Previous research has identified a set of desirable properties for inequality and segregation measures (Allison 1978; Bourguignon 1979; Coleman, Hoffer, and Kilgore 1982; Jahn et al. 1947; James and Taeuber 1985; Morgan and Norbury 1981; Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Schwartz and Winship 1980; Taeuber and Taeuber 1965; White 1986). Measures are commonly evaluated with respect to how well they meet these criteria.

First, I review the criteria concerning the conceptual and methodological qualities of measures. They address how measures should respond to distributional changes (e.g. changes to the distribution of individual incomes or the population count of each group). I organize these criteria into three categories: features of the distribution, changes to the whole distribution, and changes within the distribution. Next, I review the desirable technical qualities and quantities of measures. This second set of criteria address how a measure should be calculated and interpreted.

## Conceptual and Methodological Qualities of Measures

Measures should be invariant to the following features of a distribution (Table A1):

Table A1: Criteria Concerning Features of the Distribution

Criteria	Description	Citations
Individual Cases	All cases should be treated the same.	Symmetry requirement (Bourguignon 1979)
Population Size	Proportionate increases or decreases in the size of the population have no effect on inequality.	Symmetry axiom for population (Bourguignon 1979; Sen 1973) Size invariance (James and Taeuber 1985; Reardon and Firebaugh 2002) Population density invariance (Reardon and O’Sullivan 2004)
Aggregations of Cases	Inequality should be invariant to the aggregation of components with identical compositions into a single unit, or dividing a single unit into components with the same composition.	Organizational equivalence (James and Taeuber 1985; Reardon and Firebaugh 2002) Location equivalence (Reardon and O’Sullivan 2004) Arbitrary boundary independence (Reardon and O’Sullivan 2004)

Measures should satisfy the following criteria about changes to the whole distribution of cases (Table A2):

Table A2: Criteria Concerning Changes to the Whole Distribution

Criteria	Description	Citations
Additive Increases	Additive increases to the whole distribution should reduce inequality, because it reduces the relative difference between cases.	Scale invariance (Allison 1978)
Proportionate Increases	Multiplying the whole distribution by a constant should have no effect on inequality, because it has no effect on the relative difference between cases.	Scale invariance (Allison 1978) Income-zero-homogeneity property (Bourguignon 1979) Composition invariance (Jahn et al. 1947; James and Taeuber 1985; Morgan and Norbury 1981; Taeuber and Taeuber 1965)

The proportionate increases criterion is known as composition invariance in the segregation literature, and it has long been a source of debate. James and Taeuber (1985) explain the principle of composition invariance with reference to racial segregation in schools: “proportional changes in the numbers of students of a specific race enrolled in each school do not affect the measured level of segregation” (p. 16). By their definition, a segregation index is not composition invariant if its value is a function of the overall population composition.

However, Coleman et al. (1982) argue that under certain definitions of segregation it is substantively appropriate to standardize an index by the overall population composition. One such example is defining a segregation index in terms of the extent of inter-group contact – no inter-group contact indicates maximum segregation, and contact proportional to the overall group proportions indicates zero segregation. In a population with a small minority group, we could expect less inter-group contact than in a population with equally represented groups, and the index adjusts to these expectations. Making such an index invariant to the population composition would distort its substantive meaning.

Reardon and O’Sullivan (2004) take a reasonable stance, stating that “the traditional composition invariance criterion espoused by James and Taeuber (1985) is less important than is ensuring that a measure of segregation has a sound conceptual basis. If a segregation index measures exactly that quantity that we believe defines spatial segregation, then the index will be composition invariant by definition” (p. 134).

Measures should satisfy the following criteria about changes within the distribution (Table A3):

Table A3: Criteria Concerning Changes within the Distribution

Criteria	Description	Citations
Transfers and Exchanges	<p>1. Any transfer from a unit (e.g. individual, group, or location) with more of the relevant quantity (e.g. income) to another with less should decrease inequality, provided that the rank order remains the same.</p> <p>2. Likewise, any transfer to a unit with more of the relevant quantity should increase inequality.<sup>26</sup></p>	<p>Pigou-Dalton principle (Dalton 1920; Pigou 1912)</p> <p>Inter-group transfers (James and Taeuber 1985; Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004)</p> <p>Inter-group exchanges (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004)</p>

### Technical Qualities and Quantities of Measures

In addition to desirable conceptual and methodological qualities of measures, a second set of criteria concern the technical qualities and quantities of inequality measures. The criteria – additive decomposability, and upper and lower bounds are summarized in Table A4.

Additive decomposability is a desirable property because it allows for a deeper analysis of the sources of inequality. The relative contribution of each component or group to overall inequality can be identified, and the inequality occurring within- and between-subpopulations can be analyzed (Bourguignon 1979).

Many measures are bounded between 0 and 1, with 1 indicating maximum inequality. If a measure has known upper and lower bounds, it can be rescaled to conform to a 0 to 1 range. However, rescaling the measure may shift the definition of inequality from absolute to relative. It is most important for the bounds of the index be known and interpretable.

<sup>26</sup>For example, from Allison (1978): “measures of inequality ought to increase whenever income is transferred from a poorer person to a richer person, regardless of how poor or rich or the amount of income transferred” (p. 868).

Table A4: Technical Qualities and Quantities of Measures

Criteria	Description	Citations
Additive Decomposability	Measures should be decomposable into the sum of inequality within and between sub-populations	Aggregativity and additivity (Bourguignon 1979) Decomposition (Allison 1978) Additive decomposability <sup>27</sup> (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004)
Upper and Lower Bounds	A measure should have known upper and lower bounds and each should have a substantive interpretation.	Scale interpretability (Reardon and O’Sullivan 2004) Upper and lower bounds (Allison 1978) Principle of Directionality (Fossett and South 1983)
Relative or Absolute Inequality	Relative and absolute measures are differentiated based on whether inequality is independent of, or a function of, the number of categories (respectively).	Sensitivity to the number of components (Waldman 1977)

### Summary of the Desirable Properties of Measures

Table A5 summarizes the desirable properties of the Dissimilarity Index, Theil’s Inequality Index, the Information Theory Index, and the Divergence Index. The rows of the table correspond to the properties detailed in the previous section, as well as the comparative standard used by the measure and which types of distributions it can be used with.

The Information Theory Index does not satisfy the proportionate increases criterion according to the definition of composition invariance described by James and Taeuber (1985) – the value of the index should not be a function of the overall population composition. However, Reardon and O’Sullivan (2004) show that the index does conform to other definitions of composition invariance. For instance, it is invariant to compositional changes as long as the relationship between local population diversity and overall population diversity remains constant.

Reardon and O’Sullivan (2004) show that the Information Theory Index satisfies the transfers and exchanges criteria when used to measure aspatial segregation. None of the indexes they evaluated satisfy the transfers criterion when used to measure spatial segregation. Spatial approaches often include a proximity weighted contribution from neighboring areas in each location’s population. This makes it difficult for any index to satisfy the transfers and exchanges criteria because the local populations are not mutually exclusive. They show that the Information Theory Index satisfies the

<sup>27</sup>For segregation measures, this includes additive organizational decomposability (Reardon and Firebaugh 2002), additive grouping decomposability (Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004) and additive spatial decomposability (Reardon and O’Sullivan 2004).



Table A5: Properties of the Measures

Criteria	Dissimilarity Index	Theil Index	Information Theory Index	Divergence Index
Individual Cases	✓	✓	✓	✓
Population Size	✓	✓	✓	✓
Aggregations of Cases	✓	✓	✓	✓
Proportionate Increases	X <sup>28</sup>	✓	X	✓
Additive Increases	✓	✓	✓	✓
Transfers and Exchanges	X <sup>29</sup>	✓	✓ <sup>30</sup>	✓ <sup>30</sup>
Additive Decomposability	X	✓	✓	✓
Upper and Lower Bounds	✓ <sup>31</sup>	✓	✓	✓
Relative or Absolute Inequality	Relative	Either	Absolute	Either
Comparative Standard	Evenness (mean of the distribution)	Evenness (mean of the distribution)	Randomness	Any
Distribution Types	Discrete with nominal categories	Continuous	Discrete	Discrete or continuous

<sup>28</sup>It is debatable whether or not the Dissimilarity Index satisfies the proportionate increases criterion. Cortese et al. (1976) found that it is sensitive to the minority group proportion, while others found no such association (James and Taeuber 1985; Lieberman and Carter 1982; Taeuber and Taeuber 1965). Reardon and colleagues (Reardon and Firebaugh 2002; Reardon and O'Sullivan 2004) find that it is only composition invariant when calculated for two groups.

<sup>29</sup>The Dissimilarity Index satisfies a weak form of the transfers and exchanges criteria (Reardon and Firebaugh 2002; Reardon and O'Sullivan 2004).

<sup>30</sup>The transfers and exchanges criterion generally only applies when components are mutually exclusive, as described in the text.

<sup>31</sup>The the Dissimilarity Index is bounded between 0 and 1, but the expected value of the index is greater than 0 (Cortese et al. 1976).

Criteria	Dissimilarity Index	Theil Index	Information Theory Index	Divergence Index
Citations	Bourguignon 1979; Cortese et al. 1976; Coulter 1989; Duncan and Duncan 1955; Falk et al. 1978; Fossett and South 1983; Jahn et al. 1947; James and Taeuber 1985; Lieberman and Carter 1982; Massey and Denton 1988; Morgan 1975; Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Sakoda 1981; Taeuber and Taeuber 1965; Theil 1972; Winship 1978	Allison 1978; Bourguignon 1979; Cowell 1980b; Cowell, Flachaire, and Bandyopadhyay 2013; Shorrocks 1980, 1984, 2012; Theil 1967, 1972	Reardon and Firebaugh 2002; Reardon and O’Sullivan 2004; Theil 1967, 1972; White 1986	Bavaud 2009; Cover and Thomas 2006; Cowell 1980b; Cowell et al. 2013; Magdalou and Nock 2011; Mori et al. 2005; Shorrocks 1980, 1984, 2012; Theil 1967; Walsh and O’Kelly 1979

exchanges criterion under certain general conditions (see Reardon and O’Sullivan 2004).

The entropy-based measures (Theil Index, Information Theory Index, and Divergence Index) can be defined using logarithms to any base. The selected base defines the units of the index (Shannon 1948; Theil 1972). Log base 2 ( $\log_2$ ) is typically used in information theory, which gives results in units of binary bits of information. It is common for inequality measures to use the natural logarithm ( $\ln$ ), which has the mathematical constant ( $e$ ) as its base.

Using a fixed log base, such as base 2 ( $\log_2$ ) or  $e$  ( $\ln$ ), entropy is an absolute measure. Results are a function of the number of groups in the population (Waldman 1977). Given a uniform distribution of groups (indicating maximum diversity), entropy is an increasing function of the number of groups. At first blush, this may seem undesirable, but it has the benefit of maintaining entropy’s aggregation equivalence and independence. This means that inequality calculated for a population of two groups is the same as if there were three groups in same population, but no individuals associated with the third type.

For discrete distributions, it may be preferable to use the number of groups as the base. The result is equivalent to dividing by the maximum entropy ( $\log M$ ), given by the number of groups ( $M$ ). With the number of groups as the log base ( $\log_M$ ), results are scaled to have the same maximum entropy no matter how many groups are in the population. This transforms entropy from an absolute to a relative measure of inequality. It allows for easier comparison across results with

different numbers of groups, but comes at the cost of one of the desirable properties of entropy – aggregation equivalence and independence.

For example, using  $\log_2$  to measure white-black-Hispanic residential segregation in a city with no Hispanic residents gives the same results whether all three races are included in the measure or only the two with population. This is not the case using  $\log_M$ , because results are scaled according to the number of groups included in the index. Which of these options is preferable depends on the analytic aim of the research, but it is important to be aware of this trade-off.<sup>32</sup>

## Appendix B

### Entropy Decomposition

Entropy-based measures are additively decomposable, which is a particularly desirable property (Theil 1972).<sup>33</sup> It is simple to aggregate (and disaggregate) the entropy for multiple groups and to decompose total entropy into the entropy occurring within- and between-groups. The entropy for each component ( $i$ ) is the sum of the entropy across groups within that component ( $m$ ):

$$E_i = \sum_{m=1}^M \pi_{im} \log \frac{1}{\pi_{im}}$$

The entropy for all components is the mean of the individual entropies, weighted by the relative size of each component:

$$\bar{E}_i = \sum_{i=1}^N \frac{\tau_i}{T} E_i$$

Theil (1972) showed that total entropy can be calculated for any subdivision of the population and written as the sum of a between-subdivision entropy and the average within-subdivision entropies. For example, if the groups are aggregated into supergroups ( $S_g$ ), where  $\Pi_{ig} = \sum_{m \in S_g} \pi_{im}$  is the proportion in each supergroup ( $g$ ) within component ( $i$ ). The entropy within supergroup  $g$  for component  $i$  is:

$$E_{ig} = \sum_{m \in S_g} \frac{\pi_{im}}{\Pi_{ig}} \log \frac{\Pi_{ig}}{\pi_{im}}$$

And the between-supergroup entropy is:

$$E_{i0} = \sum_{g=1}^G \frac{\Pi_{ig}}{\pi_i} \log \frac{\pi_i}{\Pi_{ig}}$$

---

<sup>32</sup>This choice does not affect results of the information theory index, because the log appears both in the numerator and denominator of the equation.

<sup>33</sup>The additivity of entropy comes from one of the properties of logarithms:  $\log(\pi_1 \cdot \pi_2) = \log(\pi_1) + \log(\pi_2)$

The total entropy for component  $i$  can then be written as the between-supergroup entropy ( $E_{i0}$ ) plus the average within-supergroup entropy ( $E_{ig}$ ):

$$E_i = E_{i0} + \sum_{g=1}^G \frac{\Pi_{ig}}{\pi_i} E_{ig}$$

## Appendix C

### Comparing the Theil Index and Divergence Index

Theil's inequality index ( $I$ ) and the Divergence Index ( $D$ ) both measure inequality relative to a defined standard. The Theil Index measures the difference between the observed shares of income across individuals or groups and a theoretical uniform distribution – one in which everyone's income is equal to the mean.

There is a straightforward equivalency between  $I$  and  $D$  for continuous distributions, such as income.<sup>34</sup> Theil's index can be written like the Divergence Index, where  $P_i$  is  $i$ 's share of total aggregate income,  $\frac{\tau x_i}{T\bar{x}}$ , and  $Q_i$  is the theoretical uniform share  $\frac{\tau}{T}$ :

$$\begin{aligned} D(P \parallel Q) &= \sum_{m=1}^M P_m \log \frac{P_m}{Q_m} \\ I &= \sum_{i=1}^N P_i \log \frac{P_i}{Q_i} \\ &= \sum_{i=1}^N \frac{\tau_i x_i}{T\bar{x}} \log \frac{\frac{\tau_i x_i}{T\bar{x}}}{\frac{\tau_i}{T}} \\ &= \frac{1}{T} \sum_{i=1}^N \frac{\tau_i x_i}{\bar{x}} \log \frac{x_i}{\bar{x}} \end{aligned}$$

If  $\tau_i = 1$  and  $T = N$ , then we get:

$$I = \frac{1}{N} \sum_{i=1}^N \frac{x_i}{\bar{x}} \log \frac{x_i}{\bar{x}}$$

We can see that  $I$  is a specific case of  $D$  applied to measuring income inequality, using uniform shares of income as the comparative standard.

---

<sup>34</sup>Moreover, the equivalency applies to any distribution for which a mean can be calculated, such as a discrete simplification of a continuous distribution.

## Appendix D

### Population Composition in the Detroit Metro Area

Table D1: Population by Race and Ethnicity in Detroit and the Metro Area

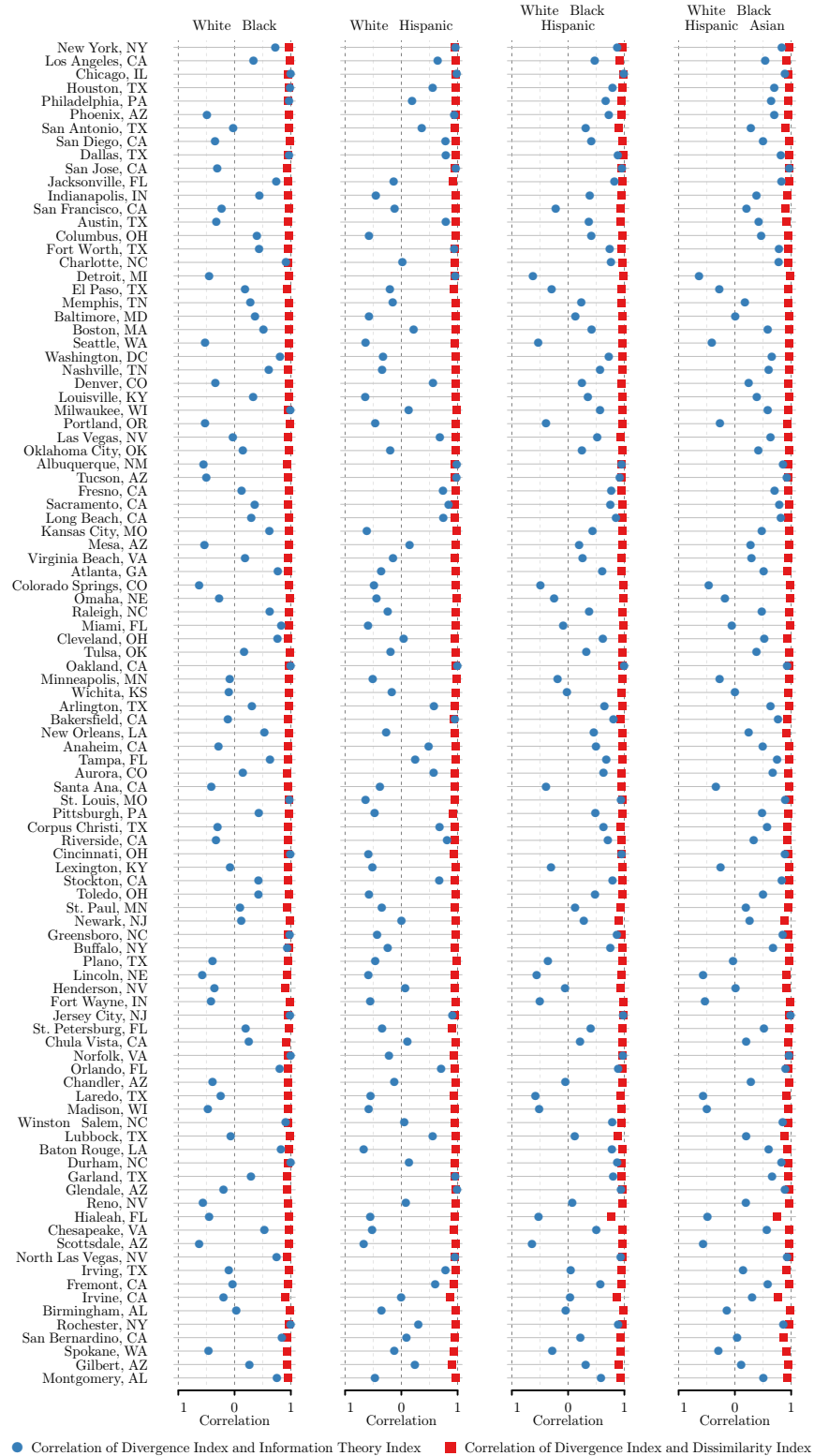
	Detroit	Metro Area
Total Population	713,777	4,296,250
White	7.8%	67.9%
Black	82.2%	22.6%
Hispanic	6.8%	3.9%
Asian	1.0%	3.3%
American Indian	0.3%	0.3%
Pacific Islander	0.0%	0.0%
Other Race	0.1%	0.1%
Multiple Races	1.7%	1.9%

Table D2: White and Black Population in the Detroit Metro Area

	Proportion White	Proportion Black
Metro Area	0.75	0.25
Detroit	0.09	0.91
Suburbs	0.88	0.12

# Appendix E

Figure E1: Correlation between Results for Census Tracts within the 100 Largest U.S. Cities



## References

- Abascal, Maria and Delia Baldassarri. 2015. "Love Thy Neighbor? Ethnoracial Diversity and Trust Reexamined." *American Journal of Sociology* 121(3):722–82.
- Akaike, Hirotugu. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19(6):716–23.
- Allison, Paul D. 1978. "Measures of Inequality." *American Sociological Review* 43(6):865–80.
- Bavaud, François. 2009. "Information Theory, Relative Entropy and Statistics." Pp. 54–78 in *Formal theories of information*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bischoff, Kendra. 2008. "School District Fragmentation and Racial Residential Segregation How Do Boundaries Matter?" *Urban Affairs Review* 44(2):182–217.
- Bloome, D. 2014. "Racial Inequality Trends and the Intergenerational Persistence of Income and Family Structure." *American Sociological Review* 79(6):1196–1225.
- Bourguignon, Francois. 1979. "Decomposable Income Inequality Measures." *Econometrica* 47(4):901–20.
- Breen, Richard and Inkwan Chung. 2015. "Income Inequality and Education." *Sociological Science* 2:454–77.
- Breen, Richard and Leire Salazar. 2011. "Educational Assortative Mating and Earnings Inequality in the United States." *The American Journal of Sociology* 117(3):808–43.
- Burkhauser, Richard V. and Jeff Larrimore. 2014. "Median Income and Income Inequality: From 2000 and Beyond." Pp. 1–35 in *Diversity and disparities: America enters a new century*. Russell Sage Foundation.
- Coleman, James S., T. Hoffer, and S. Kilgore. 1982. "Achievement and Segregation in Secondary-Schools - a Further Look at Public and Private School Differences." *Sociology of Education* 55(2-3):162–82.
- Cortese, Charles F., R. Frank Falk, and Jack K. Cohen. 1976. "Further Considerations on the Methodological Analysis of Segregation Indices." *American Sociological Review* 41(4):630–37.
- Coulter, Philip B. 1989. *Measuring Inequality: A Methodological Handbook*. Boulder : Westview Press.
- Cover, T. M. and Joy A. Thomas. 2006. *Elements of Information Theory*. Hoboken, N.J.: Wiley-Interscience.
- Cowell, Frank A. 1980a. "Generalized Entropy and the Measurement of Distributional Change." *European Economic Review* 13(1):147–59.
- Cowell, Frank A. 1980b. "On the Structure of Additive Inequality Measures." *Review of Economic Studies* 47(3):521–31.
- Cowell, Frank A. and K. Kuga. 1981. "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement." *Journal of Economic Theory* 25(1):131–43.
- Cowell, Frank A., Emmanuel Flachaire, and Sanghamitra Bandyopadhyay. 2013. "Reference distributions and inequality measurement." *Journal of Economic Inequality* 11(4):421–37.
- Dalton, Hugh. 1920. "The Measurement of the Inequality of Incomes." *The Economic Journal* 30(119):348–61.
- Duncan, Otis Dudley and Beverly Duncan. 1955. "A Methodological Analysis of Segregation

- Indexes." *American Sociological Review* 20(2):210–17.
- Falk, R. Frank, Charles F. Cortese, and Jack K. Cohen. 1978. "Utilizing standardized indices of residential segregation: comment on Winship." *Social Forces* 57:713.
- Farrell, Chad R. 2008. "Bifurcation, Fragmentation or Integration? The Racial and Geographical Structure of US Metropolitan Segregation, 1990–2000." *Urban Studies* 45(3):467–99.
- Fischer, Claude S., Gretchen Stockmayer, Jon Stiles, and Michael Hout. 2004. "Distinguishing the Geographic Levels and Social Dimensions of U.S. Metropolitan Segregation, 1960–2000." *Demography* 41(1):37–59.
- Fischer, Mary J. 2008. "Shifting Geographies: Examining the Role of Suburbanization in Blacks' Declining Segregation." *Urban Affairs Review* 43(4):475–96.
- Fossett, Mark and Scott J. South. 1983. "The Measurement of Intergroup Income Inequality: A Conceptual Review." *Social Forces* 61(3):855–71.
- Frey, William H. 2015. *Diversity Explosion: How New Racial Demographics Are Remaking America*. Brookings Institution Press.
- Hall, Matthew, Laura M. Tach, and Barrett A. Lee. 2016. "Trajectories of Ethnoracial Diversity in American Communities, 1980–2010." *Population and Development Review* 42(2):271–97.
- Handcock, Mark S. and Martina Morris. 1998. "Relative Distribution Methods." *Sociological Methodology* 28:53–97.
- Handcock, Mark Stephen and Martina Morris. 1999. *Relative Distribution Methods in the Social Sciences*. New York : Springer.
- Hao, Lingxin and Daniel Q. Naiman. 2010. *Assessing Inequality*. Los Angeles : SAGE.
- Iceland, John and R. Wilkes. 2006. "Does Socioeconomic Status Matter? Race, Class, and Residential Segregation." *Social Problems* 53(2):248–73.
- Jahn, Julius A., Calvin F. Schmid, and Clarence C. Schrag. 1947. "The Measurement of Ecological Segregation." *American Sociological Review* 12(3):293–303.
- James, David R. and Karl E. Taeuber. 1985. "Measures of Segregation." *Sociological Methodology* 15:1–32.
- Kullback, Solomon. 1987. "Letters to the Editor." *The American Statistician* 41:338–41.
- Lee, Barrett A., Sean F. Reardon, Glenn Firebaugh, Chad R. Farrell, Stephen A. Matthews, and David O'Sullivan. 2008. "Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales." *American Sociological Review* 73(5):766–91.
- Liao, Tim Futing. 2002. *Statistical Group Comparison*. New York: Wiley-Interscience.
- Lichter, D. T., D. Parisi, and M. C. Taquino. 2015. "Toward a New Macro-Segregation? Decomposing Segregation within and between Metropolitan Cities and Suburbs." *American Sociological Review* 80(4):843–73.
- Lichter, Daniel T., Domenico Parisi, and Michael C. Taquino. 2012. "The Geography of Exclusion: Race, Segregation, and Concentrated Poverty." *Social Problems* 59(3):364–88.
- Lieberson, Stanley and Donna K. Carter. 1982. "Temporal Changes and Urban Differences in Residential Segregation: A Reconsideration." *American Journal of Sociology* 88(2):296–310.
- Magdalou, Brice and Richard Nock. 2011. "Income Distributions and Decomposable Divergence Measures." *Journal of Economic Theory* 146(6):2440–54.



- Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67(2):281–315.
- Morgan, B. S. 1975. "The Segregation of Socioeconomic Groups in Urban Areas: A Comparative Analysis." *Urban Studies* 12:47–60.
- Morgan, Barrie S. and John Norbury. 1981. "Some Further Observations on the Index of Residential Differentiation." *Demography* 18(2):251–56.
- Mori, T., K. Nishikimi, and T. E. Smith. 2005. "A Divergence Statistic for Industrial Localization." *The Review of Economics and Statistics* 87(4):635–51.
- Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press.
- Page, Scott E. 2011. *Diversity and Complexity*. Princeton University Press.
- Parisi, Domenico, Daniel T. Lichter, and Michael C. Taquino. 2011. "Multi-Scale Residential Segregation: Black Exceptionalism and America's Changing Color Line." *Social Forces* 89(3):829–52.
- Pigou, A. C. 1912. *Wealth and Welfare*. London: Macmillan.
- R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reardon, Sean F. and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." *Sociological Methodology* 32:33–67.
- Reardon, Sean F. and David O'Sullivan. 2004. "Measures of Spatial Segregation." *Sociological Methodology* 34(1):121–62.
- Reardon, Sean F., Chad R. Farrell, Stephen A. Matthews, David O'Sullivan, Kendra Bischoff, and Glenn Firebaugh. 2009. "Race and Space in the 1990s: Changes in the Geographic Scale of Racial Residential Segregation, 1990–2000." *Social Science Research* 38(1):55–70.
- Reardon, Sean F., Stephen A. Matthews, David O'Sullivan, Barrett A. Lee, Glenn Firebaugh, Chad R. Farrell, and Kendra Bischoff. 2008. "The Geographic Scale of Metropolitan Racial Segregation." *Demography* 45(3):489–514.
- Roberto, Elizabeth. 2015. "The Boundaries of Spatial Inequality: Three Essays on the Measurement and Analysis of Residential Segregation." PhD thesis, Yale University.
- Sakoda, James M. 1981. "A Generalized Index of Dissimilarity." *Demography* 18(2):245–50.
- Sasson, Isaac. 2016. "Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990–2010." *Demography* 53(2):269–93.
- Schwartz, Joseph E. and Christopher Winship. 1980. "The Welfare Approach to Measuring Inequality." *Sociological Methodology* 11:1–36.
- Sen, Amartya. 1973. *On Economic Inequality*. Oxford: Clarendon Press.
- Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27(3):379–423.
- Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures." *Econometrica* 48(3):613–25.
- Shorrocks, Anthony F. 1984. "Inequality Decomposition by Population Subgroups." *Econometrica* 52(6):1369–85.

- Shorrocks, Anthony F. 2012. "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value." *The Journal of Economic Inequality* 11(1):99–126.
- Taeuber, Karl E. and Alma F. Taeuber. 1965. *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago Aldine Pub. Co.
- Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam: North Holland.
- Theil, Henri. 1972. *Statistical Decomposition Analysis*. edited by H. Theil. Amsterdam: North-Holland Publishing Company.
- Theil, Henri and Anthony J. Finizza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts." *The Journal of Mathematical Sociology* 1(2):187–93.
- U.S. Census Bureau. 2011. "2010 Census Summary File 1—United States."
- Waldman, Loren K. 1977. "Types and Measures of Inequality." *Social Science Quarterly* 58(2):229–41.
- Walsh, J. A. and M. E. O’Kelly. 1979. "An Information Theoretic Approach to Measurement of Spatial Inequality." *Economic and Social Review* 10:267–86.
- White, Michael J. 1986. "Segregation and Diversity Measures in Population Distribution." *Population Index* 52(2):198–221.
- Winship, Christopher. 1978. "The Desirability of Using the Index of Dissimilarity or Any Adjustment of It for Measuring Segregation: Reply to Falk, Cortese, and Cohen." *Social Forces* 57(2):717–20.
- Zoloth, B. S. 1976. "Alternative Measures of School Segregation." *Land Economics*.