The Isolation index

Casilda Lasso de la Vega   Oscar Volij

Discussion Paper No. 24-14

September 8, 2024

Monaster Center for
Economic Research
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva, Israel

Fax: 972-8-6472941
Tel: 972-8-6472286

The Isolation index*

Casilda Lasso de la Vega
University of the Basque Country

Oscar Volij
Ben Gurion University of the Negev.

September 8, 2024

Abstract

We provide an axiomatic characterization of the Isolation index, an index that is increasingly used by economists to measure segregation, exposure and related concepts.

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

Keywords: Segregation, evenness, dissimilarity, isolation.

*Email addresses: casilda.lassodelavega@ehu.es; ovolij@bgu.ac.il. We thank the Spanish Ministerio de Economía y Competitividad (project PID2019-107539GB-I00) and the Gobierno Vasco (project IT1697-22) for research support.

1 Introduction

The measurement of segregation has been a topic of research in sociology and in economics for quite a long time. One of its difficulties is that the very meaning of the concept is not clear. Massey and Denton (1988) identified five dimensions of segregation, each of which refers to some different aspect of it. In retrospect, the dimension that captured most of the research attention is that of *evenness*, which involves a comparison of the distribution of groups across locations. According to this criterion, the more similarly different groups are distributed across locations, the lower the existing segregation. Still, even if we are interested in the evenness dimension, it is not clear how to measure the similarity of the groups' distributions. Evidence of this difficulty is the large number of indices that have been proposed and used in empirical studies (see Massey and Denton (1988) and Flückiger and Silber (2012)).

The first contribution toward some agreement on the way segregation should be measured was James and Taeuber (1985), who proposed a number of requirements that a segregation measure should satisfy, and evaluated existing indices based on them. The first paper to study segregation axiomatically is Philipson (1993). It provided a characterization of a large family of segregation orderings that have an additively separable representation. Later, Hutchens (2001) characterized a family of indices that satisfy a set of basic properties and Hutchens (2004) strengthened one axiom to obtain a unique segregation index, the Square Root index. Echenique and Fryer Jr (2007) characterized a segregation index based on a graph-theoretic model. Except for Philipson (1993), all the above characterizations use at least one cardinal axiom. More recently, Frankel and Volij (2011) provided a characterization of two multigroup segregation indices, the Atkinson index and the Mutual Information index, and Puerta and Urrutia (2016) characterized the Gini segregation index. Both papers use exclusively ordinal axioms.

The fact that some indices had not been justified on first principles did not prevent researchers from using them to measure various kinds of segregation, prominently residential and school segregation between ethnic groups. In the last years, these indices have also been ingeniously applied to the measurement of ideological segregation (Gentzkow and Shapiro (2011)), and partisanship in congressional speeches (Gentzkow, Shapiro, and Taddy (2019)).

Interestingly, despite the solid theoretical underpinnings of the Atkinson, Mutual Information and Gini indices, they are rarely used in empirical studies.[1] Until recently, the most widely used measure was the Dissimilarity index. But more recent work in economics has adopted the Isolation index or a variation of it as the main measure. See, for instance, Cutler, Glaeser, and Vigdor (1999), Gentzkow and Shapiro (2011), Gentzkow, Shapiro, and Taddy (2019), Athey, Ferguson, Gentzkow, and Schmidt (2021), and Monarrez (2023). This index satisfies some of the desirable properties that are expected from segregation measures. However, to the best of our knowledge, it has not been fully characterized yet. This paper aims at filling this gap. For this purpose, we add one axiom to the standard requirements of symmetry, scale invariance, transfer, and independence. The additional axiom states a simple condition for two two-neighborhood cities, composed of one black ghetto and one mixed neighborhood to be equally segregated. This condition requires that the proportion of blacks in the mixed neighborhood relative to their proportion in the population be the same for both cities. It turns out that the ratio of the proportion of blacks in the mixed neighborhood to that in the city is the seed that, once sowed in the good ground of the standard axioms, yields the Isolation index. Indeed, this paper shows that there is a unique segregation order that satisfies all of the above axioms, and it is the one represented by the Isolation index. Furthermore, we show that these axioms are independent.

[1] For notable exceptions, see Oosterbeek, Sóvágó, and van der Klaauw (2021) and Serrati (2024).

2 Notation

Throughout the paper we use the language of urban ethnic segregation because it is one of the most widely studied. Our results apply in other contexts as well.

A *city* is a pair $\langle N, (B_n, W_n)_{n\in N}\rangle$, where N is a finite set of neighborhoods and for each $n \in N$, (B_n, W_n) is n's ethnic composition. The first and second components of (B_n, W_n) are the numbers of blacks and whites, respectively, in n, at least one of which is assumed to be positive.

Given a city $X = \langle N, (B_n, W_n)_{n\in N}\rangle$, we denote by B_X and W_X the total numbers of blacks and whites in X. When it is clear to which city we are referring, we will write simply B and W. We restrict attention to cities in which $B > 0$ and $W > 0$. Also, the following notation will be useful.

$$\begin{aligned}
P &= \frac{B}{B+W}\text{: the proportion of blacks in the city,}\\
p_n &= \frac{B_n}{B_n+W_n}\text{: the proportion of blacks in neighborhood } n,\\
\pi_n &= \frac{B_n+W_n}{B+W}\text{: the proportion of the population that live in neighborhood } n,\\
b_n &= \frac{B_n}{B}\text{: the proportion of the city's blacks that live in neighborhood } n,\\
w_n &= \frac{W_n}{W}\text{: the proportion of the city's whites that live in neighborhood } n.
\end{aligned}$$

For any city $X = \langle N, (B_n, W_n)_{n\in N}\rangle$ and any positive constant α, αX denotes the city that results from multiplying the number of blacks and whites in each neighborhood of X by α: $\alpha X = \langle N, (\alpha B_n, \alpha W_n)_{n\in N}\rangle$. For any two cities $X = \langle N_X, (B_n, W_n)_{n\in N_X}\rangle$ and $Y = \langle N_Y, (B_n, W_n)_{n\in N_Y}\rangle$, with disjoint sets of neighborhoods, let $X \cup Y$ denote the concatenation of the two: $X \cup Y = \langle N_X \cup N_Y, (B_n, W_n)_{n\in N_X\cup N_Y}\rangle$. We say that two cities, $X = \langle N, (B_n, W_n)_{n\in N}\rangle$ and $Y = \langle M, (B_m, W_m)_{m\in M}\rangle$, are *equivalent* if there is a one to one mapping $\phi : N \to M$ such that for all $n \in N$, $(B_n, W_n) = (B_{\phi(n)}, W_{\phi(n)})$. We will consider any two equivalent cities as one and the same city. For this reason, we will denote a city $X = \langle N, (B_n, W_n)_{n\in N}\rangle$, simply by, $\langle (B_n, W_n)_{n\in N}\rangle$. Neighborhood n is said to be a *ghetto* if contains either 0 blacks or 0 whites.

3 Segregation orderings and measures

We denote the set of all cities by $\mathcal{C}$. A segregation order, $\succcurlyeq$, is a complete and transitive binary relation on $\mathcal{C}$. We interpret $X \succcurlyeq Y$ to mean "city X is at least as segregated as city Y." The relations $\sim$ and $\succ$ are derived from $\succcurlyeq$ in the usual way.[2]

Segregation orders are usually represented by segregation indices. A segregation index assigns to each city a nonnegative number which is meant to capture its level of segregation. Given a segregation index S, the associated segregation order is defined by $X \succeq Y \Leftrightarrow S(X) \geq S(Y)$. Clearly, a segregation order may be represented by more than one index.

3.1 Examples of Segregation Indices

The following indices have been used to study segregation (Massey and Denton (1988)).

The Index of Dissimilarity This index measures the proportion of either racial group that would need to be reallocated across neighborhoods in order to obtain perfect integration. It is given by

$$D(X) = \frac{1}{2} \sum_{n \in N} |b_n - w_n| \,. \tag{1}$$

This index was introduced to the literature by Jahn, Schmid, and Schrag (1947).

Square Root The Square Root index is defined as:

$$A(X) = 1 - \sum_{n \in N} \sqrt{b_n w_n} \,. \tag{2}$$

A characterization of the Square Root index can be found in Hutchens (2004).

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

Mutual Information The Mutual Information index is defined as

$$MI(X) = h(P, 1-P) - \sum_{n\in N} \pi_n h\left(p_n, 1-p_n\right) \tag{3}$$

where $h(q, 1-q) = q\log_2\left(\frac{1}{q}\right) + (1-q)\log_2\left(\frac{1}{1-q}\right)$ is the entropy function.This index was first proposed by Theil et al. (1971) and has been applied, among others, by Fuchs (1975) and Mora and Ruiz-Castillo (2003, 2004). This index, is characterized by Frankel and Volij (2011).

Entropy The Entropy index is obtained by normalizing the Mutual Information index so that it takes on values between 0 and 1:

$$H(X) = \frac{MI(X)}{h(P, 1-P)}.$$

This index was proposed by Theil (1972) and Theil and Finizza (1992).

Isolation The Isolation index is defined by

$$I(X) = \sum_{n\in N} b_n \frac{p_n - P}{1-P}. \tag{4}$$

It was proposed by Bell (1954).

3.2 The many formulations of the Isolation index

The Isolation index appears in the literature under various equivalent formulations. Bell (1954) used the formulation that appears in (4). Gentzkow and Shapiro (2011) formulates the Isolation index as

$$I(X) = \sum_{n\in N} \frac{B_n}{B}\frac{B_n}{B_n+W_n} - \sum_{n\in N} \frac{W_n}{W}\frac{B_n}{B_n+W_n}.$$

Reardon (2011) shows that the Isolation index can also be written as

$$I(X) = 1 - \frac{\sum_{n\in N} \pi_n p_n\left(1-p_n\right)}{P\left(1-P\right)}. \tag{5}$$

Letting x be the random variable that takes on the value 1 if an individual is black and 0 if he is white, we see that $var(x) = P(1-P)$ and $var(x|n) = p_n(1-p_n)$. Recall that the variance decomposition states that $var(x) = E\left[var[x|n]\right] + var(E\left[x|n\right])$. Therefore the Isolation index can be written as

$$I(X) = \frac{var(E\left[x|n\right])}{var(x)} \tag{6}$$

and is also known as the variance ratio. It is clear from this formulation that $I(X) \in [0,1]$.

Finally, by some algebraic manipulation, it can be checked that the isolation index can be equivalently written as

$$I(X) = 1 - \sum_{n \in N} \frac{b_n w_n}{\pi_n}. \tag{7}$$

For a proof, see the appendix.

4 Axioms and main result

In this section we propose a number of properties that a segregation order may satisfy.[3] Most of them are widely accepted in the literature.

Symmetry (SYM) A segregation ordering satisfies Symmetry if the segregation in a city is unaffected by relabeling the races: $\langle (B_n, W_n)_{n\in N}\rangle \sim \langle (W_n, B_n)_{n\in N}\rangle$.

It is clear from formulation (7) that the Isolation index satisfies Symmetry. (It is also clear from (5) and (6)).

Scale Invariance (SI) A segregation ordering satisfies Scale Invariance if a city's segregation is unchanged when the number of agents is multiplied by a positive number in all neighborhoods: for any $\alpha > 0$, $\langle (B_n, W_n)_{n\in N}\rangle \sim \langle (\alpha B_n, \alpha W_n)_{n\in N}\rangle$.

Since the ethnic distribution of a neighborhood is unaffected by a multiplication of its residents by a positive constant, the Isolation index satisfies SI.

[3] With some abuse of language, we will say that a segregation index satisfies a property if its induced segregation order does.

The next axiom states that dividing a neighborhood into two, cannot decrease the segregation of tha city. Moreover, it increases its segregation if and only if the two neighborhoods have the same ethnic distribution. This axiom is equivalent to the principles of Organization Equivalence and of Transfers advocated by James and Taeuber (1985).

Neighborhood division property (NDP) Let $X \in \mathcal{C}$ be a city and let n be one of its neighborhoods. Let Y be the city that results from dividing n into two neighborhoods, n_1 and n_2. If n_1 and n_2 have the same group distributions then X and Y are equally segregated. Otherwise, Y is more segregated than X. Formally, if $p_{n_1} = p_{n_2}$, then $Y \sim X$. Otherwise, $Y \succ X$.

Since the function $f(x) = x(1-x)$ is strictly concave, using the formulation (5) it is easy to see that the Isolation index satisfies NDP. Indeed, let X be a city and let Y be the city that is obtained from X by parititioning neighborhood n into two neigborhoods, n_1 and n_2. Then, denoting by P their common proportion of blacks,

$$\begin{aligned} I(Y) - I(X) &= \frac{\pi_n f(p_n) - \pi_{n_1} f(p_{n_1}) - \pi_{n_2} f(p_{n_2})}{f(P)} \\ &= \frac{\pi_n \left(f(p_n) - \frac{\pi_{n_1}}{\pi_n} f(p_{n_1}) - \frac{\pi_{n_2}}{\pi_n} f(p_{n_2}) \right)}{f(P)} \\ &\geq 0. \end{aligned}$$

where the inequality follows from the concavity of f and from the fact that $p_n = \frac{\pi_{n_1}}{\pi_n} p_{n_1} + \frac{\pi_{n_2}}{\pi_n} p_{n_2}$. If $p_{n_1} \neq p_{n_2}$, the inequality is strict, by strict concavity. If $p_{n_1} = p_{n_2}$, then $p_{n_1} = p_{n_2} = p_n$ and $I(Y) - I(X) = 0$.

The next axiom states that under limited conditions, adjoining the same set of neighborhoods to each of two different cities does not affect which of the two cities is more segregated. Alternatively, it states that if a city is composed of two subcities, and the residents of one of them are redistributed within the subcity itself in a way that its segregation decreases, then the segregation of the whole city decreases as well.

Independence (IND) Let X, and Y be two cities with the same number of blacks and the same number of whites. Then, for all cities Z, $X \succcurlyeq Y$ if and only if $X \cup Z \succcurlyeq Y \cup Z$.

The Isolation index satisfies IND. To see this, let X and Y be two cities with the same population and the same ethnic distribution. That is, $B_X = B_Y$ and $W_X = W_Y$ (and therefore $P(X) = P(Y)$). Let Z be another city. Letting $f(x) = x(1-x)$ and using the formulation (5)

$$\begin{aligned}
I(X \cup Z) \quad &\geq \quad I(Y \cup Z) \Leftrightarrow 1 - \frac{\sum_{n \in N(X \cup Z)} \pi_n f(p_n)}{f(P(X \cup Z))} \geq 1 - \frac{\sum_{n \in N(Y \cup Z)} \pi_n f(p_n)}{f(P(Y \cup Z))} \\
&\Leftrightarrow \quad \sum_{n \in N(X)} \frac{B_n + W_n}{B_{X \cup Z} + W_{X \cup Z}} f(p_n) \leq \sum_{n \in N(Y)} \frac{B_n + W_n}{B_{Y \cup Z} + W_{Y \cup Z}} f(p_n) \\
&\Leftrightarrow \quad \sum_{n \in N(X)} \frac{B_n + W_n}{B_X + W_X} f(p_n) \leq \sum_{n \in N(Y)} \frac{B_n + W_n}{B_Y + W_Y} f(p_n) \\
&\Leftrightarrow \quad f(P(X)) - \sum_{n \in N(X)} \frac{B_n + W_n}{B_X + W_X} f(p_n) \geq f(P(Y)) - \sum_{n \in N(Y)} \frac{B_n + W_n}{B_Y + W_Y} f(p_n) \\
&\Leftrightarrow \quad \frac{f(P(X)) - \sum_{n \in N(X)} \frac{B_n + W_n}{B_X + W_X} f(p_n)}{f(P(X))} \geq \frac{f(P(Y)) - \sum_{n \in N(Y)} \frac{B_n + W_n}{B_Y + W_Y} f(p_n)}{f(P(Y))} \\
&\Leftrightarrow \quad I(X) \geq I(Y).
\end{aligned}$$

The last axiom states conditions under which two two-neighborhood cities, each with one black ghetto and one mixed neighborhood, are equally segregated. Consider one such city. If the city were fully integrated, the exposure of whites to blacks would be P. However, since whites are exposed to blacks only in the mixed neighborhood, their actual exposure to blacks is the proportion of blacks in that particular neighborhood, denoted p_m. The axiom posits that two two-neighborhood cities, each with one black ghetto and one mixed neighborhood, are equally segregated if the ratio of actual to potential exposure to blacks is the same. Formally,

Relative Exposure (RE) Let $X = \langle (B_m, W_m), (B_g, 0) \rangle$ be a city with a total population of 1, composed of one mixed neighborhood and one black ghetto, and denote by P the proportion of blacks in the city and by p_m the proportion of blacks in the mixed neighborhood. Similarly, let $X' = \left\langle (B'_m, W'_m), \left(B'_g, 0\right)\right\rangle$ another such city, and denote by P' the proportion of blacks in it and by p'_m the proportion of blacks in its mixed neighborhood. The segregation ordering satisfies Realtive Exposure if

$$\frac{p_m}{P} = \frac{p'_m}{P'} \Rightarrow X \sim X'.$$

The Isolation index satisfies Relative Exposure. Indeed,

$$\begin{aligned}
\frac{p_m}{P} &= \frac{p'_m}{P'} \Rightarrow \frac{\frac{B_m}{B_m+W_m}}{B_m+B_g} = \frac{\frac{B'_m}{B'_m+W'_m}}{B'_m+B'_g} \\
&\Rightarrow \frac{\frac{B_m}{B_m+B_g}}{B_m+W_m} = \frac{\frac{B'_m}{B'_m+B'_g}}{B'_m+W'_m} \\
&\Rightarrow \frac{b_m}{\pi_m} = \frac{b'_m}{\pi'_m} \Rightarrow I(X) = I(X')
\end{aligned}$$

where the last implication uses formulation (7) and the fact that $w_m = w'_m = 1$.

We are now ready to state our main result.

Theorem 1 *The Isolation ordering is the only ordering that satisfies Symmetry, Scale Invariance, Neighborhood Division Property, Independence, and Relative Exposure.*

Proof. We have already seen that the Isolation index satisfies the foregoing axioms. In order to see that it is the only one so do so, let $\succcurlyeq$ be a segregation ordering that satisfies SYM, SI, NDP, IND, and RE. We will show that it is the Isolation ordering. We first define for each $P \in (0,1)$, two special cities with the same ethnic distribution. The first one is $\overline{X}(P) = \langle (P,0), (0,1-P)\rangle$ and the second one is $\underline{X}(P) = \langle (P, 1-P)\rangle$. The former is a completely segregated city that consists of two ghettos, and the latter is a completely integrated, one-neighborhood city.

The following two lemmas are direct consequences of the axioms. Similar versions appear in Frankel and Volij (2011). The first one follows from SI and NDP, and its proof is left to the reader.

Lemma 1 *Let X be a city with ethnic distribution $(P, 1-P)$. Then*

1. $\overline{X}(P) \succsim X \succsim \underline{X}(P)$.

2. *If all the neighborhoods of X have the same ethnic distribution, then $X \sim \underline{X}(P)$.*

3. *If every neighorhood of X is a ghetto, then $\overline{X}(P) \sim X$.*

Note that for all $\alpha \in (0,1)$, $\alpha\underline{X}(P) \cup (1-\alpha)\overline{X}(P) = \{(\alpha,\alpha),(1-\alpha,0),(0,1-\alpha)\}$. With some abuse of notation we shall identify $\alpha\underline{X}(P) \cup (1-\alpha)\overline{X}(P)$ with $\underline{X}$, if $\alpha = 1$, and with $\overline{X}$, if $\alpha = 0$. More generally, we shall identify a collection $\langle (B_n, W_n)_{n \in N} \rangle$ with the city that is obtained from it by ignoring the pairs $(B_n, W_n) = (0,0)$.

Lemma 2 *For any $\alpha, \beta \in [0,1]$, $\alpha > \beta$, and for any $P \in (0,1)$,*

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

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

$$\beta\underline{X}(P) \cup (\alpha-\beta)\overline{X}(P) \cup (1-\alpha)\,\overline{X}(P) \succ \beta\underline{X}(P) \cup (\alpha-\beta)\underline{X}(P) \cup (1-\alpha)\,\overline{X}(P).$$

But since by NDP,

$$\beta\underline{X}(P) \cup (\alpha-\beta)\overline{X}(P) \cup (1-\alpha)\,\overline{X}(P) \sim \beta\underline{X}(P) \cup (1-\beta)\,\overline{X}(P)$$

and

$$\beta\underline{X}(P) \cup (\alpha-\beta)\underline{X}(P) \cup (1-\alpha)\,\overline{X}(P) \sim \alpha\underline{X}(P) \cup (1-\alpha)\,\overline{X}(P)$$

the result follows by transitivity. ■

Proposition 1 *Let $X = \langle (Pb, (1-P)w), (P(1-b), 0), (0, (1-P)(1-w)) \rangle$ be a city with one mixed neighborhood and two ghettos where $(b,w) \in [0,1]^2$ and $P \in (0,1)$. There exists a unique $\alpha_X \in [0,1]$ such that $X \sim \alpha_X\underline{X}(P) \cup (1-\alpha_X)\overline{X}(P)$. Further, this unique α_X is $\frac{bw}{Pb+(1-P)w}$.*

Proof. Uniqueness follows from Lemma 2. For existence, there are two cases.

Case 1: Suppose $b = 0$ or $w = 0$. In this case X is composed exclusively of ghettos. Therefore, by Lemma 1 $X \sim \overline{X}(P)$ and we have $\alpha_X = 0 = \frac{bw}{Pb+(1-P)w}$.

Case 2. Suppose $bw \neq 0$. By SYM, we can assume w.l.o.g. that $b \geq w$. Let $\alpha = \frac{bw}{\pi_m}$, where $\pi_m = Pb + (1-P)w$ is the proportion of the population that resides in the mixed neighborhood, and let

$$\begin{aligned} Y &= \alpha \underline{X}(P) \cup (1-\alpha)\overline{X}(P) \\ &= \langle (\alpha P, \alpha(1-P)), ((1-\alpha)P, 0), (0, (1-\alpha)(1-P)) \rangle \\ &= \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), \left(\left(1-\frac{bw}{\pi_m}\right)P, 0\right), \left(0, \left(1-\frac{bw}{\pi_m}\right)(1-P)\right) \right\rangle. \end{aligned}$$

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

The number of whites in the ghetto of X and in the ghetto of Y are, respectively, $(1-P)(1-w)$ and $\left(1-\frac{bw}{\pi_m}\right)(1-P)$. Since $b \geq w$, their difference is non-negative and given by

$$\Delta_w = (1-P)(1-w) - \left(1-\frac{bw}{\pi_m}\right)(1-P) = \frac{w(1-P)^2(b-w)}{\pi_m} \geq 0.$$

Similarly, the number of blacks in the ghetto of Y and in the ghetto of X are, respectively, $\left(1-\frac{bw}{\pi_m}\right)P$ and $P(1-b)$. Since $b \geq w$, their difference is non-negative and given by

$$\Delta_b = \left(1-\frac{bw}{\pi_m}\right)P - P(1-b) = \frac{bP^2(b-w)}{\pi_m} \geq 0.$$

Then, we can divide the white ghetto of X and the black ghetto of Y as follows:

$$\begin{aligned} (0, (1-P)(1-w)) &\to (0, \Delta_w) \cup \left(0, \left(1-\frac{bw}{\pi_m}\right)(1-P)\right), \\ \left(\left(1-\frac{bw}{\pi_m}\right)P, 0\right) &\to (\Delta_b, 0) \cup (P(1-b), 0). \end{aligned}$$

Therefore, by NDP,

$$\begin{aligned} X &= \langle (Pb, (1-P)w), (P(1-b), 0), (0, (1-P)(1-w)) \rangle \\ &\sim \left\langle (Pb, (1-P)w), (P(1-b), 0), (0, \Delta_w), \left(0, \left(1-\frac{bw}{\pi_m}\right)(1-P)\right) \right\rangle \end{aligned}$$

and

$$
\begin{aligned}
Y &= \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), \left(\left(1-\frac{bw}{\pi_m}\right)P, 0\right), \left(0, \left(1-\frac{bw}{\pi_m}\right)(1-P)\right) \right\rangle \\
&\sim \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), (\Delta_b, 0), (P(1-b), 0), \left(0, \left(1-\frac{bw}{\pi_m}\right)(1-P)\right) \right\rangle.
\end{aligned}
$$

Therefore, by IND and SYM,

$$
\begin{aligned}
X \sim Y &\Leftrightarrow \langle (Pb, (1-P)w), (0, \Delta_w) \rangle \sim \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), (\Delta_b, 0) \right\rangle \\
&\Leftrightarrow \langle ((1-P)w, Pb), (\Delta_w, 0) \rangle \sim \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), (\Delta_b, 0) \right\rangle \quad (8)
\end{aligned}
$$

In order to show that (8) holds, we will make use of RE. Denote

$$
\begin{aligned}
Z &= \langle ((1-P)w, Pb), (\Delta_w, 0) \rangle \\
Z' &= \left\langle \left(\frac{bw}{\pi_m}P, \frac{bw}{\pi_m}(1-P)\right), (\Delta_b, 0) \right\rangle.
\end{aligned}
$$

The black proportion in the mixed neighborhood of Z and of Z' are, respectively,

$$
\begin{aligned}
p_m(Z) &= \frac{(1-P)w}{(1-P)w + Pb} \\
p_m(Z') &= P.
\end{aligned}
$$

The overall black proportion of Z and of Z' are, respectively,

$$
\begin{aligned}
P(Z) &= \frac{(1-P)w + \Delta_w}{((1-P)w + Pb) + \Delta_w} = \frac{w(1-P)}{bP^2 + w(1-P^2)} \\
P(Z') &= \frac{\frac{bw}{\pi_m}P + \Delta_b}{\frac{bw}{\pi_m} + \Delta_b} = \frac{P\pi_m}{bP^2 + w(1-P^2)}.
\end{aligned}
$$

Therefore,

$$
\frac{p_m(Z)}{P(Z)} = \frac{\frac{w(1-P)}{Pb+(1-P)w}}{\frac{w(1-P)}{bP^2+w(1-P^2)}} = \frac{P^2b + (1-P^2)w}{Pb + (1-P)w}
$$

and

$$
\frac{p_m(Z')}{P(Z')} = \frac{P}{\frac{P(bP+w(1-P))}{bP^2+w(1-P^2)}} = \frac{P^2b + (1-P^2)w}{Pb + (1-P)w}.
$$

Therefore, by SI and RE, (8) is true, and consequently $X \sim Y$. ■

The next proposition generalizes Proposition 1 to cities with more that one mixed neighborhood.

Proposition 2 *Let* $X = \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i) \rangle \cup \langle (P(1-\sum_{i=1}^n b_i), 0), (0, (1-P)(1-\sum_{i=1}^n w_i)) \rangle$ *be a city with* n *mixed neighborhoods and two ghettos, one for each group, where* $P \in (0,1)$*, and* $b_i, w_i \in [0,1]$ *for* $i = 1, \ldots, n$*, with* $\sum_{i=1}^n b_i \leq 1$ *and* $\sum_{i=1}^n w_i \leq 1$*. There is a unique* $\alpha_X \in [0,1]$ *such that* $X \sim \alpha_X \underline{X}(P) \cup (1-\alpha_X)\overline{X}(P)$*. Further, this unique* α_X *is* $\sum_{i=1}^n \frac{b_i w_i}{Pb_i + (1-P)w_i}$*.*

Proof. Let $X = \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i) \rangle \cup \langle (P(1-\sum_{i=1}^n b_i), 0), (0, (1-P)(1-\sum_{i=1}^n w_i)) \rangle$ be a city with n mixed neighborhoods and two ghettos. If we adjoin $(n-1)\overline{X}(P) = \langle (P(n-1), 0), (0, (1-P)(n-1)) \rangle$ to it, we obtain

$$\begin{aligned}
& X \cup (n-1)\overline{X}(P) \\
= & \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i) \rangle \cup \\
& \qquad \Big\langle \Big(P\big(1 - \sum_{i=1}^n b_i\big), 0\Big), \Big(0, (1-P)\big(1 - \sum_{i=1}^n w_i\big)\Big) \Big\rangle \cup \langle (P(n-1), 0), (0, (1-P)(n-1)) \rangle \\
\sim & \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i) \rangle \cup \Big\langle \Big(P\big(n - \sum_{i=1}^n b_i\big), 0\Big), \Big(0, (1-P)\big(n - \sum_{i=1}^n w_i\big)\Big) \Big\rangle \qquad \text{by NDP} \\
= & \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i) \rangle \cup \Big\langle \Big(P \sum_{i=1}^n (1-b_i), 0\Big), \Big(0, (1-P) \sum_{i=1}^n (1-w_i)\Big) \Big\rangle \\
\sim & \bigcup_{i=1}^{n} \langle (Pb_i, (1-P)w_i), (P(1-b_i), 0), (0, (1-P)(1-w_i)) \rangle \qquad \text{by NDP.}
\end{aligned}$$

By Proposition 1, for each $i = 1, \ldots, n$, there exists $\alpha_i = \frac{b_i w_i}{Pb_i + (1-P)w_i} \in [0,1]$ such that

$$\langle (Pb_i, (1-P)w_i), (P(1-b_i), 0), (0, (1-P)(1-w_i)) \rangle \sim \alpha_i \underline{X}(P) \cup (1-\alpha_i)\overline{X}(P).$$

Therefore, by IND applied n times

$$\begin{aligned}
X \cup (n-1)\overline{X}(P) & \sim \bigcup_{i=1}^{n} \big\langle \alpha_i \underline{X}(P) \cup (1-\alpha_i)\overline{X}(P) \big\rangle \\
& \sim \Big(\sum_{i=1}^n \alpha_i\Big)\underline{X}(P) \cup \Big(\sum_{i=1}^n (1-\alpha_i)\Big)\overline{X}(P) \qquad \text{by NDP} \\
& \sim \Big(\sum_{i=1}^n \alpha_i\Big)\underline{X}(P) \cup \Big(n - 1 + 1 - \sum_{i=1}^n \alpha_i\Big)\overline{X}(P) \\
& \sim \Big(\sum_{i=1}^n \alpha_i\Big)\underline{X}(P) \cup \big(n-1\big)\overline{X}(P) \cup \Big(1 - \sum_{i=1}^n \alpha_i\Big)\overline{X}(P) \qquad \text{by NDP.}
\end{aligned}$$

Finally, by IND

$$X \sim \Big(\sum_{i=1}^{n} \alpha_i\Big)\underline{X}(P) \cup \Big(1 - \sum_{i=1}^{n} \alpha_i\Big)\overline{X}(P) = \alpha_X \underline{X}(P) \cup (1 - \alpha_X)\overline{X}(P)$$

where $\alpha_X = \sum_{i=1}^{n} \frac{b_i w_i}{P b_i + (1-P) w_i}$. ■

The previous propositions allowed us to compare cities with the same ethnic distribution. The next one will allow us to compare cities with different ethnic distributions.

Proposition 3 *For all* $\alpha \in [0,1]$ *and for all* $P, P' \in (0,1)$*, we have that*

$$\alpha \underline{X}(P) \cup (1-\alpha)\overline{X}(P) \sim \alpha \underline{X}(P') \cup (1-\alpha)\overline{X}(P').$$

Proof. Let $\alpha \in [0,1]$ and let $P, P' \in (0,1)$. Find $b \in [0,1]$ such that

$$X \sim \alpha \underline{X}(P) \cup (1-\alpha)\overline{X}(P) \tag{9}$$

where $X = \langle (P(1-b), 0), (Pb, 1-P) \rangle$. By Proposition 1,

$$b = \frac{(1-P)\,\alpha}{1 - P\alpha}.$$

Similarly, find $b' \in [0,1]$ such that

$$X' \sim \alpha \underline{X}(P') \cup (1-\alpha)\overline{X}(P') \tag{10}$$

where $X' = \langle (P'(1-b'), 0), (P'b', (1-P')) \rangle$. By Proposition 1,

$$b' = \frac{(1-P')\,\alpha}{1 - P'\alpha}.$$

It can be checked that

$$\frac{\frac{Pb}{Pb+(1-P)}}{P} = \alpha = \frac{\frac{P'b'}{P'b'+(1-P')}}{P'}.$$

Therefore, by RE, $X \sim X'$, which implies, by (9–10), $\alpha \underline{X}(P) \cup (1-\alpha)\overline{X}(P) \sim \alpha \underline{X}(P') \cup (1-\alpha)\overline{X}(P')$. ■

The next corollary will allow us to define an index that represents our segregation ordering.

Corollary 1 *Let $X \in \mathcal{C}$. There is a unique α_X such that $X \sim \alpha_X \underline{X}(1/2) \cup (1 - \alpha_X)\overline{X}(1/2)$.*

Proof. Let $X \in \mathcal{C}$. Let $(P, (1-P))$ be its ethnic distribution. By SI, we can assume w.l.o.g. that it has a population of 1. Further, by NDP, we can assume that it has the following form

$$X = \Big\langle \big(Pb_1, (1{-}P)w_1\big), \cdots, \big(Pb_n, (1{-}P)w_n\big), \Big(P\big(1{-}\sum_{i=1}^{n} b_i\big), 0\Big), \Big(0, (1{-}P)\big(1{-}\sum_{i=1}^{n} w_i\big)\Big)\Big\rangle.$$

By Proposition 2 there is a unique α_X such that $X \sim \alpha_X \underline{X}(P) \cup (1 - \alpha_X)\overline{X}(P)$. By Proposition 3, $X \sim \alpha_X \underline{X}(1/2) \cup (1 - \alpha_X)\overline{X}(1/2)$. ■

We now continue with the proof of the theorem. Let $X, Y \in \mathcal{C}$. By Corollary 1, there are numbers α_X and α_Y such that $X \sim \alpha_X \underline{X}(1/2) \cup (1 - \alpha_X)\overline{X}(1/2)$ and $Y \sim \alpha_Y \underline{X}(1/2) \cup (1 - \alpha_Y)\overline{X}(1/2)$. By Lemma 2

$$X \succcurlyeq Y \Leftrightarrow 1 - \alpha_X \geq 1 - \alpha_Y$$

which implies that the index $S : \mathcal{C} \to \mathbb{R}$ defined by $S(X) = 1 - \alpha_X$, where α_X is the number identified in Proposition 1 represents the segregation ordering $\succcurlyeq$. By Proposition 2

$$1 - \alpha_X = 1 - \sum_{i=1}^{n} \frac{b_i w_i}{Pb_i + (1-P)w_i}.$$

Since $Pb_i + (1 - P)w_i = \pi_i$, we obtain that $1 - \alpha_X = I(X)$. ■

4.1 Independence of the axioms

In this section we show that the axioms used in Theorem 1 are independent.

Symmetry Consider the index

$$\begin{aligned} S_1(X) &= 1 - \sum_{n \in N} \frac{\sqrt{b_n} w_n}{\sqrt{\pi_n}} \\ &= 1 - \frac{\sum_{n \in N} \pi_n \sqrt{p_n}\,(1 - p_n)}{\sqrt{P}\,(1 - P)}. \end{aligned}$$

This index satisfies all the axioms except for Symmetry. It satisfies SI since the black and white proportions do not change when the population is uniformly multiplied by a constant. The fact that it satisfies NDP follows from the strict concavity of the function $f(p) = \sqrt{p}(1-p)$. The proof that it satisfies IND follows from the additivity of the index and is analogous to the proof that the Isolation index satisfies the axiom. Finally, to see that it satisfies RE, let $X = \langle (B_1, W), (B_2, 0) \rangle$ and $X' = \langle (B_1', W'), (B_2', 0) \rangle$ be two cities with unit population, composed of one mixed neighborhood and one black ghetto. Denote by P the proportion of blacks in X and by π_m and p_m respectively, the proportion of the population that resides in the mixed neighborhood, and the proportion of all blacks there. Similarly, Let P', π_m', and p_m' the same proportions for X', and assume that $p_m/P = p_m'/P'$. Note that for these cities, we have that

$$\pi_m \frac{1-p_m}{1-P} = (B_1 + W) \frac{W}{B_1 + W} \frac{1}{W} = 1 \tag{11}$$

$$\pi_m' \frac{1-p_m'}{1-P'} = (B_1' + W') \frac{W'}{B_1' + W'} \frac{1}{W'} = 1 \tag{12}$$

and that the proportion of whites in the ghettos is 0. Therefore,

$$\begin{aligned}
\frac{p_m}{P} &= \frac{p_m'}{P'} \Rightarrow 1 - \frac{\sqrt{p_m}}{\sqrt{P}} = 1 - \frac{\sqrt{p_m'}}{\sqrt{P'}} \\
&\Leftrightarrow 1 - \pi_m \frac{\sqrt{p_m}}{\sqrt{P}} \frac{1-p_m}{1-P} = 1 - \pi_m' \frac{\sqrt{p_m'}}{\sqrt{P'}} \frac{1-p_m'}{1-P'} \qquad \text{by (11-12)} \\
&\Leftrightarrow S_1(X) = S_1(X').
\end{aligned}$$

Scale Invariance Consider the index defined by $S_2(X) = T_X * I(X)$, where T_X stands for the total population of X. Since I satisfies SYM, GDP, IND and RE, so does S_2. It is clear that it violates SI.

Neighborhood Division Property Consider the index defined by $S_3(X) = 1 - I(X)$. Since I satisfies SI, SYM, IND and RE, so does S_3. It violates NDP.

Independence For any cith X, let $a(X)$ denote the city that is obtained from X by aggregating all the neighborhoods with the same ethnic distribution into one neighborhood. Consider the index defined by $S_4(X) = I(X) +$ number of ghettos $(a(X))$. It satisfies SYM, SI, NDP and RE since so does I. To see that is does not satisfy IND, consider the following cities:

$$\begin{aligned} X &= \langle (1,2), (1,0) \rangle, \\ Y &= \langle (2,1), (0,1) \rangle, \\ Z &= \langle (0,1) \rangle. \end{aligned}$$

We see that X and Y have the same population and ethnic distribution and that both have a single ghetto. By SYM and IND of I, $I(X) = I(Y)$, and $I(X \cup Z) = I(Y \cup Z)$. Therefore

$$S_4(X) = I(X) + 1 = I(Y) + 1 = S_4(Y).$$

However, since $a(X \cup Z) = \langle (1,2), (1,0), (0,1) \rangle$ and $a(Y \cup Z) = \langle (2,1), (0,2) \rangle$ we have that

$$S_4(X \cup Z) = I(X \cup Z) + 2 > I(Y \cup Z) + 1 = S_4(Y \cup Z).$$

Relative Exposure It is well known that the Entropy index satisfies SI, SYM, GDP, and IND. By Theorem 1, it violates RE.

5 Appendix

The Isolation index can be written as in (7). Indeed, By (5),

$$\begin{aligned} I(X) &= 1 - \frac{\sum_{n \in N} \pi_n p_n (1 - p_n)}{P(1 - P)} \\ &= 1 - \sum_{n \in N} \frac{\frac{B_n + W_n}{B + W} \frac{B_n}{B_n + W_n} \frac{W_n}{B_n + W_n}}{\frac{B}{B+W} \frac{W}{B+W}} \\ &= 1 - \sum_{n \in N} (B + W) b_n w_n \frac{1}{(B_n + W_n)} \\ &= 1 - \sum_{n \in N} \frac{b_n w_n}{\pi_n}. \end{aligned}$$

References

Athey, S., B. Ferguson, M. Gentzkow, and T. Schmidt (2021): "Estimating experienced racial segregation in US cities using large-scale GPS data," *Proceedings of the National Academy of Sciences*, 118(46), e2026160118.

Bell, W. (1954): "A probability model for the measurement of ecological segregation," *Social Forces*, 32(4), 357–364.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor (1999): "The rise and decline of the American ghetto," *Journal of Political Economy*, 107(3), 455–506.

Echenique, F., and R. G. Fryer Jr (2007): "A measure of segregation based on social interactions," *The Quarterly Journal of Economics*, 122(2), 441–485.

Flückiger, Y., and J. Silber (2012): *The Measurement of Segregation in the Labor Force*. Springer Science & Business Media.

Frankel, D. M., and O. Volij (2011): "Measuring school segregation," *Journal of Economic Theory*, 146(1), 1–38.

Fuchs, V. R. (1975): "A note on sex segregation in professional occupations," in *Explorations in Economic Research, Volume 2, number 1*, pp. 105–111. NBER.

Gentzkow, M., and J. M. Shapiro (2011): "Ideological segregation online and offline," *The Quarterly Journal of Economics*, 126(4), 1799–1839.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019): "Measuring group differences in high-dimensional choices: method and application to congressional speech," *Econometrica*, 87(4), 1307–1340.

Hutchens, R. (2001): "Numerical measures of segregation: desirable properties and their implications," *Mathematical Social Sciences*, 42(1), 13–29.

——— (2004): "One measure of segregation," *International Economic Review*, 45(2), 555–578.

Jahn, J., C. F. Schmid, and C. Schrag (1947): "The measurement of ecological segregation," *American Sociological Review*, 12(3), 293–303.

James, D. R., and K. E. Taeuber (1985): "Measures of segregation," *Sociological Methodology*, 15, 1–32.

Massey, D. S., and N. A. Denton (1988): "The dimensions of residential segregation," *Social Forces*, 67(2), 281–315.

Monarrez, T. E. (2023): "School Attendance Boundaries and the Segregation of Public Schools in the United States," *American Economic Journal: Applied Economics*, 15(3), 210–237.

Mora, R., and J. 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, 147–179.

——— (2004): "Gender segregation by occupations in the public and the private sector. The case of Spain," *Investigaciones Económicas*, 28(3), 399–428.

Oosterbeek, H., S. Sóvágó, and B. van der Klaauw (2021): "Preference heterogeneity and school segregation," *Journal of Public Economics*, 197, 104400.

Philipson, T. (1993): "Social welfare and measurement of segregation," *Journal of Economic Theory*, 60(2), 322–334.

Puerta, C., and A. Urrutia (2016): "A characterization of the Gini segregation index," *Social Choice and Welfare*, 47, 519–529.

Reardon, S. F. (2011): "Measures of income segregation," *Unpublished Working Paper. Stanford Center for Education Policy Analysis.*

Serrati, P. S. (2024): "School and residential segregation in the reproduction of urban segregation: A case study in Buenos Aires," *Urban Studies*, 61(2), 313–330.

Theil, H. (1972): *Statistical Decomposition Analysis*. North-Holland Publishing, Amsterdam.

Theil, H., et al. (1971): *Principles of econometrics*. New York: Wiley.

Theil, H., and A. Finizza (1992): "A note on the measurement of racial integration in schools," *Lecture Notes in Economics and Mathematical Systems*, 1(389), 133–133.