INTERNATIONAL ECONOMIC REVIEW
Vol. 61, No. 4, November 2020 DOI: 10.1111/iere.12466

THE MEASUREMENT OF INCOME SEGREGATION*

By Casilda Lasso de la Vega and Oscar Volij[1]

University of the Basque Country, Spain; Ben Gurion University, Israel

We examine the problem of measuring the extent to which students with different income levels attend separate schools. Unless rich and poor attend the same schools in the same proportions, some segregation will exist. Since income is a continuous cardinal variable, however, the rich–poor dichotomy is necessarily arbitrary and renders any application of a binary segregation measure artificial. This article provides an axiomatic characterization of a measure of income segregation that takes into account the cardinal nature of income. This measure satisfies an empirically useful decomposition by subdistricts.

1. INTRODUCTION

Segregation is an attribute of school districts.[2] It refers to the extent to which pupils belonging to different demographic groups attend separate schools. Segregation measures compare districts which may differ both in the demographic distributions of their pupils and in the allocation of pupils across schools. When demographic groups are classified according to ethnicity, we are dealing with ethnic segregation. When they are classified according to gender, segregation is labeled as gender segregation. In this article, we are interested in income segregation, which can be observed when groups are classified according to income levels.

The criterion according to which we choose to classify individuals is not an innocuous one. When dealing with ethnicity or gender, for instance, there is no natural order of groups and indeed most of the ethnic segregation indices in the literature treat ethnic groups symmetrically. In other contexts, however, groups can be ordered according to some natural criterion. For example, pupils could be classified according to the educational level of their parents into having completed a primary, secondary or higher education. In these cases, it may not be appropriate to treat groups symmetrically, and in fact, indices have been developed that take into account the ordering of the groups. A richer context yet is the one of income segregation. Not only does income induce an order of the groups but it also induces a natural metric on them. Here too, segregation indices have been proposed that take into account the ordering of income levels and also their magnitude.[3]

School segregation, and its counterpart school diversity, are twin topics that regularly arise in political forums and in the media. Diversity and segregation are not restricted to race. In

*Manuscript received January 2019; revised October 2019.

[1] We are indebted to David M. Frankel. This article benefited greatly from his generous collaboration. We also thank the editor in charge for his useful comments. Finally, we thank the Spanish Ministerio de Economía y Competitividad (project ECO2015-67519-P) and the Gobierno Vasco (project IT1367-19) for research support. This research has used as a source of information databases of the Education Quality Agency of Chile (Agencia de Calidad de la Educación). The authors thank the Education Quality Agency for access to information. All the results of the study are the responsibility of the authors and do not compromise the institution. Please address correspondence to: Oscar Volij, Department of Economics, Ben-Gurion University, P.O.B 653, Beer-Sheva 84105 (Israel). E-mail: *ovolij@bgu.ac.il*.

[2] More generally, segregation is an attribute of a collection of organizational units. For expositional purposes, we focus on school districts, whose organizational units are, unsurprisingly, schools.

[3] For a necessarily incomplete list of ethnic segregation indices, see Massey and Denton (1988) and Reardon and Firebaugh (2002). For segregation among ordered categories, see, for example, Reardon (2009, 2011). For indices that exploit the cardinal nature of income, see Jargowsky (1996) and Kim and Jargowsky (2009).

the United States, for instance, programs exist that aim at increasing socioeconomic diversity in schools and creating more integrated public schools. Recently, in its concern that elite institutions enroll students who are diverse in every aspect except economically, the *New York Times* has developed the College Access Index, which attempts to measure economic diversity at top colleges, and which is published every year.

Recent empirical studies suggest that income segregation may affect educational outcomes. Students who have higher-quality peer groups tend to have better educational outcomes (Coleman et al., 1966), an effect for which evidence has been found to be causal (Hanushek et al., 2003; Imberman et al., 2012; Lavy et al., 2012). As pupils with higher family income tend to have higher ability, income segregation may be a significant source of differential peer effects across schools. Indeed, the findings of Mayer (2002) suggest that an increase in income segregation between census tracts or school districts tends to lower the achievement of low-ability pupils and raise that of high-ability pupils.

Despite the potential importance of income segregation, there is wide disagreement about how to measure it. Several income segregation indices have been proposed in the literature and some of their properties have been pointed out. Some researchers have used ethnic segregation indices, such as the Dissimilarity Index of Jahn et al. (1947). Other indices, notably the rank-order information theory index of Reardon (2011), take account of the ordinal nature of income categories. Finally, some indices treat income as a cardinal variable, the main example being Jargowsky's (1996) neighborhood sorting index.

Many attempts to measure income segregation consist of transforming a district in which individuals are classified by income into one in which they are classified by dichotomous categories so that a standard ethnic segregation index can be applied. For instance, Fong and Shibuya (2000) propose a poverty line to partition the population into rich and poor and apply an existing two-group segregation index. A more sophisticated approach measures income segregation by averaging the two-group segregation indices associated with all possible poverty lines. See, for instance, Reardon (2011) and the references therein. The main problem with these approaches is that small changes in pupils' incomes may induce large changes in the demographic distribution of both district and schools with the corresponding large change in income segregation. Hence, the resulting segregation orders fail to satisfy continuity, which is a property that any income segregation order should satisfy.

In this article, instead of directly adapting an existing ethnic segregation index to the context of income segregation, we adapt the *properties* of standard ethnic segregation measures to the new context and investigate their implications. Specifically, we show that these properties characterize a continuous index of income segregation, which we call the School Separation index. This index measures segregation as the difference between the district's variability and the average variability of its schools, variability being measured by the mean logarithmic deviation. To the best of our knowledge, this is the first axiomatic derivation of an income segregation measure.[4]

Before we move to the formal model, we discuss the concept of income segregation we have in mind and its relation to income inequality. Income segregation is an attribute of school districts, which may differ both in their allocation of pupils across schools and in the distribution of income across pupils. Although changes in the former will not affect the district's income inequality, changes in the latter will typically affect, however measured, its income segregation. Yet, some authors propose to disentangle the two concepts as much as possible. Reardon (2011), for instance, proposes that income segregation be maximal if and only if within each school, all pupils have the same income, no matter what the income distribution of the district may be. In order to illustrate this requirement, which Reardon (2011) calls *scale interpretability*, consider the following districts:

[4] Measures of segregation among unordered categories such as ethnic groups have been axiomatized by Echenique and Fryer (2007), Frankel and Volij (2011), and Hutchens (2001, 2004).

X	\$10	\$20	Y	\$10	$\$20 \times 10^6$
School 1	100	0	School 1	100	0
School 2	0	100	School 2	0	100

Both districts have two schools, one attended by the rich and the other by the poor. However, whereas the poor in both districts have an income of \$10, the rich have an income of \$20 in district X and an income of \$20 million in district Y. By virtue of scale interpretability, they are equally and maximally segregated. This is so even though the difference between rich and poor in X is negligible compared to the corresponding difference in Y. The idea of income segregation that we have in mind, however, is inconsistent with the above requirement. In fact our axioms will imply that district X exhibits less income segregation than district Y since, although in both districts poor and rich attend separate schools, district Y exhibits a much higher income inequality than X. In other words, according to our concept of income segregation, the extent to which students with different incomes attend different schools is magnified by the inequality of students' incomes.

Indices that satisfy scale interpretability aim to capture a concept of *pure segregation*, namely, one that is independent of income inequality. In contrast, the idea of segregation we are aiming for in this article is a hybrid one in which pure segregation and income inequality interact. Needless to say, both kinds of indices may provide useful information.

In order to further illustrate the difference between the concept of pure income segregation and the one that we propose, consider the following two districts:

X	\$200	\$300	Y	\$100	\$200	\$300	\$400
School 1	20	0	School 1	10	10	0	0
School 2	0	20	School 2	0	0	10	10

District X consists of two schools, one attended by the rich and one attended by the poor. Since all the poor have an income of \$200 and all the rich have an income of \$300, according to the above requirement, X has maximum segregation. If we now make half the poor even poorer and half the rich even richer, we obtain district Y. According to scale interpretability segregation is reduced. The reason for this reduction is that although nobody moved from one school to the other, and although the poor and the rich still go to separate schools, both schools became "more diverse" as a result of the pauperization of half of the already poor and of the enrichment of half of the already rich. In contrast, according to our notion of segregation, the increase in income inequality observed in the transition from X to Y may magnify instead of reduce the income segregation already existing in X; the poor and the rich still attend separate schools, and the difference between rich and poor became more striking.

The article is organized as follows. After introducing the basic notation in Section 2, Section 3 presents, among others, the income segregation indices that are the focus of the article. In Section 4, we propose a list of axioms which are used in Section 5 to present our characterization result. Finally, Section 6 offers an empirical illustration of the behavior of the School Separation Index (SSI).

2. NOTATION

As mentioned in the Introduction, segregation is an attribute of a collection of organizational units, such as schools or neighborhoods, which are taken as given. For empirical purposes, segregation measures are most reliable when there is a natural choice of these units. In the case of residential segregation, for instance, census tracts seem to be arbitrary and consequently, the resulting segregation measurements are vulnerable to the *modifiable areal unit problem*; slight changes in the definition of an organizational unit may lead to large changes in measured segregation. In the case of school segregation, however, schools seem an unquestionable choice as organizational units. For this reason, we focus on school segregation.[5]

A *school* is a finite collection of pairs $\langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$, where for each $g = 1, \ldots, G$, $y_g > 0$ is an income level and $n_g \geq 0$ represents the number of pupils with income level y_g. Pairs (n_g, y_g) will be called income groups. For notational convenience, we allow schools to have two income groups with the same income. In that case, however, if they are combined, the school does not change; for example, the schools $\langle (n, y), (n', y) \rangle$ and $\langle (n + n', y) \rangle$ are regarded as the same school. Also, if we permute the income groups the school does not change; that is, for any permutation $\pi : \{1, \ldots, G\} \to \{1, \ldots, G\}$, $\langle (n_1, y_1), \ldots, (n_G, y_G) \rangle = \langle (n_{\pi(1)}, y_{\pi(1)}), \ldots, (n_{\pi(G)}, y_{\pi(G)}) \rangle$.

For any school $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$, let $|c| = \sum_{g=1}^{G} n_g y_g$ denote the total income of school c, and $n_c = \sum_{g=1}^{G} n_g$ its total enrollment. If $n_c = 0$, c is an empty school. Empty schools will play no role in the article but are needed for notational convenience. If c is not empty, we denote by $\mu_c = |c|/n_c$ its mean income, and by $\overline{c} = \langle (n_c, \mu_c) \rangle$ the *smoothed* school that is obtained from c by redistributing c's total income equally among its pupils. For any school $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$ and scalar $\lambda > 0$, let $\lambda c = \langle (\lambda n_1, y_1), \ldots, (\lambda n_G, y_G) \rangle$ denote the school that is obtained from c by multiplying the number of people in each income group by λ and let $c * \lambda = \langle (n_1, \lambda y_1), \ldots, (n_G, \lambda y_G) \rangle$ denote the school that is obtained from c by multiplying each pupil's income by λ. For any two schools $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$ and $c' = \langle (n'_1, y'_1), \ldots, (n'_{G'}, y'_{G'}) \rangle$, let $c \mathbin{+\!\!+} c' = \langle (n_1, y_1), \ldots, (n_G, y_G), (n'_1, y'_1), \ldots, (n'_{G'}, y'_{G'}) \rangle$ denote the result of combining the two schools into a single school. We say that a sequence of schools $c^m = \langle (n_1^m, y_1^m), \ldots, (n_G^m, y_G^m) \rangle$ converges to school $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$, denoted $c^m \to c$, if for all $g = 1, \ldots, G$, the sequence of pairs (n_g^m, y_g^m) converges to (n_g, y_g). We denote by $\mathcal{C}$ the class of nonempty schools where pupils have positive incomes.

A *district* $\{c_1, \ldots, c_K\}$ is a finite collection of schools at least one of which is not empty. We identify any district with the district that is obtained from it by deleting all its empty schools. If we permute the schools, the district does not change; for example, for any permutation $\pi : \{1, \ldots, K\} \to \{1, \ldots, K\}$, $\{c_1, \ldots, c_K\} = \{c_{\pi(1)}, \ldots, c_{\pi(K)}\}$. With some abuse of notation we will denote a typical district by $X = \{c_k\}_{k \in K}$. For any district X, let $n_X = \sum_{c \in X} n_c$ denote the total attendance of X, let $|X| = \sum_{c \in X} |c|$ denote its total income, and let $\mu_X = |X|/n_X$ denote its mean income. For any district X and scalar $\lambda > 0$, let $\lambda X = \{\lambda c\}_{c \in X}$ denote the district that is obtained from X by multiplying the number of people in each school by λ and let $X * \lambda = \{c * \lambda\}_{c \in X}$ denote the district that is obtained from X by multiplying each pupil's income by λ. For any two districts $X = \{c_1, \ldots, c_K\}$ and $Y = \{c'_1, \ldots, c'_{K'}\}$, let $X \uplus Y = \{c_1, \ldots, c_K, c'_1, \ldots, c'_{K'}\}$ denote the district that results from combining the schools of X and Y into a single district. We denote by $\mathcal{D}$ the set of all districts where all students have positive incomes.

A district is *simple* if it is of the form $\{\langle (n_1, y_1) \rangle, \ldots, \langle (n_K, y_K) \rangle\}$ for some K, that is, if each school contains a single income group. For any district $X = \{c_1, \ldots, c_K\}$, let $c(X) = c_1 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c_K$ be the school that is obtained by combining the schools of X into a single school, and let $R(X) = \{\frac{n_{c_1}}{n_X} c(X), \ldots, \frac{n_{c_K}}{n_X} c(X)\}$ be the district that is obtained from X by reallocating its students so that each school maintains its enrollment and all schools have the same income distribution (that of X). We will henceforth call $R(X)$ the completely integrated version of X. For instance, if $c_1 = \langle (3, 1), (3, 2) \rangle$ and $c_2 = \langle (12, 4) \rangle$ are two schools, and $X = \{c_1, c_2\}$, then $c(X) = c_1 \mathbin{+\!\!+} c_2 = \langle (3, 1), (3, 2), (12, 4) \rangle$, and $R(X) = \{\langle (1, 1), (1, 2), (4, 4) \rangle, \langle (2, 1), (2, 2), (8, 4) \rangle\}$.

[5] Also, school segregation, as opposed to residential segregation, is a context in which geography seems not to play a role. For an attempt to measure segregation taking geography into account, see Dawkins (2007).

3. SEGREGATION AND INEQUALITY

3.1. *Inequality Indices.* Income segregation is related to income inequality in two ways. On the one hand, the higher the income inequality of a district is, the higher the potential for income segregation in it. On the other hand, ceteris paribus, for any given level of income inequality of a district, the more economically diverse are its schools, the lower its income segregation. Given this relation, before we define measures of income segregation, we need to introduce indices of income inequality.

An *inequality index* I assigns to each school c a real number, $I(c)$ which is meant to capture its level of income inequality. The following are examples of prominent income inequality indices. The first one consists of the class of generalized entropy indices and the second is the variance (see Shorrocks, 1980).

EXAMPLE 1. For $\alpha \in \mathbb{R}$ the *Generalized Entropy index*, $I^\alpha : \mathcal{C} \to [0, \infty)$, is defined as follows: For all $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle \in \mathcal{C}$,

$$I^\alpha(c) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \sum_{g=1}^{G} \frac{n_g}{n_c} \left[\left(\frac{y_g}{\mu_c} \right)^\alpha - 1 \right] & \text{if } \alpha \notin \{0,1\} \\ \sum_{g=1}^{G} \frac{n_g}{n_c} \ln \left(\frac{\mu_c}{y_g} \right) & \text{if } \alpha = 0 \\ \sum_{g=1}^{G} \frac{n_g y_g}{|c|} \ln \left(\frac{y_g}{\mu_c} \right) & \text{if } \alpha = 1. \end{cases}$$

When $\alpha = 0$, the associated generalized entropy index I^0 is known as *Theil's second measure of income inequality* (Theil, 1967).

EXAMPLE 2. The *variance* assigns to each school the variance of its (absolute) income distribution. Formally, var, is defined as follows: For all $c = \langle (n_1, y_1), \ldots, (n_G, y_G) \rangle$,

$$\text{var}(c) = \frac{1}{n_c} \sum_{g=1}^{G} n_g (y_g - \mu_c)^2.$$

We shall sometimes speak of the income inequality in the whole district, and to measure it, we will apply an inequality index to the combination of all its schools into a single school. In particular, with a slight abuse of notation for any district $X = \{c_1, \ldots, c_K\}$, we will write $I(X)$ for $I(c_1 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c_K)$, the inequality of the district's income distribution.

3.2. *Segregation Orders and Indices.* A segregation order defined on $\mathcal{D}$ is a complete and transitive relation $\succcurlyeq$ on $\mathcal{D}$. An *income segregation index*, or segregation index for short, $\mathcal{S}$ assigns to each district, X a real number, $\mathcal{S}(X)$, which is meant to capture its level of segregation. We shall maintain the convention of using calligraphic capital letters to denote segregation indices. A segregation index represents a segregation order $\succcurlyeq$ if for any two districts $X, Y, X \succcurlyeq Y \Leftrightarrow \mathcal{S}(X) \geq \mathcal{S}(Y)$. The following are examples of income segregation indices. For any district $X = \{c_1, \ldots, c_K\}$,

- the *school separation index* is defined by

$$\mathcal{SSI}(X) = \sum_{c \in X} \frac{n_c}{n_X} \ln \left(\frac{\mu_X}{\mu_c} \right);$$

- the *Variance segregation index* is defined by

$$\mathcal{V}(X) = \frac{1}{n_X} \sum_{c \in X} n_c (\mu_c - \mu_X)^2;$$

- *Jargowsky's neighborhood sorting index* is defined (on the class of districts whose income distribution has positive variance) by [6]

$$\mathcal{NSI}(X) = \sqrt{\frac{\mathcal{V}(X)}{\text{var}(X)}}.$$

Finally, given any income inequality index I,

- the *segregation index induced by I* is defined by

$$\mathcal{I}(X) = I(X) - \sum_{c \in X} \frac{n_c}{n_X} I(c).$$

In order to understand the idea behind the last class of indices, note that the sum $\sum_{c \in X} \frac{n_c}{n_X} I(c)$ is an average of the level of income inequality, as measured by I, within the schools of X, and can be seen as a measure of the economic diversity of such schools. Clearly, this diversity cannot contribute to the segregation of X. Thus, the segregation of X as measured by $\mathcal{I}$ is what remains from the district's income inequality after we deduct the economic diversity exhibited by the schools.

An interesting feature of the segregation index induced by I is that when each school has zero income variation, the district's segregation coincides with its income inequality as measured by I, and as a result the higher the income inequality is, the higher the district's segregation is. This means that the segregation index induced by I is not pure since it does not fulfill the requirement that segregation be maximal when schools exhibit no diversity. Nevertheless, one could use the segregation index induced by I to define a measure of pure segregation by the ratio $\mathcal{I}(X)/I(X)$. With this definition, we see that the segregation index induced by an inequality index I is in fact the product of an index of pure segregation and the associated income inequality of the district.

Interestingly, the segregation index induced by the generalized entropy index I^0 is $\mathcal{SSI}$. To see this, note that

$$\begin{aligned}
I^0(X) &= I^0(c_1 \mathbin{+\!\!\!+} \cdots \mathbin{+\!\!\!+} c_K) \\
&= \sum_{c \in X} \sum_{g=1}^{G(c)} \frac{n_g}{n_X} \ln\left(\frac{\mu_X}{y_g}\right) \\
&= \sum_{c \in X} \frac{n_c}{n_X} \sum_{g=1}^{G(c)} \frac{n_g}{n_c} \ln\left(\frac{\mu_X}{\mu_c}\frac{\mu_c}{y_g}\right) \\
&= \sum_{c \in X} \frac{n_c}{n_X} \ln\left(\frac{\mu_X}{\mu_c}\right) + \sum_{c \in X} \frac{n_c}{n_X} \sum_{g=1}^{G(c)} \frac{n_g}{n_c} \ln\left(\frac{\mu_c}{\mu_g}\right) \\
&= \mathcal{SSI}(X) + \sum_{c \in X} \frac{n_c}{n_X} I^0(c),
\end{aligned}$$

which, after a rearrangement of terms, yields the desired result. A similar proof shows that the segregation index induced by the variance, *var*, is $\mathcal{V}$. [7]

[6] Kim and Jargowski (2009) propose an analogous segregation index, where the standard deviation is replaced by the Gini coefficient.

[7] This implies that Jargowsky's $\mathcal{NSI}$ is ordinally equivalent to the pure segregation index induced by *var*.

4. AXIOMS

We now list several desirable properties of an income segregation order. We start with three fundamental axioms that convey the basic idea of what it means for a district to be segregated. Specifically, they express the idea that there cannot be segregation unless there are at least two schools with different income distributions. Recall that for any district X, $R(X)$ is the district that is obtained from X by reallocating its pupils so that all schools keep their original enrollment, and share the same relative income distribution. The first axiom requires that if pupils are reallocated so that all schools have the same income distributions, segregation does not increase.

Equal Allocation Property (EAP): For any district X, $X \succcurlyeq R(X)$.

EAP is a very weak axiom. It restricts $\succcurlyeq$ only when comparing a district to its completely integrated version. It says nothing about districts with different allocations of pupils across income groups, or with different number of schools. Also, it seems a very natural requirement for any segregation measure to satisfy. Indeed, one would expect a reallocation of pupils that is independent of their income not to result in an increased level of segregation.

The next axiom identifies a class of districts all of whose members exhibit the same level of segregation.

Equivalence of Single-School Districts (SSD): If X and Y are single-school districts, then $X \sim Y$.

The next axiom refers to the class of (simple) districts that have an egalitarian income distribution. A district $X = \{c_1, \ldots, c_K\}$ is said to have an egalitarian income distribution if all pupils have the same income, namely, if $c_k = \langle (n_k, \mu_X) \rangle$ for $k = 1, \ldots K$. The axiom says that all such districts are equally segregated, and more segregated than any other simple district.

Equivalence of Uniform Distribution Districts (UDD): Let X and Y be two simple districts with the same number of pupils and the same income. Assume further that X has an egalitarian income distribution. Then, $Y \sim X$ if and only if Y also has an egalitarian income distribution.

An immediate consequence of these two axioms is that the egalitarian districts and the single-school districts are all equally segregated.

Next, we list two axioms that require invariance to certain changes in units of measurement. The first one states that changes in population that leave the relative attendances of the schools unchanged do not affect segregation.

Population Homogeneity (PH): For any district X and scalar $\lambda > 0$, $X \sim \lambda X$.

The next axiom states that changes in household incomes that keep the students' relative incomes unchanged do not affect segregation.

Income Homogeneity (IH): For any district X and scalar $\lambda > 0$, $X \sim X * \lambda$.

It can be easily checked that $\mathcal{SSI}$ satisfies this and the previous axioms.

The next two axioms require segregation comparisons to be independent of irrelevant subdistricts. In order to motivate the first one, consider a school district partitioned into two subdistricts. Suppose that a reorganization within each subdistrict reduces segregation in every one of them. Though not unimaginable, it is reasonable to expect that such a reorganization does not result in a higher districtwide segregation. Otherwise we would be witnessing a rather perverse outcome of an otherwise well-intended policy. The next axiom requires that no such outcomes can ever occur.

Independence (IND): For any two districts X, Y with the same population and the same total income, and for any arbitrary district Z, $X \succcurlyeq Y \Leftrightarrow X \uplus Z \succcurlyeq Y \uplus Z$.

Independence guarantees that any policy that reduces segregation in one subdistrict does not result in a higher districtwide segregation. Versions of this axiom appear in several contexts. For instance, Hutchens (2001) and Frankel and Volij (2011) use variations of this axiom in their characterizations of ethnic segregation measures. Shorrocks (1988) and Foster and Shorrocks (1991) subgroup consistency axioms are essentially the independence axiom in the context of income inequality and poverty measurement, respectively.

The $\mathcal{SSI}$ satisfies IND. To see that this, let X, Y, and Z be three districts as described in the axiom. Denoting $n_X = n_Y = n$ and $\mu_X = \mu_Y = \mu$, and taking into account that $n_{X\uplus Z} = n_{Y\uplus Z}$ and $\mu_{X\uplus Z} = \mu_{Y\uplus Z}$, we have

$$\begin{aligned}
\mathcal{SSI}(X \uplus Z) \geq \mathcal{SSI}(Y \uplus Z) &\Leftrightarrow \sum_{c\in X\uplus Z} \frac{n_c}{n_{X\uplus Z}} \ln\left(\frac{\mu_{X\uplus Z}}{\mu_c}\right) \geq \sum_{c'\in Y\uplus Z} \frac{n_{c'}}{n_{Y\uplus Z}} \ln\left(\frac{\mu_{Y\uplus Z}}{\mu_{c'}}\right) \\
&\Leftrightarrow \sum_{c\in X} \frac{n_c}{n_{X\uplus Z}} \ln\left(\frac{\mu_{X\uplus Z}}{\mu_c}\right) \geq \sum_{c'\in Y} \frac{n_{c'}}{n_{Y\uplus Z}} \ln\left(\frac{\mu_{Y\uplus Z}}{\mu_{c'}}\right) \\
&\Leftrightarrow \sum_{c\in X} \frac{n_c}{n} \ln\left(\frac{\mu_{X\uplus Z}}{\mu}\frac{\mu}{\mu_c}\right) \geq \sum_{c'\in Y} \frac{n_{c'}}{n} \ln\left(\frac{\mu_{Y\uplus Z}}{\mu}\frac{\mu}{\mu_{c'}}\right) \\
&\Leftrightarrow \ln\left(\frac{\mu_{X\uplus Z}}{\mu}\right) + \sum_{c\in X} \frac{n_c}{n} \ln\left(\frac{\mu}{\mu_c}\right) \geq \ln\left(\frac{\mu_{Y\uplus Z}}{\mu}\right) + \sum_{c'\in Y} \frac{n_{c'}}{n} \ln\left(\frac{\mu}{\mu_{c'}}\right) \\
&\Leftrightarrow \sum_{c\in X} \frac{n_c}{n} \ln\left(\frac{\mu}{\mu_c}\right) \geq \sum_{c'\in Y} \frac{n_{c'}}{n} \ln\left(\frac{\mu}{\mu_{c'}}\right) \\
&\Leftrightarrow \mathcal{SSI}(X) \geq \mathcal{SSI}(Y).
\end{aligned}$$

Though similar, the next axiom is different from independence. Consider a district composed of two subdistricts $X \uplus Y$ and assume that a policy is applied to Y, thereby transforming it into Z. Further assume that this policy left attendance unchanged. The axiom states that whether or not this policy increases districtwide segregation does not depend on the segregation within subdistrict X.

Separability (SEP): For any three districts X, Y, Z such that $n_Y = n_Z$,

$$X \uplus Y \succcurlyeq X \uplus Z \Leftrightarrow R(X) \uplus Y \succcurlyeq R(X) \uplus Z.$$

This is a weak axiom, at least to the extent that any segregation index induced by an inequality index, in particular $\mathcal{SSI}$, satisfies separability. To see this, consider three districts as described in the axiom, and let us denote $n = n_X + n_Y = n_X + n_Z$. Then,

$$\begin{aligned}
\mathcal{I}(X \uplus Y) \geq \mathcal{I}(X \uplus Z) &\Leftrightarrow I(X \uplus Y) - \sum_{c\in X\uplus Y} \frac{n_c}{n} I(c) \geq I(X \uplus Z) - \sum_{c\in X\uplus Z} \frac{n_c}{n} I(c) \\
&\Leftrightarrow I(X \uplus Y) - \sum_{c\in Y} \frac{n_c}{n} I(c) \geq I(X \uplus Z) - \sum_{c\in Z} \frac{n_c}{n} I(c) \\
&\Leftrightarrow I(R(X) \uplus Y) - \sum_{c\in R(X)\uplus Y} \frac{n_c}{n} I(c) \geq I(R(X) \uplus Z) - \sum_{c\in R(X)\uplus Z} \frac{n_c}{n} I(c) \\
&\Leftrightarrow \mathcal{I}(R(X) \uplus Y) \geq \mathcal{I}(R(X) \uplus Z).
\end{aligned}$$

The last axiom requires that similar districts have similar levels of segregation. It will allow us to find a continuous representation of an order that satisfies the previous axioms.

Continuity (CONT): Let $X = \{c_1, \dots, c_K\}$ be a district and let $X^n = \{c_1^n, \dots, c_K^n\}$, for $n = 1, 2, \dots$ be a sequence of districts such that $c_k^n \to c_k$ for $k = 1, \dots, K$. For any district Y, if $X^n \succcurlyeq Y$ for all n, then $X \succcurlyeq Y$, and if $Y \succcurlyeq X^n$ for all n, then $Y \succcurlyeq X$.

All of the indices mentioned above satisfy this axiom.

5. AN ORDINAL CHARACTERIZATION OF $\mathcal{SSI}$

We now state our main result.

THEOREM 1. *Let $\succcurlyeq$ be a segregation order on $\mathcal{D}$. It satisfies the equal allocation property, equivalence of single-school districts, equivalence of uniform-distribution districts, independence, separability, population homogeneity, income homogeneity, and continuity if and only if it is represented by the school separation index. Namely, $X \succcurlyeq Y \Leftrightarrow \mathcal{SSI}(X) \geq \mathcal{SSI}(Y)$ for all districts X, Y.*

PROOF. As was shown earlier, the order represented by $\mathcal{SSI}$ satisfies all the axioms listed in Theorem 1. We now show that the only order that satisfies this list is $\mathcal{SSI}$. Let $\succcurlyeq$ be an order that satisfies all the axioms listed in Theorem 1. The proof consists of four steps. First, we build an index $\mathcal{S}$ that represents $\succcurlyeq$. Second, we prove that $\mathcal{S}$ satisfies a very strong separability property. Third, we show that when restricted to the family of simple districts, $\mathcal{S}$ satisfies several properties and, as a result, it has a particular form. Finally, we show that the only extension of this restricted index to the class of all districts, if is to satisfy all the axioms, and in particular the separability property uncovered in the second step, is $\mathcal{SSI}$.

The following piece of notation will be useful. For any school $c = \langle (n_1, y_1), \dots, (n_G, y_G) \rangle$, let $d(c) = \{\langle (n_g, y_g) \rangle\}_{g \in G}$ denote the simple district that results from placing each income group in c into its own school. Also, for a district $X = \{c_1, \dots, c_K\}$, let $d(X) = \uplus_{c \in X} d(c)$, namely, the district that results from applying the operation d to each school in X.

Let $\succcurlyeq$ be a segregation order on $\mathcal{D}$ that satisfies all the forgoing axioms. The next claim shows that merging schools with the same distribution of income does not affect segregation.

CLAIM 1. *For any district X and for any $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$, $\alpha X \uplus \beta X \sim (\alpha + \beta) X$.*

PROOF. Let $X = \{c_1, \dots, c_K\}$ be a district and let $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$.

$$\begin{aligned} \alpha X \uplus \beta X &= \{\alpha c_1, \dots, \alpha c_K\} \uplus \{\beta c_1, \dots, \beta c_K\} && \text{by definition} \\ &\sim \{\alpha \bar{c}_1, \dots, \alpha \bar{c}_K\} \uplus \{\beta \bar{c}_1, \dots, \beta \bar{c}_K\} && \text{by SSD and IND} \\ &\sim \uplus_{i=1}^K \{\alpha \bar{c}_i, \beta \bar{c}_i\}. \end{aligned}$$

By UDD we have that $\{\alpha \bar{c}_i, \beta \bar{c}_i\} \sim \{(\alpha + \beta)\bar{c}_i\}$ for $i = 1, \dots K$. Therefore, by IND and SSD

$$\begin{aligned} \uplus_{i=1}^K \{\alpha \bar{c}_i, \beta \bar{c}_i\} &\sim \uplus_{i=1}^K \{(\alpha + \beta)\bar{c}_i\} \\ &\sim \uplus_{i=1}^K \{(\alpha + \beta) c_i\} = (\alpha + \beta) X. \end{aligned}$$

□

As a corollary we obtain that districts whose schools have the same income distribution are as segregated as single-school districts. As a result of EAP then, these districts exhibit the minimum level of segregation.

COROLLARY 1. *For any X and for any single-school district $\{c\}$, we have that $R(X) \sim \{c\}$.*

PROOF. Note that $R(X) = \alpha_1\{c(X)\} \uplus \cdots \uplus \alpha_K\{c(X)\}$, where $\alpha_k = n_{c_k}/n_X$. By Claim 1 and by SSD

$$R(X) = \alpha_1\{c(X)\} \uplus \cdots \uplus \alpha_K\{c(X)\} \sim \{(\alpha_1 + \cdots + \alpha_K)c(X)\} \sim \{c\}.$$

□

We now start building the index. Let $\mathcal{D}_1$ denote the class of districts X with $n_X = |X| = 1$. Also let $X_0 = \{\langle(1, 1)\rangle\}$ be the district with a single school, which has a single student with income 1. Note that by Corollary 1 and EAP, $X \succcurlyeq X_0$ for all districts X.

LEMMA 1. *Let $X' \in \mathcal{D}_1$ be a district such that $X' \succ X_0$. If $0 \leq \alpha < \beta < 1$, then $\beta X' \uplus (1 - \beta)X_0 \succ \alpha X' \uplus (1 - \alpha)X_0$.*

PROOF. By PH, $(\beta - \alpha)X' \succ (\beta - \alpha)X_0$. By IND,

$$\alpha X' \uplus (\beta - \alpha)X' \uplus (1 - \beta)X_0 \succ \alpha X' \uplus (\beta - \alpha)X_0 \uplus (1 - \beta)X_0.$$

By Claim 1 and IND, $\beta X' \uplus (1 - \beta)X_0 \succ \alpha X' \uplus (1 - \alpha)X_0$. □

LEMMA 2. *Let X' be a district in $\mathcal{D}_1$ such that $X' \succ X_0$. For any district X such that $X' \succcurlyeq X$, there is a unique $\alpha' \in [0, 1]$ such that $X \sim \alpha' X' \uplus (1 - \alpha')X_0$.*

PROOF. The sets $\{\alpha \in [0, 1] : \alpha X' \uplus (1 - \alpha)X_0 \succcurlyeq X\}$ and $\{\alpha \in [0, 1] : X \succcurlyeq \alpha X' \uplus (1 - \alpha)X_0\}$ are closed by CONT. Since $X' \succcurlyeq X \succcurlyeq X_0$, they are not empty. Since $\succcurlyeq$ is complete, their union is [0,1]. Therefore, since the unit interval is connected, the intersection of the two sets is not empty. By Lemma 1, this intersection must contain a single element. This single element is the α' we are looking for. □

LEMMA 3. *Let X' and X'' be two districts in $\mathcal{D}_1$ such that $X'' \succcurlyeq X' \succ X_0$. Let X be a district such that $X' \succcurlyeq X$, and let α' and α'' be the unique numbers identified in Lemma 2 defined, respectively, by $X \sim \alpha' X' \uplus (1 - \alpha')X_0$ and $X \sim \alpha'' X'' \uplus (1 - \alpha'')X_0$. Let β be the unique number identified in Lemma 2 such that $X' \sim \beta X'' \uplus (1 - \beta)X_0$. Then, $\alpha'' = \alpha'\beta$.*

PROOF. By definition of α' and IND, $X \sim \alpha'(\beta X'' \uplus (1 - \beta)X_0) \uplus (1 - \alpha')X_0$. Therefore, by Claim 1, $X \sim \alpha'\beta X'' \uplus (1 - \alpha'\beta)X_0$. □

We can now proceed to the definition of a segregation index. Fix the following district: $X_{1/2} = \{\langle(1/2, 1/2)\rangle, \langle(1/2, 3/2)\rangle\} \in \mathcal{D}_1$. Let X be a district, and let $X' \in \mathcal{D}_1$ be a district that satisfies $X' \succcurlyeq X$ and $X' \succcurlyeq X_{1/2}$. Let α' and β' be the unique numbers identified in Lemma 2 that satisfy $X \sim \alpha' X' \uplus (1 - \alpha')X_0$ and $X_{1/2} \sim \beta' X' \uplus (1 - \beta')X_0$. Note that by UDD, $X_{1/2} \succ X_0$, and as a result, $\beta' > 0$. We can thus assign to every district X the number α'/β'. It turns out that this number does not depend on the choice of X'. Indeed, let $X'' \in \mathcal{D}_1$ be another district such that $X'' \succcurlyeq X$ and $X'' \succcurlyeq X_{1/2}$ and let α'' and β'' be defined by $X \sim \alpha'' X'' \uplus (1 - \alpha'')X_0$ and $X_{1/2} \sim \beta'' X'' \uplus (1 - \beta'')X_0$. Assume without loss of generality that $X'' \succcurlyeq X'$. Let δ be defined by $X' \sim \delta X'' \uplus (1 - \delta)X_0$. By Lemma 3, $\alpha'' = \alpha'\delta$ and $\beta'' = \beta'\delta$. Therefore, $\alpha'/\beta' = \alpha''/\beta''$. The above discussion allows us to define the segregation index $\mathcal{S}$ by $\mathcal{S}(X) = \alpha'/\beta'$, where α'/β' is the ratio built above.

LEMMA 4. *The index $\mathcal{S}$ represents the segregation order $\succcurlyeq$.*

PROOF. Let X and X' be two districts and assume that $X' \succ X$. Let $X'' \in \mathcal{D}_1$ be a district such that $X'' \succcurlyeq X'$ and $X'' \succcurlyeq X_{1/2}$. Let α and α' be defined by $X \sim \alpha X'' \uplus (1 - \alpha)X_0$ and $X' \sim \alpha' X'' \uplus (1 - \alpha')X_0$. By Lemma 1, $\alpha' > \alpha$, which implies that $\mathcal{S}(X') > \mathcal{S}(X)$. □

The following property follows from the way the index was constructed and by SSD. The proof is left to the reader.

CLAIM 2. *For any district X, $\mathcal{S}(X) \geq 0$. Furthermore, if X is a single-school district, then $\mathcal{S}(X) = 0$.*

We now start the second part of the proof. The next proposition shows that the index $\mathcal{S}$ satisfies a very strong separability property.

PROPOSITION 1. *Let X and X' be two districts. Then,*

$$\mathcal{S}(X \uplus X') = \frac{n_X}{n_{X \uplus X'}}\mathcal{S}(X) + \mathcal{S}(R(X) \uplus X').$$

PROOF. Let X and X' be two districts with populations $n_X = n$ and $n_{X'} = m$, respectively. By PH, we can assume without loss of generality that $n + m = 1$. Since $\succcurlyeq$ satisfies IH, we can also assume without loss of generality that $|X \uplus X'| = 1$. Let X'' be a district in $\mathcal{D}_1$ such that $X'' \succcurlyeq X$, $X'' \succcurlyeq X \uplus X'$ and $X'' \succcurlyeq X_{1/2}$, and let α, γ and δ be such that

$$X \sim \alpha X'' \uplus (1-\alpha)X_0, \tag{1}$$

$$R(X) \uplus X' \sim \gamma X'' \uplus (1-\gamma)X_0, \tag{2}$$

$$X \uplus X' \sim \delta X'' \uplus (1-\delta)X_0. \tag{3}$$

Then, $S(X) = \alpha/\beta$, $S(R(X) \uplus X') = \gamma/\beta$, and $S(X \uplus X') = \delta/\beta$ for some $\beta > 0$. In order to prove the result it is enough to show that $\delta = n\alpha + \gamma$.

Denote $X_0^* = \{\langle (n, \frac{|X|}{n}) \rangle\}$. This district has the same population and income as X and it is obtained from X_0 by multiplying its population by n and the income of each pupil by $|X|/n$. Also denote $X^* = nX'' * (|X|/n)$. This district has the same population and income as X. It is obtained from X'' by multiplying its population by n and by multiplying the income of each pupil by $|X|/n$. It follows from 1, using PH and IH, that

$$X \sim \alpha X^* \uplus (1-\alpha)X_0^*. \tag{4}$$

Choose $k \in \mathbb{N}$ such that $k > n + \gamma$. By concatenating $(k-1)X_0$ to both sides of Equation (2), we obtain

$$\begin{array}{llll}
R(X) \uplus \overbrace{X' \uplus (k-1)X_0}^{Y} & \sim & \gamma X'' \uplus (k-\gamma)X_0 & \text{by IND and Claim 1} \\
& \sim & \frac{\gamma}{n}X^* \uplus \frac{k-\gamma}{n}X_0^* & \text{by IH} \\
& \sim & \frac{\gamma}{n}X^* \uplus \frac{k-\gamma}{n}R(X) & \text{by Corollary 1 and IND} \\
& \sim & R(X) \uplus \overbrace{\frac{\gamma}{n}X^* \uplus \left(\underbrace{\frac{k-\gamma}{n} - 1}_{>0}\right) R(X)}^{Z} & \text{by Claim 1.}
\end{array}$$

Note that since $k > n+\gamma$ subdistrict Z is well defined. Since $n_Y = n_Z = m + (k-1)$, by SEP,

$$X \uplus \overbrace{X' \uplus (k-1)X_0}^{Y} \sim X \uplus \overbrace{\frac{\gamma}{n}X^* \uplus (\frac{k-\gamma}{n} - 1)R(X)}^{Z}.$$

By Equations (3), (4), and IND,

$$\begin{aligned} \delta X'' \uplus (1-\delta)X_0 \uplus (k-1)X_0 &\sim \alpha X^* \uplus (1-\alpha)X_0^* \uplus \frac{\gamma}{n}X^* \uplus \left(\frac{k-\gamma}{n} - 1\right) R(X) \\ &\sim \alpha X^* \uplus (1-\alpha)X_0^* \uplus \frac{\gamma}{n}X^* \uplus \left(\frac{k-\gamma}{n} - 1\right) X_0^* \\ &\sim n\alpha X'' \uplus n(1-\alpha)X_0 \uplus \gamma X'' \uplus (k-\gamma-n)X_0, \end{aligned}$$

where the second line follows from IND and Corollary 1, and the last one from IH. Applying Claim 1 to both sides, we obtain

$$\delta X'' \uplus (k-\delta)X_0 \sim (n\alpha + \gamma)X'' \uplus (k-\gamma-n\alpha)X_0.$$

By PH and Lemma 2, we conclude that $\delta = n\alpha + \gamma$. □

COROLLARY 2. *Let $X_1, \ldots, X_J$ be J districts. Then,*

$$\mathcal{S}\left(\uplus_{j=1}^{J} X_j\right) = \sum_{j=1}^{J} \frac{n_{X_j}}{n_X}\mathcal{S}(X_j) + \mathcal{S}\left(\uplus_{j=1}^{J} R(X_j)\right). \tag{5}$$

PROOF. See the Appendix. □

COROLLARY 3. *For any district $X = \{c_1, \ldots, c_K\}$,*

$$\mathcal{S}(X) = \mathcal{S}(d(c_1 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c_K)) - \sum_{k=1}^{K} \frac{n_{c_k}}{n_X}\mathcal{S}(d(c_k)).$$

PROOF. By Corollary 2, $\mathcal{S}(\uplus_{k=1}^K d(c_k)) = \mathcal{S}(\uplus_{k=1}^K R(d(c_k))) + \sum_{k=1}^K \frac{n_{d(c_k)}}{n_X}\mathcal{S}(d(c_k))$. Also, by Corollary 1, $\{c_k\} \sim R(d(c_k))$ for $k = 1, \ldots, K$. Since these two districts have the same attendance and total income, by IND, $X = \uplus_{k=1}^K \{c_k\} \sim \uplus_{k=1}^K R(d(c_k))$, which implies that $S(X) = S(\uplus_{k=1}^K R(d(c_k)))$. Noting that $n_{c_k} = n_{d(c_k)}$, and $d(c_1 \mathbin{+\!\!+} \ldots, \mathbin{+\!\!+} c_K) = \uplus_{k=1}^K d(c_k)$, rearranging yields the desired result. □

We now start the third part of the proof. We will show that, restricted to the class of simple districts $\mathcal{S}$ has a very particular form. Let I be the inequality index defined by $I(c) = \mathcal{S}(d(c))$. Corollary 3 says that $\mathcal{S}$ is the segregation index induced by I just defined. We now show that I is a monotone transformation of a member of the generalized entropy family defined in Example 1. The proof is based on Theorem 5 in Shorrocks (1984, p. 1381). In order to apply it, we will show that the inequality index $I(c)$ satisfies the following properties.

Anonymity: $I((n_1, y_1), \ldots, (n_G, y_G)) = I((n_{\pi(1)}, y_{\pi(1)}), \ldots, (n_{\pi(G)}, y_{\pi(G)}))$*for all permutations* $\pi : G \to G$.The reason is that $\langle (n_1, y_1), \ldots, (n_G, y_G)\rangle = \langle (n_{\pi(1)}, y_{\pi(1)}), \ldots, (n_{\pi(G)}, y_{\pi(G)})\rangle$.

Normalization: *For any school c, we have that $I(\bar{c}) = 0$.*Indeed, $I(\bar{c}) = \mathcal{S}(d(\bar{c})) = 0$, where the last equality follows from Claim 2.

Replication invariance: *For any c, we have that $I(c) = I(c \mathbin{+\!\!\!+} c)$.*To see this, note that by PH and Claim 1, $\mathcal{S}(d(c)) = \mathcal{S}(2d(c)) = \mathcal{S}(d(c) \uplus d(c)) = \mathcal{S}(d(c \mathbin{+\!\!\!+} c))$.

Homogeneity: *For any $\alpha > 0$ and any c, $I(c * \alpha) = I(c)$.* Indeed, by IH, $\mathcal{S}(d(c * \alpha)) = \mathcal{S}(d(c) * \alpha) = \mathcal{S}(d(c))$.

Aggregativity: *There is a continuous aggregator $A : R \to \mathbb{R}$ for some subset $R \subset \mathbb{R}^6_+$ such that for all schools c, c'*

$$I(c \mathbin{+\!\!\!+} c') = A(I(c), n_c, \mu_c, I(c'), n_{c'}, \mu_{c'}). \tag{6}$$

Furthermore this aggregator is increasing in its first and fourth arguments. Indeed, consider the function $A : R \to \mathbb{R}$ defined by $A(x, n, \mu, y, m, \nu) = \mathcal{S}(X \uplus Y)$ for some districts X, Y such that $\mathcal{S}(X) = x$, $n_X = n$, $\mu_X = \mu$, and $\mathcal{S}(Y) = y$, $n_Y = m$, $\mu_Y = \nu$. This function is well defined. Indeed, if we let Z and W be two districts such that $(\mathcal{S}(W), n_W, \mu_W) = (\mathcal{S}(X), n_X, \mu_X)$ and $(\mathcal{S}(Z), n_Z, \mu_Z) = (\mathcal{S}(Y), n_Y, \mu_Y)$, by IND applied twice, $\mathcal{S}(X \uplus Y) = \mathcal{S}(X \uplus Z) = \mathcal{S}(W \uplus Z)$. To see that the aggregator A is increasing in its first argument note that by IND, $\mathcal{S}(W \uplus Y) > \mathcal{S}(X \uplus Y)$ whenever $\mathcal{S}(W) > \mathcal{S}(X)$ and $(n_W, \mu_W) = (n_X, \mu_X)$. A similar argument shows that A is increasing in its fourth argument.

To see that Equation (6) holds, note that $I(c \mathbin{+\!\!\!+} c') = \mathcal{S}(d(c \mathbin{+\!\!\!+} c')) = \mathcal{S}(d(c) \uplus d(c'))$ and that by definition of the aggregator A

$$\mathcal{S}(d(c) \uplus d(c')) = A(\mathcal{S}(d(c)), n_c, \mu_c, \mathcal{S}(d(c')), n_{c'}, \mu_{c'}) = A(I(c), n_c, \mu_c, I(c'), n_{c'}, \mu_{c'}).$$

The next proposition show that I satisfies the Pigou–Dalton principle of transfers. Namely, if school c is obtained from school c' by means of a progressive transfer, then $I(c) < I(c')$. Formally,

Proposition 2. *For any two schools $c = \langle (n_1, y_1), (n_2, y_2) \rangle$ and $c' = \langle (n_1, y_1 - \Delta/n_1), (n_2, y_2 + \Delta/n_2) \rangle$ such that $0 < y_1 \le y_2$ and $\Delta \in (0, n_1 y_1)$, we have that $I(c') > I(c)$.*

Proof. Let $b_1 = \frac{n_2 \Delta}{(n_1+n_2)\Delta + n_1 n_2 (y_2 - y_1)}$ and $b_2 = \frac{n_1 \Delta}{(n_1+n_2)\Delta + n_1 n_2 (y_2 - y_1)}$, and consider the following subdivision of school $c_1 = \langle (n_1, y_1) \rangle$ into

$$c_{11} = \langle ((1 - b_1) n_1, y_1 - \Delta/n_1) \rangle \text{ and } c_{12} = \big\langle (b_1 n_1, y_2 + \Delta/n_2) \big\rangle$$

Since $n_1 = n_{c_{11}} + n_{c_{12}}$ and $|c_1| = |c_{11}| + |c_{12}|$, this subdivision is feasible. By UDD and Claim 2 we have that $\mathcal{S}(\{c_{11}, c_{12}\}) > \mathcal{S}(\{c_1\})$. Similarly, if we subdivide school $c_2 = \langle (n_2, y_2) \rangle$ into the following two schools:

$$c_{21} = \langle (b_2 n_2, y_1 - \Delta/n_1) \rangle,\ \ c_{22} = \big\langle ((1 - b_2) n_2, y_2 + \Delta/n_2) \big\rangle$$

we obtain that $\mathcal{S}(\{c_{21}, c_{22}\}) > \mathcal{S}(\{c_2\})$. Therefore, by IND

$$\begin{aligned} \mathcal{S}(d(c)) = \mathcal{S}(\{c_1\} \uplus \{c_2\}) &< \mathcal{S}(\{c_{11}, c_{12}\} \uplus \{c_{21}, c_{22}\}) \\ &= \mathcal{S}(\{c_{11}, c_{21}\} \uplus \{c_{12}, c_{22}\}). \end{aligned} \tag{7}$$

By UDD, $\mathcal{S}(\{c_{11}, c_{21}\}) = \mathcal{S}(\{\langle (n_1, y_1 - \Delta/n_1) \rangle\})$ and $\mathcal{S}(\{c_{12}, c_{22}\})) = \mathcal{S}(\{\langle (n_2, y_2 + \Delta/n_2) \rangle\})$. Since $n_{c_{11}} + n_{c_{21}} = n_1$ and $n_{c_{12}} + n_{c_{22}} = n_2$, by IND,

$$\begin{aligned} \mathcal{S}(\{c_{11}, c_{21}\} \uplus \{c_{12}, c_{22}\}) &= \mathcal{S}\big(\big\{\langle (n_1, y_1 - \Delta/n_1) \rangle\big\} \uplus \big\{\langle (n_2, y_2 + \Delta/n_2) \rangle\big\}\big) \\ &= \mathcal{S}(d(c')). \end{aligned} \tag{8}$$

From inequalities (7) and (8), we obtain that $I(c) = \mathcal{S}(d(c)) < \mathcal{S}(d(c')) = I(c')$ which is what we wanted to show. □

Finally, the next proposition uses the fact that I satisfies the Pigou–Dalton principle to show that it is continuous.

PROPOSITION 3. *For all $c = \langle (n_1, y_1), \ldots, (n_G, y_G)\rangle$, the value $I(c)$ depends continuously on its arguments (n_g, y_g).*

PROOF 1. Let $c = \langle (n_1, y_1), \ldots, (n_G, y_G)\rangle$ be a school and let $c^k = \langle (n_1^k, y_1^k), \ldots, (n_G^k, y_G^k)\rangle$, for $k = 1, 2, \ldots$ be a sequence of schools that converges to c. We need to show that $I(c^k) \to I(c)$. We can assume without loss of generality that $|c| = |c^k| = 1$ and $n_c = n_{c^k} = 1$, for $k = 1, 2, \ldots$. Indeed, since $\succcurlyeq$ satisfies IH we can define $\hat{c}$, and $\hat{c}^k$ to be the schools that are obtained from c and c^k, respectively by normalizing both their attendance and income to be one as follows: $\hat{c} = (1/n_c)\, c * (n_c/|c|)$ and $\hat{c}^k = (1/n_{c^k})\, c^k * (n_{c^k}/|c^k|)$. Since $c^k \to c$ we have that $\{\hat{c}^k\} \to \hat{c}$. By PH and by IH, $I(\hat{c}^k) = I(c^k)$ and $I(\hat{c}) = I(c)$ for all k.

Let $n = \min\{n_1, \ldots, n_G\}$ and $y = \min\{y_1, \ldots, y_G\}$. Let $\varepsilon \in (0, \min\{n, y\})$ and let k_0 be such that for all $k > k_0$, $||c^k - c|| < \varepsilon$. Consider school $c^* = \langle (n_p, y_p), (n_r, y_r)\rangle$, where $n_p = 1 - (n - \varepsilon)$, $y_p = y - \varepsilon$, and (n_r, y_r) is chosen so that $n_{c^*} = |c^*| = 1$. School c^* is a school with two income groups, the poor being poorer than every student both in c and in c^k, and the rich being richer than the rich both in c and c^k, for every $k > k_0$. Also, the number of rich in c^* is smaller than the number of members in every income group both in c and c^k.

By construction, the income distribution of c^* is Lorenz-dominated by that of c^k for all $k > k_0$. That is, the Lorenz curve associated with c^* is nowhere above the one associated with c^k. Therefore, there is a sequence of schools $c_0, c_1, \ldots, c_N$ with $c_0 = c^*$ and $c_N = c^k$ such that c_{t+1} is obtained from c_t by means of a progressive transfer. Therefore, by Proposition 2, $I(c^*) > I(c^k)$. As a result, $d(c^*) \succ d(c^k)$ for all $k > k_0$. Furthermore, since c^k converges to c, by CONT $d(c^*) \succcurlyeq d(c)$.

Let now $X' \in \mathcal{D}_1$ such that $X' \succcurlyeq X_{1/2}$ and $X' \succcurlyeq d(c^*)$. By Lemma 2 there are unique $\alpha, \alpha^k, \beta \in [0, 1]$ such that

$$d(c) \sim \alpha X' \uplus (1-\alpha) X_0, \tag{9}$$

$$d(c^k) \sim \alpha^k X' \uplus (1-\alpha^k) X_0 \qquad k > k_0, \tag{10}$$

$$X_{1/2} \sim \beta X' \uplus (1-\beta) X_0. \tag{11}$$

We end the proof by showing that α^k converges to α. Since $I(c^k) = \alpha^k/\beta$ and $I(c) = \alpha/\beta$ this will imply that $I(c^k)$ converges to $I(c)$, which is what we want to show. The argument is standard. Since $\alpha^k \in [0, 1]$ for $k > k_0$, the sequence $\{\alpha^k\}_{k>k_0}$ has a convergent subsequence. We now argue that all its convergent subsequences converge to α. Assume by contradiction that there is a subsequence $\alpha^{k(\ell)}$ that converges to $\lambda > \alpha$ (the case where $\lambda < \alpha$ is similar and left to the reader). Let $\widehat{\lambda} = \frac{\lambda+\alpha}{2}$. Since $\widehat{\lambda} > \alpha$ by Lemma 1 and Equation (9)

$$\widehat{\lambda} X' + (1-\widehat{\lambda}) X_0 \succ \alpha X' + (1-\alpha) X_0 \sim d(c). \tag{12}$$

Since $\alpha^{k(\ell)}$ converges to $\lambda > \widehat{\lambda}$, there is an ℓ_0 such that for all $\ell > \ell_0$, $\alpha^{k(\ell)} > \widehat{\lambda}$. Therefore, by equation (10) and Lemma 1, for all $\ell > \ell_0$,

$$d(c^{k(\ell)}) \sim \alpha^{k(\ell)} X' \uplus (1 - \alpha^{k(\ell)}) X_0 \succ \widehat{\lambda} X' \uplus (1 - \widehat{\lambda}) X_0.$$

Since $c^{k(\ell)}$ converges to c, by CONT we have that $d(c) \succcurlyeq \widehat{\lambda} X' \uplus (1 - \widehat{\lambda}) X_0$, which contradicts (12). □

Since I satisfies the above properties on $\mathcal{C}$, they are also satisfied on the subclass of schools $\mathcal{C}_{\mathbb{Z}}$, where the population n_g of each of its groups is an integer. It now follows from Theorem 5 in Shorrocks (1984, p. 1381) that there exists a parameter α in $\mathbb{R}$ and an increasing, continuous function $F : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $F(0) = 0$ such that for any school c in $\mathcal{C}_{\mathbb{Z}}$,

$$I(c) = F[I^\alpha(c)], \tag{13}$$

where I^α is the generalized entropy inequality index with parameter α.

We now show that Equation (13) also holds for all schools c where the number of pupils in each school is a rational number. To see this, note that when the number of pupils in each group of school c is rational, $kc \in \mathcal{C}_{\mathbb{Z}}$ for some positive integer k. By replication invariance, $I(c) = I(\underbrace{c \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c}_{k})$, which, by combining all the groups with the same income into one group, can be written as $I(kc)$. Then, using Equation (13) we have that $I(c) = I(kc) = F[I^\alpha(kc)] = F[I^\alpha(c)]$, where the last equality follows from the fact that I^α also satisfies replication invariance. Finally, Equation (13) also holds for all schools $c \in \mathcal{C}$ since $F \circ I^\alpha$ is continuous and $\mathbb{Q}$ is dense in $\mathbb{R}$.

We now start the last step of the proof. We show that $\mathcal{S}$ is a positive multiple of the $\mathcal{SSI}$. Given that Equation (13) holds for all schools $c \in \mathcal{C}$, applying Corollary 3, we obtain that the segregation index is of the form

$$\mathcal{S}(X) = F[I^\alpha(X)] - \sum_{c \in X} \frac{n_c}{n_X} F(I^\alpha(c)), \tag{14}$$

that is, $\mathcal{S}$ is the segregation index induced by $F(I^\alpha)$. Since the $\mathcal{SSI}$ is the segregation index induced by I^0, it is enough to show that F is linear and that $\alpha = 0$.

In order to show that F is linear we will make use of the following well-known decomposability property of the generalized entropy indices I^α. See, for instance, equation (20) in Shorrocks (1988).

OBSERVATION 1. *For any two schools c_1 and c_2, let $c = c_1 \mathbin{+\!\!+} c_2$. Then*

$$I^\alpha(c) = \frac{n_{c_1}}{n_c}\left(\frac{\mu_{c_1}}{\mu_c}\right)^\alpha I^\alpha(c_1) + \frac{n_{c_2}}{n_c}\left(\frac{\mu_{c_2}}{\mu_c}\right)^\alpha I^\alpha(c_2) + I^\alpha(\overline{c}_1 \mathbin{+\!\!+} \overline{c}_2).$$

The proof of this observation follows from a routine manipulation of the formula of I^α and is left to the reader.

We now show that F must be both concave and convex. Let z, z' be in the range of I^α (which is known to be an interval), and $\gamma \in (0, 1)$. Assume without loss of generality that $z > z'$. Pick two simple districts, $X = \{c_1, \ldots, c_K\}$ and $Y = \{c'_1, \ldots, c'_{K'}\}$, each with unit population and unit income, such that $I^\alpha(X) = z$ and $I^\alpha(Y) = z'$. Since X and Y are simple districts, we have that $I^\alpha(c) = 0$ for all $c \in X$ and for all $c' \in Y$. Therefore, since F is increasing, $\mathcal{S}(X) = F(I^\alpha(X)) = F(z) > F(z') = F(I^\alpha(Y)) = \mathcal{S}(Y)$. Since $\mathcal{S}(X) > 0$, X has at least two schools. Pick one school, say c_1, and transfer a proportion p of pupils from each of the other schools to school c_1 to obtain district $X(p) = \{c_1 \mathbin{+\!\!+} p(c_2 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c_n), (1-p)c_2, \cdots (1-p)c_n\}$. Denoting $c_1(p) = c_1 \mathbin{+\!\!+} p(c_2 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} c_n)$ we have by Equation (14)

$$\begin{aligned}\mathcal{S}(X(p)) &= F(I^\alpha(X(p))) - (n_1 + p(1-n_1)F(I^\alpha(c_1(p))) \\ &= F(I^\alpha(X)) - (n_1 + p(1-n_1)F(I^\alpha(c_1(p))).\end{aligned}$$

Note that when $p = 0$, $\mathcal{S}(X(0)) = \mathcal{S}(X) = F(z) > F(z') = \mathcal{S}(Y)$, and when $p = 1$, $\mathcal{S}(X(1)) = F(I^\alpha(X)) - F(I^\alpha(X)) = 0$. Consequently, by the intermediate value theorem, there is a $p^* \in (0, 1)$ such that $\mathcal{S}(X(p^*)) = \mathcal{S}(Y)$. Let $Z = X(p^*)$ and note that $n_Z = n_X = 1 = n_Y$, $|Z| = |X| = |Y| = 1$ and $I^\alpha(Z) = I^\alpha(X) = z$. Then, given that Y is a simple district,

$$\begin{aligned}\mathcal{S}(\gamma Z \uplus (1-\gamma)Y) &= F(I^\alpha(\gamma Z \uplus (1-\gamma)Y) - \sum_{c\in Z} \gamma n_c F(I^\alpha(\gamma c)) \\ &= F(\gamma I^\alpha(Z) + (1-\gamma)I^\alpha(Y)) - \sum_{c\in Z} \gamma n_c F(I^\alpha(\gamma c)),\end{aligned} \tag{15}$$

where the second equality made use of Observation 1 and the fact that $\mu_{\gamma Z} = \mu_{(1-\gamma)Y}$. On the other hand, since by PH, $\mathcal{S}(\gamma Z) = \mathcal{S}(\gamma Y)$, by IND,

$$\begin{aligned}\mathcal{S}(\gamma Z \uplus (1-\gamma)Y) &= \mathcal{S}(\gamma Y \uplus (1-\gamma)Y) \\ &= \mathcal{S}(Y) \\ &= \gamma\mathcal{S}(Y) + (1-\gamma)\mathcal{S}(Y) \\ &= \gamma\mathcal{S}(Z) + (1-\gamma)\mathcal{S}(Y),\end{aligned}$$

where the second equality follows from Claim 1 and the fourth one from IND. Using Equation (14), and taking into account that Y is a simple district,

$$\begin{aligned}\mathcal{S}(\gamma Z \uplus (1-\gamma)Y) &= \gamma\left[F(I^\alpha(Z)) - \sum_{c\in Z} n_c F(I^\alpha(c))\right] + (1-\gamma)F(I^\alpha(Y)) \\ &= \gamma F(I^\alpha(Z)) + (1-\gamma)F(I^\alpha(Y)) - \sum_{c\in Z} \gamma n_c F(I^\alpha(c)).\end{aligned} \tag{16}$$

Comparing Equations (15) and (16), and taking into account that $I^\alpha(\gamma c) = I^\alpha(c)$, we conclude that $\gamma F(I^\alpha(Z)) + (1-\gamma)F(I^\alpha(Y)) = F(\gamma I^\alpha(Z) + (1-\gamma)I^\alpha(Y))$. Recalling that $I^\alpha(Z) = z$ and $I^\alpha(Y) = z'$ we conclude that F is both concave and convex. Furthermore, since $F(0) = 0$, we have that $F(z) = az$ for some $a > 0$.

It remains to show that $\alpha = 0$. We will show that unless this is the case, there exist two schools, c_1 and c_2 such that $\mathcal{S}(\{c_1, c_2\}) < 0$, which contradicts Claim 2.

Let $\alpha \neq 0$. Let $n_1 = n_2 = 1$, let $\mu_1 > 0$ be such that $\mu_1^\alpha \in (0, 1)$, and let μ_2 be implicitly defined by $n_1\mu_1 + n_2\mu_2 = 1$. Also let $c_1 = \langle(p, \varepsilon\mu_1), ((1-p), \frac{\mu_1(1-\varepsilon p)}{(1-p)})\rangle$ and $c_2 = \langle(1, \mu_2)\rangle$ be two schools where $0 < p < 1$ and $0 < \varepsilon < 1$. School c_1 has two income groups. The proportion of pupils in the lower income group is p. The total population is 1 and the mean income is μ_1. It can be checked that the closer p is to 1 and ε to 0, the higher is the income inequality as measured by I^α, both because the proportion of low-income pupils becomes large and their incomes become low. For the moment assume that p is chosen to be close enough to 1 and ε is chosen to be close enough to 0 so that

$$I^{\alpha}(c_1) > \frac{I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2)}{1/2\,(1-\mu_1^{\alpha})}. \tag{17}$$

We will later show that this can be done. Now let $X = \{c_1, c_2\}$. Then, using Equation (14) and the fact that $F(z) = az$, we have that

$$\mathcal{S}(\{c_1, c_2\}) = a\left(I^{\alpha}(X) - \sum_{s=1}^{2} \frac{I^{\alpha}(c_s)}{2}\right).$$

By Observation 1 and since $I^{\alpha}(c_2) = 0$,

$$\begin{aligned}
\mathcal{S}(\{c_1, c_2\}) &= a\left(I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2) + \frac{1}{2}\mu_1^{\alpha} I^{\alpha}(c_1) + \frac{1}{2}\mu_2^{\alpha} I^{\alpha}(c_2) - \frac{1}{2} I^{\alpha}(c_1) - \frac{1}{2} I^{\alpha}(c_2)\right) \\
&= a\left(I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2) + \frac{1}{2}\mu_1^{\alpha} I^{\alpha}(c_1) - \frac{1}{2} I^{\alpha}(c_1)\right) \\
&= a\left(I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2) - \frac{1}{2}(1-\mu_1^{\alpha}) I^{\alpha}(c_1)\right) \\
&< 0,
\end{aligned}$$

where the last inequality follows from inequality (17). As mentioned before, this inequality contradicts Claim 2.

It remains to show that $p < 1$ and $\varepsilon > 0$ can be chosen so that inequality (17) holds. To see this, note first that since $n_{\bar{c}_1} = 1$ and $\mu_{\bar{c}_1} = \mu_1$, we have that $I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2)$ is independent of p and of ε. Also, by direct computation, we have that

$$I^{\alpha}(c_1) = \begin{cases} \frac{p\varepsilon^{\alpha} + (1-p)^{1-\alpha}(1-\varepsilon p)^{\alpha} - 1}{(\alpha-1)\alpha} & \text{if } \alpha \neq 1 \\ p\varepsilon \log(\varepsilon) + (1-p\varepsilon)\log\left(\frac{1-p\varepsilon}{1-p}\right) & \text{if } \alpha = 1. \end{cases}$$

Case 1: $\alpha \geq 1$. In this case we have that $\lim_{p\to 1} I^{\alpha}(c_1) = \infty$ and therefore, inequality (17) can be satisfied.

Case 2: $\alpha < 0$. In this case we have that $\lim_{p\to 1} I^{\alpha}(c_1) = \frac{\varepsilon^{\alpha}-1}{(\alpha-1)\alpha}$ and therefore for ε close enough to 0, inequality (17) holds.

Case 3: $\alpha \in (0, 1)$. In this case we have that $\lim_{\substack{p\to 1 \\ \varepsilon\to 0}} I^{\alpha}(c_1) = \frac{1}{(1-\alpha)\alpha}$, and noting that $\frac{(1-\mu_1^{\alpha})+(1-\mu_2^{\alpha})}{(1-\mu_1^{\alpha})} < 1$ we have that

$$\frac{1}{(1-\alpha)\alpha} > \frac{\left(1-\mu_1^{\alpha}\right)+\left(1-\mu_2^{\alpha}\right)}{\left(1-\mu_1^{\alpha}\right)} \frac{1}{(1-\alpha)\alpha} = \frac{I^{\alpha}(\bar{c}_1 \mathbin{+\!\!\!+} \bar{c}_2)}{1/2\,\left(1-\mu_1^{\alpha}\right)}.$$

Therefore, for p close enough to 1 and ε close enough to 0, inequality (17) holds. This completes the proof of the theorem.

$\square$

TABLE 1
INDEPENDENCE OF THE AXIOMS

	PH	IH	EAP	SSD	UDD	IND	SEP	CONT
$\mathcal{SSI}$	√	√	√	√	√	√	√	√
$n\mathcal{SSI}$		√	√	√	√	√	√	√
$\mathcal{V}$	√		√	√	√	√	√	√
$-\mathcal{SSI}$	√	√		√	√	√	√	√
$\mathcal{W}$	√	√	√		√	√	√	√
$\mathcal{N}$	√	√	√	√		√	√	√
$\mathcal{F}$	√	√	√	√	√		√	√
$\mathcal{C}_{\mathcal{V}}$	√	√	√	√	√	√		√

5.1. *Discussion.*

(a) *Independence of the axioms*. Table 1 lists a number of segregation indices that satisfy all the axioms but one. Apart from the segregation indices introduced in Section 3, the table includes the following indices:
- $n\mathcal{SSI}(X) = n_X \mathcal{SSI}(X)$
- $\mathcal{W}(X) = -\sum_{c\in X} \frac{n_c}{n_X} I^0(c)$
- $\mathcal{N}(X) = 1 - \sum_{c\in X} (\frac{n_c}{n_X})^2$
- $\mathcal{F}(X) = \sqrt{I^0(X)} - \sum\limits_{c\in X} \frac{n_c}{n_X}\sqrt{I^0(c)}$
- $\mathcal{C}_{\mathcal{V}}(X) = \frac{\sqrt{\mathcal{V}(X)}}{\mu_X}$

It can be seen that $\mathcal{SSI}$ satisfies all the axioms introduced in Section 4 and that if we do not require either PH, IH, the single-school property, equivalence of single-school districts, independence, or separability, then a segregation index can be found that satisfies all the remaining axioms. We have not been able to show that continuity is not implied by the other axioms. Our main results states that $\mathcal{SSI}$ index is essentially the only segregation index that satisfies all of them.

(b) *Strength of the axioms*. Our axioms impose restrictions on segregation indices. Nevertheless, restricted to the class of simple districts (those with no income variation within schools), any segregation index $\mathcal{S}$ naturally induces an income inequality index. Theorem 1 implies that our axioms characterize a segregation index whose induced index of income inequality is Theil's second measure. One may wonder whether our axioms restricted to the class of simple districts are strong enough to characterize directly Theil's second measure or any other income inequality index. The answer is negative. Indeed, EAP, SSD, and SEP are axioms that are toothless when applied to indices defined on the class of simple districts since they deal with comparisons between districts that are not in that class. The axioms that have any bite on the subclass of simple districts are UDD, PH, IH, IND, and CONT, which are not sufficient to imply the well-known additive separability of I_0, and not even the Pigou–Dalton principle.

(c) *Our characterization is ordinal*. Specifically, we characterize the segregation order represented by the $\mathcal{SSI}$ instead of the index itself. Therefore, any increasing transformation of the $\mathcal{SSI}$ yields an ordinally equivalent index. In particular, if one is interested in an index that satisfies all our axioms and that is also bounded between 0 and 1, one can apply the transformation $f(x) = x/(1+x)$.

(d) *Additive separability of the index*. Given a partition of a district into two subdistricts, the within-district segregation is the population-weighted average of the segregation of the districts. The between-district segregation, on the other hand, is the segregation that would result if the segregation within each of the subdistricts were to be eliminated. Corollary 2 shows that among the equivalent representations of the order characterized in Theorem 1 the positive multiples of $\mathcal{SSI}$ satisfy a very useful additive separability

TABLE 2
SEGREGATION IN SELECTED CHILEAN REGIONS FOR 2013

	$\mathcal{SSI}$ Segregation Breakdown					$\mathcal{V}$ Segregation Breakdown				
Region	Total 1	Between 2	Within 3	Inequal. 4	Pure 5	Total 6	Between 7	Within 8	Inequal. 9	Pure 10
Biobío	0.259	0.023	0.236	0.441	0.587	0.177	0.007	0.170	0.278	0.635
Valparaíso	0.236	0.016	0.220	0.402	0.586	0.211	0.007	0.204	0.332	0.637
Santiago	0.390	0.024	0.365	0.560	0.696	0.736	0.022	0.714	0.971	0.758

NOTES: Columns 2 and 3 show its decomposition into segregation between- and within-provinces for the $\mathcal{SSI}$. Columns 7 and 8 show the same decomposition for $\mathcal{V}$. Columns 4 and 5 show the decomposition into income inequality and pure segregation induced by $\mathcal{SSI}$. Columns 9 and 10 show this decomposition for $\mathcal{V}$. For the calculation of the variance we measured income in millions of 2013 Chilean pesos.

property. Namely, the index is the sum of the between-district and within-district segregation. It can be checked that this property is also satisfied by $\mathcal{V}$. The next section illustrates this separability property using data from Chile.

6. AN EMPIRICAL ILLUSTRATION

In this section, we illustrate the decomposability property of the $\mathcal{SSI}$ and $\mathcal{V}$ mentioned above. We use data from SIMCE (Sistema de medición de la calidad de la educación), which contains student data from virtually all schools in Chile. Chile has 54 provinces, grouped into 15 regions. For our analysis, we restrict attention to all provinces of the regions of Santiago, Valparaíso, and Biobío (except for the province of Isla de Pascua, which has only three schools). These three regions represent a 60% of the Chilean population. Data include for each student, the school he attends and the income bracket his parents belong to. Income levels, which we measure in millions of Chilean pesos are partitioned into 15 income brackets.[8] For each province, we estimate the mean income in each bracket by assuming that income is distributed according to a log-normal distribution, as follows. For an initial guess $(\bar{y}_1, \ldots, \bar{y}_{15})$ of the mean incomes, we fit a log-normal distribution assuming that all households in income bracket i have an income of $\bar{y}_i$, for $i = 1, \ldots, 15$. Then, we calculate the mean incomes of each bracket induced by the estimated distribution, and repeat the process using the estimated mean incomes as a new guess until the process converges. Chilean schools are classified according to their degree of dependence on public funding into three categories: public, semipublic, and private.[9]

Table 2 shows for the regions of Santiago, Biobío, and Valparaíso, their income segregation as measured both by the $\mathcal{SSI}$ and $\mathcal{V}$ (columns 1 and 6), and its decomposition into segregation between provinces (columns 2 and 7) and segregation within them (column 3 and 8). As can be seen, the Metropolitan region of Santiago exhibits more segregation than the other two, both according to the $\mathcal{SSI}$ and $\mathcal{V}$. However, these two indices do not order the regions identically. Also, for all the three regions, more than 90% of the segregation can be attributed to the segregation within provinces, reflecting the fact that for each region the mean incomes of its provinces are roughly the same. Recall that any segregation index that is induced by an inequality index can be factored into a pure segregation and an inequality indices. Columns

[8] One Chilean peso was equivalent to around US $500 in 2013.
[9] Public schools are funded by the city, and the semipublic category consist of private schools that are subsidized by public funds.

TABLE 3
SEGREGATION IN SELECTED CHILEAN PROVINCES FOR 2013

	$\mathcal{SSI}$	Segregation Breakdown				$\mathcal{V}$	Segregation Breakdown			
Province	Total 1	Between 2	Within 3	Inequal. 4	Pure 5	Total 6	Between 7	Within 8	Inequal. 9	Pure 10
Arauco	0.123	0.054	0.070	0.310	0.398	0.032	0.021	0.010	0.082	0.386
Biobío	0.223	0.143	0.080	0.411	0.542	0.136	0.102	0.034	0.226	0.604
Concepcion	0.269	0.189	0.080	0.443	0.607	0.253	0.194	0.059	0.380	0.666
Ñuble	0.233	0.123	0.110	0.424	0.551	0.100	0.043	0.056	0.186	0.537
Chacabuco	0.617	0.555	0.062	0.805	0.766	1.812	1.769	0.043	2.086	0.869
Cordillera	0.147	0.066	0.081	0.298	0.493	0.106	0.057	0.048	0.249	0.425
Maipo	0.275	0.190	0.085	0.441	0.623	0.287	0.228	0.059	0.420	0.683
Melipilla	0.207	0.171	0.036	0.380	0.544	0.147	0.122	0.025	0.236	0.625
Santiago	0.402	0.300	0.101	0.573	0.701	0.828	0.711	0.118	1.096	0.756
Talagante	0.248	0.160	0.088	0.419	0.593	0.247	0.185	0.062	0.378	0.652
Los Andes	0.213	0.160	0.052	0.393	0.542	0.190	0.149	0.041	0.353	0.539
Marga Marga	0.170	0.104	0.065	0.351	0.484	0.121	0.086	0.035	0.262	0.463
Petorca	0.074	0.038	0.037	0.256	0.290	0.021	0.009	0.012	0.095	0.217
Quillota	0.189	0.139	0.050	0.361	0.523	0.136	0.104	0.032	0.245	0.555
S. Antonio	0.113	0.048	0.065	0.285	0.396	0.050	0.024	0.026	0.141	0.351
San Felipe	0.180	0.137	0.042	0.346	0.520	0.130	0.101	0.029	0.217	0.598
Valparaiso	0.311	0.252	0.059	0.462	0.673	0.350	0.294	0.056	0.479	0.731

NOTES: Columns 2 and 3 show its decomposition into segregation between- and within-school categories for the $\mathcal{SSI}$. Columns 7 and 8 show the same decomposition for $\mathcal{V}$. Columns 4 and 5 show the decomposition into income inequality and pure segregation induced by $\mathcal{SSI}$. Columns 9 and 10 show this decomposition for $\mathcal{V}$. For the calculation of the variance we measured income in millions of 2013 Chilean pesos.

4 and 5 report the result of this factorization for each of the regions for the case of the $\mathcal{SSI}$, and columns 9 and 10 report it for $\mathcal{V}$.[10] As can be seen, the tiny difference in the segregation exhibited by the regions of Biobío and Valparaíso is mainly due to a difference in their income inequality instead of a difference in their pure segregation.

The fact that most of the regions segregation is attributed to the segregation within their provinces suggests an analysis of this component. Table 3 reports for each of the provinces of the above three regions, their income segregation in 2013 as measured both by the $\mathcal{SSI}$ and $\mathcal{V}$, and its decomposition into between- and within-school categories.[11]

Though similar, the ordering of the provinces according to the two indices are not identical. We can see that in most provinces, a large proportion of income segregation both according to $\mathcal{SSI}$ and $\mathcal{V}$, is due to the segregation between categories. This indicates that the mean incomes of the public, semipublic, and private schools are substantially different from each other. The mean incomes of the schools within each category, on the other hand, are similar to each other as evidenced by the small segregation within categories exhibited by most provinces.

[10] Since the pure segregation associated with $\mathcal{V}$ is the square of Jargowsky's $\mathcal{NSI}$, column 10 corresponds to this index.

[11] As mentioned above, schools are classified into public, semipublic, and private.

APPENDIX

PROOF OF COROLLARY 2. Applying Proposition 1 to the district $X_1 \uplus R(X_2)$, we obtain the statement for $J = 2$. Therefore, for any $m > 1$,

$$\mathcal{S}\left(\biguplus_{j=1}^{m} X_j\right) = \mathcal{S}\left(R\left(\biguplus_{j=1}^{m-1} X_j\right) \uplus R(X_m)\right) + \frac{\sum_{j=1}^{m-1} n_{X_j}}{n}\mathcal{S}\left(\biguplus_{j=1}^{m-1} X_j\right) + \frac{n_{X_m}}{n}\mathcal{S}(X_m).$$

Assume that the statement is also true for $J = m - 1$. Then, denoting $n = n_X$,

$$\mathcal{S}\left(\biguplus_{j=1}^{m} X_j\right) = \mathcal{S}\left(R\left(\biguplus_{j=1}^{m-1} X_j\right) \uplus R(X_m)\right) + \frac{\sum_{j=1}^{m-1} n_{X_j}}{n}\left[\mathcal{S}\left(\biguplus_{j=1}^{m-1} R(X_j)\right) + \sum_{j=1}^{m-1} \frac{n_{X_j}}{\sum_{j=1}^{m-1} n_{X_j}}\mathcal{S}(X_j)\right] +$$

$$+\frac{n_{X_m}}{n}\mathcal{S}(X_m)$$

$$= \mathcal{S}\left(R\left(\biguplus_{j=1}^{m-1} X_j\right) \uplus R(X_m)\right) + \frac{\sum_{j=1}^{m-1} n_{X_j}}{n}\mathcal{S}\left(\biguplus_{j=1}^{m-1} R(X_j)\right) + \sum_{j=1}^{m} \frac{n_{X_j}}{n}\mathcal{S}(X_j). \tag{A.1}$$

Applying this expression to $\biguplus_{j=1}^{m} R(X_j)$, and noting that $R(R(X_j)) = R(X_j)$, we obtain that

$$\mathcal{S}\left(\biguplus_{j=1}^{m} R(X_j)\right) = \mathcal{S}\left(R\left(\biguplus_{j=1}^{m-1} R(X_j)\right) \uplus R(X_m)\right) + \frac{\sum_{j=1}^{m-1} n_{X_j}}{n}\mathcal{S}\left(\biguplus_{j=1}^{m-1} R(X_j)\right) + \sum_{j=1}^{m} \frac{n_{X_j}}{n}\mathcal{S}(R(X_j)).$$

Since $R(\biguplus_{j=1}^{m-1} R(X_j)) = R(\biguplus_{j=1}^{m-1} X_j)$ and since by Claim 2 $\mathcal{S}(R(X_j)) = 0$, rearranging we obtain $\mathcal{S}(R(\biguplus_{j=1}^{m-1} X_j) \uplus R(X_m)) = \mathcal{S}(\biguplus_{j=1}^{m} R(X_j)) - \frac{\sum_{j=1}^{m-1} n_{X_j}}{n}\mathcal{S}(\biguplus_{j=1}^{m-1} R(X_j))$. Replacing this expression in Equation (A.1) we get

$$\mathcal{S}\left(\biguplus_{j=1}^{m} X_j\right) = \mathcal{S}\left(\biguplus_{j=1}^{m} R(X_j)\right) + \sum_{j=1}^{m} \frac{n_{X_j}}{n}\mathcal{S}(X_j).$$

□

REFERENCES

COLEMAN, J. S., E. CAMPBELL, C. HOBSON, J. MCPARTLAND, A. MOOD, F. WEINFELD, AND R. YORK, *Equality of Educational Opportunity Study* (Washington, DC: United States Department of Health, Education, and Welfare, 1966).

DAWKINS, C. J., "Space and the Measurement of Income Segregation," *Journal of Regional Science* 47 (2007), 255–72.

ECHENIQUE, F., AND R. G. FRYER, "A Measure of Segregation Based on Social Interactions," *The Quarterly Journal of Economics* 122 (2007), 441–85.

FONG, E., AND K. SHIBUYA, "The Spatial Separation of the Poor in Canadian Cities," *Demography* 37 (2000), 449–59.

FOSTER, J. E., AND A. F. SHORROCKS, "Subgroup Consistent Poverty Indices," *Econometrica* 59 (1991), 687–709.

FRANKEL, D. M., AND O. VOLIJ, "Measuring School Segregation," *Journal of Economic Theory* 146 (2011), 1–38.

HANUSHEK, E. A., J. F. KAIN, J. M. MARKMAN, AND S. G. RIVKIN, "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (2003), 527–44.

HUTCHENS, R., "Numerical Measures of Segregation: Desirable Properties and Their Implications," *Mathematical Social Sciences* 42 (2001), 13–29.

———, "One Measure of Segregation," *International Economic Review* 45 (2004), 555–78.

IMBERMAN, S. A., A. D. KUGLER, AND B. I. SACERDOTE, "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees," *The American Economic Review* 102 (2012), 2048–82.

JAHN, J., C. F. SCHMID, AND C. SCHRAG, "The Measurement of Ecological Segregation," *American Sociological Review* 12 (1947), 293–303.

JARGOWSKY, P. A., "Take the Money and Run: Economic Segregation in US Metropolitan Areas," *American Sociological Review* 61 (1996), 984–98.

KIM, J., AND P. A. JARGOWSKY, "The Gini coefficient and segregation on a continuous variable," in Y. Flückiger, S. F. Reardon, and J. Silber, eds., *Occupational and Residential Segregation* (Research on Economic Inequality, Vol. 17) (Bingley, United Kingdom: Emerald Group Publishing Limited, 2009), 57–70.

LAVY, V., M. D. PASERMAN, AND A. SCHLOSSER, "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom," *The Economic Journal* 122 (2012), 208–37.

MASSEY, D. S., AND N. A. DENTON, "The Dimensions of Residential Segregation," *Social Forces* 67 (1988), 281–315.

MAYER, S. E., "How Economic Segregation Affects Children's Educational Attainment," *Social Forces* 81 (2002), 153–76.

REARDON, S. F., "Measures of Ordinal Segregation," in Y. Flückiger, S. F. Reardon, and J. Silber, eds., *Occupational and Residential Segregation* (Research on Economic Inequality, Vol. 17) (Bingley, United Kingdom: Emerald Group Publishing Limited, 2009), 129–55.

———, "Measures of Income Segregation," Technical Report, Stanford Center for Education Policy Analysis, 2011.

———, AND G. FIREBAUGH, "Measures of Multigroup Segregation," *Sociological Methodology* 32 (2002), 33–67.

SHORROCKS, A. F., "The Class of Additively Decomposable Inequality Measures," *Econometrica* 48 (1980), 613–25.

———, "Inequality Decomposition by Population Subgroups," *Econometrica* 52 (1984), 1369–85.

———, "Aggregation Issues in Inequality Measurement," In W. Eichhorn (Ed.), *Measurement in Economics* (Heidelberg: Springer, 1988), 429–51.

THEIL, H., *Economics and Information Theory* (Amsterdam: North-Holland Publishing Company, 1967).