Theory and Decision
https://doi.org/10.1007/s11238-026-10142-7

Characterization of the multi-group Gini Segregation index

Francesco Andreoli[1,2] · **Claudio Zoli**[1]

Received: 27 December 2025 / Accepted: 17 April 2026

Abstract

Segregation analysis is concerned with the way in which mutually exclusive groups are dissimilarly distributed across non-overlapping organizational units. We axiomatically derive a novel multi-group measure of segregation, the Gini Segregation index, which values segregation as a form of dissimilarity between groups' distributions. The index coincides with the volume spanned by the zonotope representation of the data (a generalization of the segregation curve to more than two groups) and is hence consistent with robust segregation partial orders discussed in the literature. The axiomatic model that we consider could also be useful to characterize volume-based measures of multivariate inequality and mobility.

Keywords Dissimilarity · Gini · volume · zonotope · majorization

1 Introduction

The standard approach to segregation measurement requires partitioning the population into mutually exclusive groups, each gathering individuals with the same traits such as ethnicity, gender, income or parental background, and assessing the way these groups are dissimilarly distributed across *organizational units*, such as neighborhoods (Cutler & Glaeser, 1997; Reardon & O'Sullivan, 2004), schools (Echenique et al.,

An earlier version of this paper has benefitted from the comments and suggestions by Valentino Dardanoni, Koen Decancq, Brice Magdalou, Eugenio Peluso, Alain Trannoy, Dirk Van de gaer and participants to the Social Choice and Welfare Conference, the Canazei Winter School on Inequality and Social Welfare Theory and the Public Economic Theory Conference. We are also grateful to a Reviewer for helpful comments and suggestions.

✉ Francesco Andreoli
francesco.andreoli@univr.it

Claudio Zoli
claudio.zoli@univr.it

1 Department of Economics, University of Verona, Via Cantarane 24, 37129 Verona, Italy

2 Luxembourg Institute of Socio-Economic Research (LISER), 11 Porte des Sciences, L4366 Esch-sur-Alzette, Luxembourg

Published online: 21 May 2026

2006; Frankel & Volij, 2011) or jobs (Flückiger & Silber, 1999; Hutchens, 1991, 2004). Each distribution depicts the likelihood that a group is observed in any of a finite number of organizational units. A collection of distributions, one for each group defined on the same set of units, can be arranged into a matrix, which represents a configuration of the data. Configurations correspond, for instance, to distributions of groups across neighborhoods, school districts or firms.

In empirical analysis of segregation, groups distributions can be estimated from the observed relative frequency at which each group is represented across organizational units. The extent to which groups are disproportionately represented within each unit is informative about the degree of segregation that a configuration displays. Disproportionality between groups is zero when relative frequencies observed in each unit coincide across groups in a given unit (but not necessarily across units). This is the no-segregation case. Conversely, disproportionality reaches a maximum when one and only one group has a non-zero probability of being observed in any given unit. In this case, segregation is also maximal. All cases in-between display some segregation and should be ranked accordingly.

The literature has extensively investigated criteria for ranking configurations when only two groups/distributions are involved, such as the partial order of configurations induced by non-intersecting segregation curves (Alonso-Villar & del Rio, 2010; Duncan & Duncan, 1955; Lasso de la Vega & Volij, 2014) as well as through indices (Flückiger & Silber, 1999; Massey & Denton, 1988; Puerta & Urrutia, 2016). The axiomatic properties that are used in the two-groups cases are heavily inspired by the literature on income inequality measurement. Some key axioms, such as segregation-reducing transfers, cannot however be straightforwardly extended to comparisons of segregation involving *more than two groups*. Moreover, multi-group segregation evaluations may not be simply summarized by sequences of two-groups evaluations (either looking at all possible combinations of all pairs of groups or all comparisons of each group compared with an aggregate of all other groups).

In their important contribution, Frankel and Volij (2011) have investigated an axiomatic model that draws on *informativeness* to characterize broad families of multi-group segregation indices (such as those in Reardon and Firebaugh 2002). The convex closure of the linear orders that they characterize admits a representation in terms of *matrix majorization* (see also Dahl 1999), and they provide a characterization of it. In terms of segregation, matrix majorization invokes the use of transformations represented by row-stochastic matrices which, when applied to the data of a distribution of two or more groups across units, allow to obtain another distribution that unambiguously displays less segregation than the original one. Matrix majorization is a demanding criterion, in the sense that it originates a very partial order of configurations, and it is also difficult to test in practice.

Andreoli and Zoli (2023) have introduced and characterized an alternative robust dissimilarity criterion for ranking distributions of two or more groups that is based on the *inclusion of the zonotope* representations of the data. The zonotopes inclusion criterion is a natural generalization of the segregation curve partial order to the multi-group setting. In the multi-group case, the zonotope inclusion criterion induces a partial ordering of configurations that is weaker (i.e. less partial) than matrix majorization and that is therefore implied by it. The two orders coincide when comparisons involve configurations of only two groups.

In this paper, we characterize axiomatically a multi-group segregation measure, called the *Gini Segregation index*, which corresponds to the volume of the zonotope representation of the data. The index is obviously consistent with the zonotope inclusion criterion and hence with matrix majorization. The main result offers a complete axiomatic derivation of the Gini Segregation index that rides on two fundamental principles. The first principle requires that the order in terms of segregation of two multi-group distributions is preserved when the same row-stochastic transformation is applied to each of the two distributions. This is a form of consistency of segregation indices vis-a-vis the transformations applied to the data. The second principle states that the impact on segregation of a multiplicity of row-stochastic transformations does not depend on the order in which such transformations are applied, but only on the set of transformations that are used.

We also show that similar results could be obtained by employing more elementary transformations than those implied by row-stochastic matrices, which require to merge a certain proportion of the groups observed in one unit with the proportion of groups observed in another unit, thereby increasing the size of the newly obtained unit while leveling (or leaving unchanged) disparities in groups composition therein. We require every segregation index to react to such operations by reducing in value proportionally to the share that is displaced.

In the two-group case, these two approache allow to derive a characterization of Gini-type segregation indices which is alternative to those presented in the literature (see for instance Puerta and Urrutia 2016). To our knowledge, this paper also provides the first characterization of a multidimensional volume-based measure of dispersion. Results can be hence useful to determine characterization of volume-based indices discussed in the multidimensional inequality literature (Koshevoy & Mosler, 1997, 1996), to which this paper relates.

The rest of the paper is organized as follows. The notation is reported in Section 2 whereas the Gini Segregation index and its relations with majorization conditions are described in Section 3. Axioms and the main characterization result (Theorem 1) are in Section 4. Section 4.2 delivers a new characterization based on our axiomatic model for the case with two groups. Section 5 concludes. All proofs are collected in Appendix.

2 Notation

A *distribution matrix* $\mathcal{L}$ collects the probabilities that any group $g \in \mathcal{G} = \{1, \dots, G\}$, $G \geq 2$ is observed in a given unit $i \in \mathcal{N} = \{1, \dots, N\}$, with $N \geq G$. A distribution matrix could represent, for instance, the likelihoods that people located in a given unit, such as a neighborhood or a school or a job, have to interact with the members of each group in a city, a school district or an industry, respectively. We use $\ell_{gi} \in [0, 1]$ with $\sum_{i\in\mathcal{N}} \ell_{gi} = 1$ to denote this likelihood, with $\boldsymbol{\ell}_i$ representing the G dimensional column vector collecting the likelihoods ℓ_{gi} for $g \in \mathcal{G}$. In compact notation:

$$\mathcal{L} := (\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_N) = \begin{pmatrix} \ell_{11} & \ldots & \ell_{1N} \\ \vdots & & \vdots \\ \ell_{G1} & \ldots & \ell_{GN} \end{pmatrix}.$$

Thus, every distribution matrix is row-stochastic, that is, is composed by non-negative elements whose row sum is 1, giving $\mathcal{L}\cdot\mathbf{1}_N = \mathbf{1}_G$ where $\mathbf{1}_G$ is a G dimensional column vector of 1's. The set of all distribution matrices is denoted by $\mathcal{M}_{GN}$.[1] For our analysis we focus on matrices in $\mathcal{M}_{GN}$ that do not include empty units (i.e. columns with all elements equal to 0).[2] Two distribution matrices of interest are I_G, the identity matrix of size G, and Π_G, the permutation matrix of G rows and columns (collected in the set $\mathcal{P}_G$).

3 The Gini Segregation index

We consider the problem of ranking the distribution matrices $\mathcal{L}, \mathcal{L}' \ldots$ with the same number of rows G but varying number of columns $N, N' \ldots$ in terms of segregation exhibited by the distributions of the groups across the units that each distribution matrix displays. A distribution matrix $\mathcal{L} \in \mathcal{M}_{GN}$ displays no segregation whenever all groups exhibit the same (probability) distribution across units i.e., if $\ell_{gi} = \ell_{g'i}$ for any pair $g, g' \in \mathcal{G}$ and for all $i \in \mathcal{N}$. In this case, the knowledge of the unit does not provide information about which group is predominantly occupying it, that is, for any given unit all groups have the same likelihood to be observed in that unit. Segregation is instead maximized whenever the knowledge of the unit allows to forecast the group in it, i.e., if for any $g \in \mathcal{G}, i \in \mathcal{N}$, $\ell_{gi} > 0$ implies $\ell_{g'i} = 0$ for all $g' \neq g$. All distribution matrices that do not display this latter feature, may exhibit some degree of segregation but not maximal segregation. The case of perfect equality among the elements of each vector $\boldsymbol{\ell}_i$ coincides with the case of no segregation, while maximal segregation is associated with maximal inequality among the elements of $\boldsymbol{\ell}_i$ for each unit $i \in \mathcal{N}$.

A *segregation index* is a statistic that maps information provided by a distribution matrix $\mathcal{L} \in \mathcal{M}_{GN}$ into a number, regarded to as the level of segregation displayed

[1] It is common in the segregation literature to use group frequencies to determine ℓ_{gi}, in which case $\ell_{gi} = \frac{n_{gi}}{n_g}$, where n_{gi} is the frequency of group g in unit i and $n_g = \sum_{i\in\mathcal{N}} n_{gi}$. More sophisticated models may estimate likelihoods ℓ_{gi} while taking into considerations features of the spatial distribution of the data (O'Sullivan & Wong, 2007; Wong, 1993) as well as spatial heterogeneity (Reardon & O'Sullivan, 2004), diversity (Clark et al., 2015) and the extent of centrality in spatial networks of interactions (Echenique & Fryer, 2007). The results in this paper abstract from this discussion.

[2] Given that the analysis is not restricted to comparisons with the same number of units, in the empirical case of empty units, these are dropped from the distribution matrix.

by that matrix. Many segregation indices are related to income inequality measures, in that they aggregate the degree of inequality in the distributions $\boldsymbol{\ell}_i$ within each unit $i \in \mathcal{N}$. It is however the case that segregation evaluations should not necessarily focus separately on the distributions $\boldsymbol{\ell}_i$ within each unit i but should also take into account the relations between these distributions by jointly assessing the distributions in all units or subsets of them.

We consider the multi-group *Gini Segregation* index $GS : \bigcup_{N \geq G} \mathcal{M}_{GN} \to [0, 1]$, that is inspired by the multidimensional extension of the Gini index by Koshevoy and Mosler (1997) (see also Koshevoy 1995, Koshevoy and Mosler 1996).

Definition 1 (**The Gini Segregation index**) Let $\mathcal{L} \in \mathcal{M}_{GN}$,

$$GS(\mathcal{L}) := \frac{1}{G!} \sum_{\forall \{i_1,\ldots,i_G\} \subseteq \mathcal{N}} \left| \det \left(\boldsymbol{\ell}_{i_1} \ \ldots \ \boldsymbol{\ell}_{i_G} \right) \right|. \tag{1}$$

According to the Gini Segregation (GS) index the overall segregation is computed by considering building blocks of $G \times G$ matrices collecting vectors $\boldsymbol{\ell}_i$. For a given number of groups G the case of G units identifies the minimal size of the matrices that allows to reach maximal segregation, for instance if the distribution matrix takes the form of the identity matrix I_G. Note that for square matrices $\mathcal{L} \in \mathcal{M}_{GG}$ one obtains $GS(\mathcal{L}) = |\det \mathcal{L}|$. The index measures the segregation within a configuration of G groups distributed over G units by the absolute value of the determinant of the matrix collecting these units and aggregates across all possible combination of G dimensional sets of units $\{i_1, \ldots, i_G\}$ that can be obtained from $\mathcal{N}$.

The GS index can be easily related to robust criteria of multi-group segregation discussed in the literature. Frankel and Volij (2011) have considered axiomatic characterizations of classes of multi-group segregation indices which are based on principles inspired by information theory. They have shown that the convex closure of such class of indices coincides with the partial order of distribution matrices that is obtained by *matrix majorization* (Dahl, 1999). Formally, for any pair $\mathcal{L} \in \mathcal{M}_{GN}$ and $\mathcal{L}' \in \mathcal{M}_{GN'}$, $\mathcal{L}'$ is matrix majorized by $\mathcal{L}$ whenever there exist a $N \times N'$ row-stochastic matrix $X \in \mathcal{M}_{NN'}$ such that $\mathcal{L}' = \mathcal{L} \cdot X$. The GS index is consistent with this partial order criterion.

Andreoli and Zoli (2023) have considered instead a geometric criterion that expands segregation curve dominance to the multi-group domain, which is based on the inclusion comparison of the *zonotope* representation of the data originated by a distribution matrix $\mathcal{L}$. The zonotope set $Z(\mathcal{L}) \subseteq [0, 1]^G$ of a matrix $\mathcal{L} \in \mathcal{M}_{GN}$ is a convex polytope lying on the hypercube $[0, 1]^G$ that is symmetric with respect to the point $\frac{1}{2}\mathbf{1}_G$ and whose coordinates correspond to the Minkowski sum of columns of $\mathcal{L}$ (see McMullen 1971). Formally:

$$Z(\mathcal{L}) := \left\{ \mathbf{z} := (z_1, \ldots, z_G)' : \mathbf{z} = \sum_{i=1}^{N} \theta_i \boldsymbol{\ell}_i,\ \theta_i \in [0, 1]\ \forall i = 1, \ldots, N \right\}.$$

Intuitively the criterion of zonotope inclusion $Z(\mathcal{L}') \subseteq Z(\mathcal{L})$ requires that the distribution of the different groups across the various units in $\mathcal{L}$ is more heterogeneous than in $\mathcal{L}'$ in the sense that there are combinations $\mathbf{z}$ of proportions $\theta_i \boldsymbol{\ell}_i$ that could be attained in $\mathcal{L}$ but not in $\mathcal{L}'$. This is the case because in $\mathcal{L}'$ the distribution of the groups across the units is more homogeneous and therefore the associated zonotope set is "smaller" than in $\mathcal{L}$. For instance, in the case of no segregation where $\ell_{gi} = \ell_{g'i}$ for any $g, g' \in \mathcal{G}$ it is possible to obtain only vectors $\mathbf{z} \in [0, 1]^G$ whose entries are identical, while in the case of maximal segregation the zonotope set coincides with all vectors in $[0, 1]^G$.

Likewise matrix majorization, the zonotope sets inclusion criterion gives rise to a *partial order* of distribution matrices. The zonotope inclusion criterion is, however, less partial (i.e., is capable of ranking more configuration) than matrix majorization (see Koshevoy 1995, Andreoli and Zoli 2023).

The Gini Segregation index is the volume of the zonotope of the distribution matrix $\mathcal{L}$. It is therefore a member of the class of multi-group segregation indices studied in Frankel and Volij (2011). The proof of the following remark can be derived from Andreoli and Zoli (2023).

Remark 1 Let $\mathcal{L} \in \mathcal{M}_{GN}$ and $\mathcal{L}' \in \mathcal{M}_{GN'}$: $\exists X \in \mathcal{M}_{NN'}$ such that $\mathcal{L}' = \mathcal{L} \cdot X \Rightarrow Z(\mathcal{L}') \subseteq Z(\mathcal{L}) \Rightarrow GS(\mathcal{L}') \leq GS(\mathcal{L})$.

The Gini Segregation index has useful decomposition properties discussed in Andreoli (2014). It is also empirically correlated with multi-group dissimilarity indices. Likewise other measures based on determinants, the Gini Segregation index has not been well received in the literature. This is because the index would provide a picture of lack of segregation in all cases in which two rows of a distribution matrix coincide, irrespective of whether differences between the remaining rows are large or small. In fact, as it will be clarified in the next section, the determinant acts as a volume measure of a geometric figure, the zonotope, which is defined on the G-dimensional space. Any linear dependence between (at least two) groups naturally reduces the dimensionality of the zonotope, implying that its multi-dimensional volume and the Gini Segregation index based on it both collapse to zero. Furthermore, the Gini Segregation index cannot be computed when $N < G$, i.e. when some groups' likelihoods are determined as linear combinations of the likelihoods of other groups and the matrix is not of full rank, thus making the determinant measure invalid. As emphasized in a famous paper by Shorrocks (1978), where determinant-based indices of mobility are studied, it could be argued that such hypothetical cases are insufficient grounds for dismissal of the index, insofar they are of little or none relevance for empirical applications where, typically, the number of units N is much larger than G, making the occurrence very unlikely in practice.

The Gini Segregation index when $G = 2$ has been widely studied in the segregation literature (Alonso-Villar & del Rio, 2010; Duncan & Duncan, 1955). The two-group GS index (that we denote by GS_2 in order to make explicit that is computed for two groups) is the area of a two-groups zonotope, that is, it is double the area delimited by the segregation curve and the no-segregation line. Formally:

$$GS_2(\mathcal{L}) := \frac{1}{2} \sum_{\forall \{i,j\} \subseteq \mathcal{N}} \left| \det \begin{pmatrix} \ell_{1i} & \ell_{1j} \\ \ell_{2i} & \ell_{2j} \end{pmatrix} \right|. \tag{2}$$

Puerta and Urrutia (2016) explores this analogy to develop a characterization of the two-group Gini Segregation index. Flückiger and Silber (1999) and Alonso-Villar and del Rio (2010) have considered a multi-group Gini-type index which consists of an average, taken across groups, of the extent of dissimilarity between each group distribution and that of the aggregate population as measured by two-group Gini Segregation index. Gajdos and Weymark (2005) have studied a similar aggregation procedure in the context of multivariate inequality analysis. It has been argued (see for instance Andreoli and Zoli 2020) that such procedure neglects considerations based on the correlations between the attributes (in this case, group belonging) to be evaluated. Moreover, the axiomatic models that are used to provide alternative characterizations of the two-group Gini Segregation index cannot be straightforwardly adapted to the multi-group case. This paper fills the gap.

4 Characterization

4.1 Axioms and general result

For any distribution matrix $\mathcal{L} \in \mathcal{M}_{GN}$, segregation is measured by an index $S^N(\mathcal{L})$, where $S^N : \mathcal{M}_{GN} \to [0, 1]$ denotes a sequence of *continuous functions* from the sets $\mathcal{M}_{GN}$ to the interval $[0, 1]$ for $N \geq G$.[3] Any segregation index S^N should satisfy some desirable properties that drive its behavior when some specific *transfer* operations are applied to the data. Transfers consist in movements of proportions of population of a group from one unit to another. We investigate transfers that, when applied to a distribution matrix generate another distribution matrix that cannot display more segregation than the original one.

For expositional reasons, we initially restrict the focus to the set of square distribution matrices $\mathcal{M}_{GG}$, we will then expand the analysis to matrices in $\mathcal{M}_{GN}$ with $N > G$. Matrices in $\mathcal{M}_{GG}$ are those representing configurations with G groups distributed across the minimal number of units that guarantee that segregation could range between the minimal and the maximal level.[4]

The building block of our analysis is the transfer induced by an *elementary mixture of units*, displacing a portion $1 - \alpha$ with $0 < \alpha \leq 1$ of (the population in) unit i

[3] The index "N" is useful to recall differences in size across matrices under comparison.

[4] In fact, with $N < G$ it won't be possible to associate only one group to any unit.

towards unit j, such that the resulting distribution matrix $\mathcal{L}(\alpha, i, j)$ can be written as

$$\mathcal{L}(\alpha, i, j) := (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \alpha\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G).$$

This transfer is regarded as segregation-reducing. By effect of an elementary mixture of units, the obtained matrix $\mathcal{L}(\alpha, i, j) \in \mathcal{M}_{GG}$ cannot display more segregation than $\mathcal{L}$. In fact, unless $\boldsymbol{\ell}_j$ and $\boldsymbol{\ell}_i$ are linearly dependent, the transfer levels disparities in groups compositions across the units. On the other hand, if $\boldsymbol{\ell}_j$ and $\boldsymbol{\ell}_i$ are linearly dependent, segregation is unaffected by this transformation of the data as it would not alter disproportionality in groups composition in the units.

Every elementary "mixture of units" can be formalized through a specific matrix $X \in \mathcal{M}_{GG}$ which is row-stochastic and that, when post-multiplied to $\mathcal{L}$, gives $\mathcal{L}(\alpha, i, j)$.[5] In general, the transformation embedded by a row-stochastic matrix $X \in \mathcal{M}_{GG}$ applied to the distribution matrix $\mathcal{L}$ generates a new distribution matrix $\mathcal{L}'$ whose distributions of the groups across units are mixtures of those in $\mathcal{L}$. As already noted (see for instance Frankel and Volij 2011, Andreoli and Zoli 2023), matrix majorization can be characterized on the basis of elementary "mixture of units" transformations alongside operations that consist in *evenly or unevenly splitting* a unit into more units. Within the set of matrices $\mathcal{M}_{GG}$, the literature hence regards matrix majorization as a fundamental segregation-reducing transformation in the multi-group context, thus implying that $\mathcal{L}'$ cannot display more segregation than $\mathcal{L}$.

We introduce a parsimonious axiomatic model that, consistently with the literature, also regards matrix majorization as the fundamental criterion for ordering multi-group distribution matrices. The axioms lay down the normative foundations for a segregation index S^G that values the extent of segregation displayed by matrices in $\mathcal{M}_{GG}$. In addition to "natural" normalization properties of the segregation index S^G, we introduce two novel axioms that establish consistency of the ranking of two distributions $\mathcal{L}$ and $\mathcal{L}'$ produced by S^G when both $\mathcal{L}$ and $\mathcal{L}'$ are transformed according to matrix majorization.

Axiom 1 *(UN: Upper Normalization)* $S^G(I_G) = 1$.

Axiom 2 *(LN: Lower Normalization) Let* $\mathcal{L}^0 \in \mathcal{M}_{GG}$ *be such that for any pair* $g, g' \in \mathcal{G}$, $\boldsymbol{\ell}^0_{gi} = \boldsymbol{\ell}^0_{g'i}$ $\forall i \in \mathcal{N}$, *then* $S^G(\mathcal{L}^0) = 0$.

Axiom 3 *(C: Consistency) Let* $\mathcal{L}, \mathcal{L}' \in \mathcal{M}_{GG}$ *then* $S^G(\mathcal{L}) \leq S^G(\mathcal{L}')$ *if and only if for every* $X \in \mathcal{M}_{GG}$, $S^G(\mathcal{L} \cdot X) \leq S^G(\mathcal{L}' \cdot X)$.

Axiom 4 *(IT: Invariance from permutations of Transformations) Let* $\mathcal{L} \in \mathcal{M}_{GG}$, *then for every* $X, Y \in \mathcal{M}_{GG}$, $S^G(\mathcal{L} \cdot X \cdot Y) = S^G(\mathcal{L} \cdot Y \cdot X)$.

[5] For instance see the matrices $T(\alpha_1, 1, 2)$ and $T(\alpha_2, 2, 1)$ in the proof of Theorem 2.

Upper and Lower Normalization axioms establish the cases of maximal and minimal segregation. In particular, the latter axiom regards the case in which rows of a distribution matrix display no dissimilarity as a situation of no-segregation. The Consistency axiom requires the segregation index to be consistent with the application of transformations that support matrix majorization: if a pair of distribution matrices is ranked by a measure S^G, then the same pair of matrices should be ranked in the exact same way by S^G after that both matrices are transformed by the *same* set of operations, here represented by the row-stochastic matrix X. The most demanding axiom is the IT invariance axiom, which requires that the level of segregation measured by S^G could be affected by a sequence of row-stochastic transformations applied to an original matrix, however, the achieved impact on the level of segregation should be independent on the order in which these transformations are applied to the original distribution.

The axioms place restrictions on the functional form of the index S^G and on its properties. We can identify some relevant properties which are desirable in any segregation measure and that can be induced by the application of the four axioms above. As a consequence, all indices of segregation S^G that are consistent with the axioms UN, LN, C and IT described in our model should satisfy properties P1-P6 below. These are the indices implicitly considered in the definitions of the properties. A proof of the properties is in Appendix A.1.

First and foremost, axioms UN and C jointly characterize segregation measures that are consistent with matrix majorization ordering of multi-group segregation in Frankel and Volij (2011). In fact one can obtain the following property.

Property 1 (P1) Let $\mathcal{L} \in \mathcal{M}_{GG}$, then for every $X \in \mathcal{M}_{GG}$, $S^G(\mathcal{L} \cdot X) \leq S^G(\mathcal{L})$.

Such condition can be further verified when the focus is on the impact of elementary segregation-reducing transformations as in Property P2. Recall that any sequence of such transformations originates matrix majorization, but not the reverse.

Property 2 (P2) Let $\mathcal{L} \in \mathcal{M}_{GG}$ and let $\mathcal{L}(\alpha, i, j)$ be obtained from $\mathcal{L}$ by an elementary mixture of units, then $S^G(\mathcal{L}(\alpha, i, j)) \leq S^G(\mathcal{L})$.

Property 3 (P3) Let $X^0 \in \mathcal{M}_{GG}$ be such that for any pair $g, g' \in \mathcal{G}$, $x^0_{gi} = x^0_{g'i}$ $\forall i \in \mathcal{N}$, then for every $\mathcal{L} \in \mathcal{M}_{GG}$, $S^G(\mathcal{L} \cdot X^0) = 0$.

Property P3 is implied by LN and the fact that after the transformation generated by X^0 we obtain a matrix like $\mathcal{L}^0$ no matter what is the original transformed matrix $\mathcal{L}$.[6]

Property 4 (P4) $S^G(\mathcal{L}) = S^G(\mathcal{L} \cdot \Pi_G)$ for all $\mathcal{L} \in \mathcal{M}_{GG}$ and all $\Pi_G \in \mathcal{P}_G$.

[6] Notice, in fact, that any matrix $\mathcal{L}$ post-multiplied by a matrix X^0 as in P3, delivers a new matrix $\mathcal{L}' = \mathcal{L} \cdot X^0$ such that for any pair $g, g', \ell'_{gi} = \ell'_{g'i}$ $\forall i \in \mathcal{N}$.

Property P4 is a unit anonymity property. It is an invariance property of the segregation measures with respect to the label of the units, highlighting that the index does not incorporate concerns related to other dimensions than the heterogeneity in the entries of each single column of a distribution matrix.

Property 5 (P5) $S^G(\mathcal{L}) = S^G(\Pi_G \cdot \mathcal{L})$ for all $\mathcal{L} \in \mathcal{M}_{GG}$ and all $\Pi_G \in \mathcal{P}_G$.

Property P5 is a group anonymity property. The segregation is measured irrespective of the order in which the groups appear in the rows of the distribution matrix. The index does not incorporate concerns related to the name or size of the groups but only on the comparisons of their distributions across units.

The last property makes explicit that any permutation of operations involving matrices in $\mathcal{M}_{GG}$ generates the same level of segregation.

Property 6 (P6) For every $m \geq 2$ and any sequence of m matrices $X_k \in \mathcal{M}_{GG}$ for $k = 1, 2, ..., m$ and their permutation with index $\sigma(k)$ we have that $S^G(X_1 \cdot X_2 \cdot ... \cdot X_k \cdot ... \cdot X_m) = S^G(X_{\sigma(1)} \cdot X_{\sigma(2)} \cdot ... \cdot X_{\sigma(k)} \cdot ... \cdot X_{\sigma(m)})$.

While P1-P5 are widely regarded as relevant properties of segregation indices and satisfied by a large majority of them, P6 is by far more demanding and posits significant structure on the way segregation should be measured. A product of row stochastic matrices is not in general invariant with respect to the permutation of the order of these matrices. What is required is not that any order of transformations generates the same distribution matrix but that all matrices obtained by permuting the order are indifferent in terms of the degree of segregation they exhibit.

The next proposition clarifies that the combined effect of the four axioms leads to the solution that for distributions in $\mathcal{M}_{GG}$ the segregation index should be an increasing transformation of the volume of the zonotope of the distribution matrix $\mathcal{L}$, that is an increasing transformation of the absolute value of the determinant of $\mathcal{L}$.

Proposition 1 *Let $\mathcal{L} \in \mathcal{M}_{GG}$, the segregation index S^G satisfies axioms UN, LN, C and IT if and only if there exists a continuous (strictly) increasing function ψ : $[0, 1] \to [0, 1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ such that $S^G(\mathcal{L}) := \psi(|\det \mathcal{L}|)$.*

Proof See Appendix A.2. □

Note that $|\det \mathcal{L}|$ is the *volume of the zonotope* associated with matrix $\mathcal{L} \in \mathcal{M}_{GG}$ (McMullen, 1971). The index S^G is hence an increasing and monotone transformation of that volume. Proposition 1 offers, to our knowledge, the first characterization of a *volume* index that applies to any situation in which $G \geq 2$ distributions are compared. As such, the measure is also novel with respect to a wide range of characterizations of the *determinant* of square matrices (since we deal with its absolute value).

By construction, the index $\psi(|\det \mathcal{L}|)$ is such that if two columns are linearly dependent then the index takes the value of 0. This is certainly a limitation for the use of this measure if one restricts attention to problems with G groups and G units. However, as we are going to show, when N is large compared to G (a common case in empirical analysis) the overall segregation measure boils down to 0 only if there are much more than G linearly independent column vectors among the N vectors associated to the units in $\mathcal{L}$. A sufficient case for this extreme result is obtained when (at least) two rows of $\mathcal{L}$ are identical. As argued, this case is of scarce relevance for empirical applications (in this situation, the two groups differ only in their respective labels, which defines an irrelevant partition of the population for the purpose of addressing segregation).

We introduce two additional axioms which allow to extend the set of comparable matrices to $\mathcal{M}_{GN}$. Additional notation is needed. Let $\mathcal{N}_O(\mathcal{L})$ denote the set of *units ordered* according to the order they appear as columns of $\mathcal{L}$. Any G dimensional subset of $\mathcal{L}$ of ordered units is denoted by $\{i_1, i_2, ..., i_G\} \subseteq \mathcal{N}_O(\mathcal{L})$ where the index i_k denotes a unit in position $k \leq G$ in the units order obtained by eliminating $N - G$ units from the initial ordered set of units $\{1, 2, ..., N\}$ in $\mathcal{L}$. The obtained sub-matrix derived from $\mathcal{L}$ by keeping the units $\{i_1, i_2, ..., i_G\}$ is denoted $(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})$. In general this square matrix of dimension G is not in $\mathcal{M}_{GG}$. This could be the case because all elements in a row are 0's, or more generally because for some/all rows the elements do not sum to 1. In order to accommodate the first case we consider matrices in $\mathcal{M}^0_{GG}$, a subset of matrices in $\mathcal{M}_{GG}$ where at most $G-1$ rows could contain all 0's while the other rows are stochastic and no column has all elements equal to 0. In the second case, the matrix could however be made row-stochastic by dividing each row (except those made of all 0's) by the corresponding element of the vector $\boldsymbol{\lambda}^{\{i_1,i_2,...,i_G\}}$ obtained by calculating the product

$$\boldsymbol{\lambda}^{\{i_1,i_2,...,i_G\}} := (\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G}) \cdot \mathbf{1}_G$$

where $\mathbf{1}_G$ denotes the G dimensional column vector of 1's. For simplicity of exposition we denote such row stochastic matrix as $(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G})$. Note that the generic element $\lambda_g^{\{i_1,i_2,...,i_G\}}$ of the vector $\boldsymbol{\lambda}^{\{i_1,i_2,...,i_G\}}$ corresponding to group g denotes the conditional probability of observing group g in one of the units of the set $\{i_1, i_2, ..., i_G\}$. The joint conditional probability distribution is given by the product of all elements of $\boldsymbol{\lambda}^{\{i_1,i_2,...,i_G\}}$.

We are now in the position to formalize the *Decomposition* property that requires that the aggregate segregation evaluation could be decomposed in the weighted sum of all evaluations made over all ordered matrices in $\mathcal{M}^0_{GG}$ weighted according to the joint probabilities.

Axiom 5 *(D: Decomposition)*

$$S^N(\mathcal{L}) := \sum_{\{i_1,i_2,...,i_G\}\subseteq\mathcal{N}_O(\mathcal{L})} \left(\prod_{g=1}^{G} \lambda_g^{\{i_1,i_2,...,i_G\}}\right) \cdot S^G(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G})$$

for all $\mathcal{L} \in \mathcal{M}_{GN}$*, where* S^G *is defined over* $\mathcal{M}^0_{GG}$.

The decomposition property is consistent with the logic adopted by the Gini index to measure dispersions in income distributions as a weighted average of pairwise individuals comparisons of their distances.

We introduce one additional axiom that postulates a form of invariance with respect to operations of splitting of units. This implies that segregation is neutral with respect to the partition of units as long as the newly obtained units preserve the extent of (dis)proportionality in groups relative frequencies.

Axiom 6 *(IS: Segregation Invariance from Splitting of units)*

$$S^{N+1}((\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \alpha\boldsymbol{\ell}_i, (1-\alpha)\boldsymbol{\ell}_i, .., \boldsymbol{\ell}_j, ..., \boldsymbol{\ell}_N) = S^N(\mathcal{L})$$

for any $\mathcal{L} \in \mathcal{M}_{GN}$*,* $i \in \mathcal{N}$*,* $0 < \alpha \leq 1$.

Note that the newly obtained distribution matrix is of dimension $G \times (N+1)$.

We are now ready to provide a complete characterization result for the GS index.

Theorem 1 *Let* $\mathcal{L} \in \mathcal{M}_{GN}$*, the segregation index* $S^N : \mathcal{M}_{GN} \to [0, 1]$ *satisfies axioms UN, LN, C, IT, IS, and D if and only if* $S^N(\mathcal{L}) := GS(\mathcal{L})$.

Proof See Appendix A.3. □

Theorem 1 bears two important contributions. First, the theorem provides a complete characterization of a volume-based segregation index. It shows in particular that the Gini Segregation index is the unique measure satisfying the axiomatic model of interest. The index is multi-group in nature and related to the multivariate inequality Gini index introduced by Koshevoy and Mosler (1996). The characterization rests on the properties that characterize the result in Proposition 1 that impose restrictions on the functional form of the segregation measure in reaction to matrix majorization operations. These operations were valid on the domain $\mathcal{M}_{GG}$, which we extend to $\mathcal{M}_{GN}$ (the domain of empirical distribution matrices) through axioms D and IS.

Second, the theorem contributes with an innovative axiomatic model which is capable of characterizing the Gini Segregation index even when $G = 2$ and that is alternative to the model introduced by Puerta and Urrutia (2016). To establish our result, we provide first an axiom that is grounded exclusively on elementary mixture of units operations, thus on a subset of the data transformations considered in Property P1.

4.2 A result with two groups

Every elementary mixture of units combines two operations, already discussed in Andreoli and Zoli (2023). The first operation requires splitting a given unit into two units of smaller size, however preserving the groups compositional differences. This is, in all effects, the operation underlying IS axiom, which is regarded as a source of indifference for all segregation measures. The second transformation underlying an elementary mixture of units requires that one of the split units obtained from unit i is merged with unit j, thereby leading to matrix $\mathcal{L}(\alpha, i, j)$. Such operation cannot increase segregation because every way of combining together two units likelihoods reduces the disparities among them and therefore reduces the dissimilarity across the distributions of the groups on the units and thus the level of segregation. As a result, if the original matrix $\mathcal{L}$ is in $\mathcal{M}_{GG}$ then also matrix $\mathcal{L}(\alpha, i, j)$ is still in $\mathcal{M}_{GG}$.

In line with Property P2, the axiom of "Segregation reduction through elementary Mixtures of units" postulates that the combination of these operations should not increase segregation. Beyond this, it also assumes that segregation should be reduced proportionally to the share of peoples displaced by the operation. Thus, if the original configuration is associated with a positive level of segregation these operations should strictly reduce it in proportion α.

Axiom 7 *(SM: Segregation reduction through elementary Mixtures of units)*

$$S^G(\mathcal{L}(\alpha, i, j)) = \alpha S^G(\mathcal{L}),$$

for all $\mathcal{L} \in \mathcal{M}_{GG}$, $i, j \in \mathcal{N}$, $0 < \alpha \leq 1$.

Axiom SM is a multidimensional generalization of a property satisfied by the Gini (income inequality) index when $G = 2$. Consider, for instance, the Gini index derived from the distribution matrix with $G = 2$ groups and $N = 2$ columns:

$$\mathcal{L} = \begin{bmatrix} p & 1-p \\ x & 1-x \end{bmatrix},$$

where for expositional purpose we assume that p denotes the proportion of poor individuals whose income is a share $x < p$ of the total income, and $1 - p$ is the proportion of rich individuals that own a share $1 - x$ of the society income.

The Gini index for this society is $p - x$. It could be calculated by computing the segregation curve that in this case is piecewise linear with coordinates $(0, 0)$, (p, x), $(1, 1)$. Suppose that we apply an elementary mixture transformation to the data by post multiplying matrix $\mathcal{L}$ by $\begin{bmatrix} \alpha & 1-\alpha \\ 0 & 1 \end{bmatrix}$. The obtained new matrix $\mathcal{L}(\alpha, 1, 2)$ will be

$$\mathcal{L}(\alpha, 1, 2) = \begin{bmatrix} \alpha p & (1-\alpha)\, p + 1 - p \\ \alpha x & (1-\alpha)\, x + 1 - x \end{bmatrix} = \begin{bmatrix} \alpha p & 1 - \alpha p \\ \alpha x & 1 - \alpha x \end{bmatrix}.$$

The Gini index of matrix $\mathcal{L}(\alpha, 1, 2)$ will be $\alpha p - \alpha x = \alpha \cdot (p - x)$, precisely as postulated by axiom SM.

An analogous result could be obtained by post multiplying $\mathcal{L}$ by $\begin{bmatrix} 1 & 0 \\ 1-\alpha & \alpha \end{bmatrix}$ which is the matrix related to the other possible elementary mixture operation created by splitting column 2 and merging it with column 1. In this case

$$\mathcal{L}(\alpha, 2, 1) = \begin{bmatrix} 1 - \alpha\,(1-p) & \alpha\,(1-p) \\ 1 - \alpha\,(1-x) & \alpha\,(1-x) \end{bmatrix}$$

whose Gini index is still $1 - \alpha\,(1-p) - 1 + \alpha\,(1-x) = \alpha \cdot (p - x)$.

The SM is consistent with Property 1 and more generally with axiom C because every elementary mixture of units transformation is a special case of those obtained making use of matrix X. The axiom focuses on specific row-stochastic transformations and moreover specifies their multiplicative impact on the level of segregation of a matrix $\mathcal{L}$.

The next theorem formalizes the general characterization of the GS index when $G = 2$ in a way that is different and alternative to that in Puerta and Urrutia (2016), but consistent with multi-group axioms employed to derive the main result. The results rests on the role of UN, SM and D and on the *Unit Anonymity* (UA) property, which can be formalized on the basis of Property 4:

Axiom 8 *(UA: Units Anonymity)* $S^N(\mathcal{L}) = S^N(\mathcal{L}\Pi_N)$ *for all* $\mathcal{L} \in \mathcal{M}_{GN}$ *and all* $\Pi_N \in \mathcal{P}_N$.

UA requires that segregation evaluations do not depend on the order of the units considered and thus expands P4 to matrices where the number of groups differs from G. Altogether, these axioms characterize the two-group Gini Segregation index.

Theorem 2 *Let* $\mathcal{L} \in \mathcal{M}_{2N}$*, the segregation index* $S^N : \mathcal{M}_{2N} \to [0, 1]$ *satisfies axioms UA, UN, SM and D if and only if* $S^N(\mathcal{L}) := GS_2(\mathcal{L})$*, that is the Gini index in (2).*

Proof See Appendix A.4. □

5 Concluding remarks

This paper offers a characterization of the Gini Segregation index. The index captures segregation stemming from the dissimilarity between distributions of groups across units. The index is a multi-group generalization of the Gini index that has been used in inequality and segregation analysis when only two distributions are involved. From a geometric perspective, the Gini Segregation index is characterized as the volume of the zonotope representation of the data. We investigate a novel axiomatic characterization of the index that identifies the minimal set of operations that are unambiguously dissimilarity reducing or preserving and are therefore relevant to characterize segregation measures. The Gini Segregation index is the unique measure consistent with the zonotope inclusion order and therefore with matrix majorization, that satisfies all the axioms.

The key axiom that we consider rests on a simple property: if two units are merged to form a larger unit, then newly obtained unit displays lower disproportionality in groups composition and a larger population size. The key axiom assumes that the effect of such merge is to reduce measured segregation proportionally to the merged share of population. Such operation has intuitive bite when units are schools or neighborhoods that can be split and recombined at will.

We also provide a generalization of the axiom. We consider the effect of a *combination* of such operations that can be represented through a row-stochastic matrix. We introduce a new set of axioms that introduce a form of consistency in segregation evaluations when distribution matrices are transformed according to the same set of transformations. Moreover, our axioms emphasize that only the set of transformations used, but not the order in which operations are performed, matter for determining the extent at which segregation should reduce in response to them. In the multi-group case, such transformations are always obtained through multiplications involving row-stochastic square matrices. Importantly, the Gini Segregation index is the unique index consistent with the axiomatic model that we outline in this paper. Axiom 5 (Decomposition) is, arguably, a strong one. It can be meaningfully weakened by considering a more general decomposition based on a functional (rather than a sum) which depends on all G-dimensional square matrices and their relative weights. Such a generalization could allow, for instance, for normative weights in the aggregation of the evaluations of the G-dimensional square matrices, delivering larger relevance in the evaluation to those matrices displaying larger disproportionalities in groups' likelihoods. When paired with Axiom 6 (Splitting), it is however arguable that a more general version of Axiom 5 would boil down to the linear aggregator that is actually used in Axiom 5.

Lastly, we notice that our axiomatic model does not request that entries of distribution matrices being calculated from groups relative frequencies distributions. Andreoli (2014), for instance, considers a setting in which distribution matrices are obtained from data on social interactions, specifying the probabilities that an individual drawn at random from a given unit has to interact with each of the groups. Such estimates can serve to derive interaction likelihood and the Gini Segregation index can be used to capture aspects related to the exposure dimension of segregation. This paves the way to a broader set of applications of the index.

On more general grounds, this paper provides a complete characterization of the *absolute value of the determinant* of a square, row-stochastic matrix. Such a measure is trivially related to volume-based indices discussed in the multidimensional inequality literature (Koshevoy & Mosler, 1997, 1996), to which this paper relates. Shorrocks (1978) has identified the absolute value of the determinant of a transition matrix (i.e., a row-stochastic matrix which reports probabilities that individuals born to parents in a given income class end up in each income class as adults) as a relevant mobility index, when origin independence is regarded as the highest level of mobility that a society could achieve. Our axiomatic model provides a natural starting point to construct a characterization of Shorrocks' index, if one is willing to interpret row-stochastic distribution matrices as transition matrices. Shorrocks' index is consistent with Property P1, which regards the product of two transition matrices as generating a more mobile society than the societies characterized by each transition matrix separately. The index thus values long-term mobility as the product of short-term transitions. The more complex the story of short transitions, the more likely is mobility to prevail. Proposition 1 paves the way for an axiomatic derivation of Shorrocks' mobility index.

A Appendix of proofs

A.1 Proof of Properties P1-P6

Property P1 This property derives from the combined effect of UN, C and IT. By UN we have that $S^G(\mathcal{L}) \leq S^G(I_G) = 1$ for all $\mathcal{L} \in \mathcal{M}_{GG}$. Substituting into the definition of C for $\mathcal{L}' = I_G$ and $\mathcal{L} = A$ and using the transformation matrix $Y \in \mathcal{M}_{GG}$ instead of X we obtain $S^G(A) \leq S^G(I_G)$ if and only if for every $A, Y \in \mathcal{M}_{GG}$, $S^G(A \cdot Y) \leq S^G(Y)$. Consider IT. Let $\mathcal{L} = I_G$ in IT, then for every $X, Y \in \mathcal{M}_{GG}$ we have $S^G(X \cdot Y) = S^G(Y \cdot X)$ using matrix $A \in \mathcal{M}_{GG}$ instead of X we have $S^G(A \cdot Y) = S^G(Y \cdot A)$. Combining the results we obtain $S^G(Y \cdot A) = S^G(A \cdot Y) \leq S^G(Y)$. If we let $Y = \mathcal{L}$ and $A = X$ we obtain the desired property.

Property P2 This property come from the fact that $\mathcal{L}(\alpha, i, j) = \mathcal{L} \cdot T(\alpha, i, j)$ where the transformation

$$T(\alpha, i, j) := \begin{array}{c} \begin{matrix} 1 & \dots & i & \dots & \dots & j & \dots \end{matrix} \\ \begin{bmatrix} 1 &&&&& \\ & 1 &&&& \\ && \alpha && 1-\alpha & \\ &&& 1 && \\ &&&& 1 & \\ &&&&& 1 \end{bmatrix} \end{array} \tag{3}$$

is a row stochastic matrix in $\mathcal{M}_{GG}$, and therefore according to Property P1 we have $S^G(\mathcal{L} \cdot T(\alpha, i, j)) \leq S^G(\mathcal{L})$.

Property P3 See text.

Property P4 The result is obtained making use of Property P1 and axioms C applied to transformations where X is a permutation matrix. According to Property P1 we have that $S^G(\mathcal{L} \cdot \Pi) \leq S^G(\mathcal{L})$ for a generic permutation matrix $\Pi \in \mathcal{P}_G$. Let $\mathcal{L}' = \mathcal{L}\Pi$. Since the inverse $\Pi^{-1} \in \mathcal{P}_G$ always exists[7], applying again P1 yields $S^G(\mathcal{L}' \cdot \Pi^{-1}) \leq S^G(\mathcal{L}')$. Since $\mathcal{L}' \cdot \Pi^{-1} = \mathcal{L}$, one has that both inequalities hold at the same time if and only if $S^G(\mathcal{L} \cdot \Pi) = S^G(\mathcal{L})$, which implies Property P4.[8]

Property P5 The result is an implication of Property P4 and axiom IT. See the proof of Property P6.

Property P6 This property is obtained letting $\mathcal{L} = X_1$, $X = X_2$ and $Y = X_3$ in IT, recalling that also by IT also $S^G(X \cdot Y) = S^G(Y \cdot X)$ holds and considering that any product of matrices in $\mathcal{M}_{GG}$ is still in $\mathcal{M}_{GG}$ and therefore by combining the products of pairs of matrices X_i and permuting them one can reach the conclusion that any permutation of the product of matrices X_1, X_2 and X_3 leads to the same segregation level. For instance $S^G(X_1 \cdot X_2 \cdot X_3) = S^G(X_3 \cdot X_2 \cdot X_1)$ because by IT $S^G(X_1 \cdot X_2 \cdot X_3) = S^G(X_1 \cdot X_3 \cdot X_2)$ and then $S^G(X_1 \cdot (X_3 \cdot X_2)) = S^G((X_3 \cdot X_2) \cdot X_1)$. To expand the conclusion to the case of the product of 4 matrices consider that the segregation is invariant to any permutation of 2 and any permutation of 3 matrices and use a similar logic applied to derive the result for 3 matrices. Recursively one can reach the invariance condition for any permutation of matrices for each number $m \geq 2$.

A.2 Proof of Proposition 1

Proof *Sufficiency part.* It is immediate to show that the index $S^G(\mathcal{L}) := \psi(|\det \mathcal{L}|)$ satisfies all axioms listed. Note that: (UN) $\det I_G = 1$, while (LN) $\det \mathcal{L}^0 = 0$, (C) $|\det \mathcal{L} \cdot X| = |\det \mathcal{L}| \cdot |\det X|$, moreover (IT) the product operation is commutative and so the application of any permutation of the same set of transformations lead to the same value for the index. The application of normalized increasing transformations $\psi(\cdot)$ does not affect the conclusions derived above for the index $|\det \mathcal{L}|$.

Necessity part. We construct the proof through a series of steps. The analysis is made for matrices in $\mathcal{M}_{GG}$, and so all the indices considered are $S^G(\cdot)$, to simplify the exposition we drop the superscript G and use the notation $S(\cdot)$.

Step 1. Consider axiom C and note that an implication of the axiom requires that $S(\mathcal{L} \cdot X) = S(\mathcal{L}' \cdot X)$ is equivalent to $S(\mathcal{L}) = S(\mathcal{L}')$. Continuity of the index $S(\cdot)$ in $\mathcal{M}_{GG}$, axioms C, N, LN and IT imply that for any $\mathcal{L}, X \in \mathcal{M}_{GG}$ there exists a continuous function $G : [0, 1] \times \mathcal{M}_{GG} \to [0, 1]$, strictly increasing in its first argument where

$$S(\mathcal{L} \cdot X) := G[S(\mathcal{L}), X]$$

[7] Recall that for any permutation matrix $\Pi \in \mathcal{P}_G$ there exists a positive integer k such that $\Pi^k = I_G$, implying $\Pi^{-1} = \Pi^{k-1}$ (see Humphreys and Prest 1989, Ch.4).

[8] We are grateful to an anonymous reviewer for suggesting this simplified version of our proof.

where $G[s, I_G] = s$ for any $s \in [0, 1]$. Moreover, we have that $S(I_G \cdot X) = S(X) = G[S(I_G), X] = G[1, X]$. It follows that by construction $S(\mathcal{L}) := G[1, \mathcal{L}]$ for any $\mathcal{L} \in \mathcal{M}_{GG}$. In addition if we consider $\mathcal{L}^0$ and X^0 we have that $S(\mathcal{L}^0 \cdot X) = S(\mathcal{L} \cdot X^0) = 0 = G[0, X] = G[s, X^0]$ for any $X^0 \in \mathcal{M}_{GG}$ and any $s \in [0, 1]$.

The result is generated by the continuity of the index on $\mathcal{M}_{GG}$ and makes use of the Properties P1 and P3 that imply that for any $\mathcal{L}$ and associated value $S(\mathcal{L})$ through the transformations X it is possible to obtain matrices $\mathcal{L} \cdot X$ in $\mathcal{M}_{GG}$ whose indices $S(\mathcal{L} \cdot X)$ range from $S(\mathcal{L})$ to 0. It is then possible to identify set of matrices that exhibit the same level for the segregation index.

Step 2. Let $\mathcal{L}, \mathcal{L}', X, Y \in \mathcal{M}_{GG} \backslash I_G$ such that $S(\mathcal{L}) = S(\mathcal{L}')$ and $S(X) = S(Y)$. Consider property IT. As argued in the discussion of Property P6 one has that in general $S(X \cdot Y) = S(Y \cdot X)$ for every $X, Y \in \mathcal{M}_{GG}$. Moreover, according to C we have that $S(\mathcal{L} \cdot X) = S(\mathcal{L}' \cdot X)$ is equivalent to $S(\mathcal{L}) = S(\mathcal{L}')$ and analogously one has $S(X \cdot \mathcal{L}) = S(Y \cdot \mathcal{L})$ is equivalent to $S(X) = S(Y)$. Combining these facts we obtain that $S(\mathcal{L} \cdot X) = S(\mathcal{L}' \cdot X) = S(X \cdot \mathcal{L}) = S(X \cdot \mathcal{L}')$ and $S(X \cdot \mathcal{L}) = S(Y \cdot \mathcal{L}) = S(\mathcal{L} \cdot X) = S(\mathcal{L} \cdot Y)$. We then obtain that $S(\mathcal{L}' \cdot X) = S(\mathcal{L} \cdot X) = S(\mathcal{L} \cdot Y) = S(Y \cdot \mathcal{L})$ for all $X, Y, \mathcal{L}, \mathcal{L}', \in \mathcal{M}_{GG}$ such that $S(X) = S(Y)$ and $S(\mathcal{L}') = S(\mathcal{L})$. This consideration allows to specify the result in Step 1. In fact one obtains that the role of the matrix X in the argument of the function $G(\cdot)$ depends on the value of $S(X)$, moreover, one also has that $G[S(\mathcal{L}), X] = G[S(X), \mathcal{L}]$. As a result for any $\mathcal{L}, X \in \mathcal{M}_{GG}$ there exists a continuous function $V : [0, 1] \times [0, 1] \to [0, 1]$, strictly increasing in both arguments and symmetric such that

$$S(\mathcal{L} \cdot X) := V[S(\mathcal{L}), S(X)]$$

where $V(s, 1) = V(1, s) = s$ for any $s \in [0, 1]$ and $V(s, 0) = V(0, s) = 0$ for any $s \in [0, 1]$.

Step 3. Consider axiom IT and the related Property P6. Let $X, Y, Z \in \mathcal{M}_{GG}$, note that $S(X \cdot Y \cdot Z) = S((X \cdot Y) \cdot Z) = S(X \cdot (Y \cdot Z))$, moreover the value is invariant with respect to any permutation of the three matrices. If we apply the result in Step 2 we obtain

$$\begin{aligned} S(X \cdot Y \cdot Z) &= V[S(X \cdot Y), S(Z)] = V[V[S(X), S(Y)], S(Z)] \\ &= V[S(X), S(Y \cdot Z)] = V[S(X), V[S(Y), S(Z)]] \end{aligned}$$

for any $X, Y, Z \in \mathcal{M}_{GG}$. If we denote by $S(X) = x$; $S(Y) = y$, and $S(Z) = z$ with $x, y, z \in [0, 1]$ and substitute we obtain the following functional equation related to the function $V(\cdot, \cdot)$:

$$V[V(x, y), z] = V[x, V(y, z)]$$

for all $x, y, z \in [0, 1]$, with V continuous, strictly increasing in both arguments and symmetric, with $V(1, s) = s$ and $V(0, s) = 0$ for any $s \in [0, 1]$. Moreover, by Property 4 the function is also invariant with respect to any permutation of x, y, z.

This is a specification of the Generalized Associativity Equation analyzed in Chapter 6.2 of Aczel (1966).

The general solution of the functional equation requires that there exists a continuous and strictly increasing functions $\Psi : (0, 1] \rightarrow (-\infty, 1]$ such that

$$V(x, y) := \Psi^{-1}[\Psi(x) + \Psi(y)]$$

for all $x, y \in (0, 1]$, with $\Psi(1) = 0$ and $\lim_{x\rightarrow 0^+} \Psi(x) = -\infty$, and $V(0, s) = V(s, 0) = 0$ for any $s \in [0, 1]$.

If we let $\Psi(x) := \ln \phi(x)$ for any $x \in (0, 1]$ we can express the result in more compact terms by obtaining that there exists a continuous and strictly increasing functions $\phi : [0, 1] \rightarrow [0, 1]$ such that

$$V(x, y) := \phi^{-1}[\phi(x) \cdot \phi(y)]$$

for all $x, y \in [0, 1]$, with $\phi(1) = 1$ and $\phi(0) = 0$ where the latter condition allows to obtain that $V(0, s) = V(s, 0) = 0$ for any $s \in [0, 1]$.

Step 4. We now recover explicitly the functional form for $S(\mathcal{L})$. Recall that from Step 2 and Step 3 we have that $S(X \cdot Y) = V(x, y)$ for any $X, Y \in \mathcal{M}_{GG}$ where $S(X) = x$ and $S(Y) = y$ with $x, y \in [0, 1]$. Combining with the obtained solution of the functional equation $V(x, y)$ we have that

$$S(X \cdot Y) = \phi^{-1}[\phi \circ S(X) \cdot \phi \circ S(Y)]$$

for all $X, Y \in \mathcal{M}_{GG}$. Recalling that $\phi : [0, 1] \rightarrow [0, 1]$ is strictly increasing we can write $\phi \circ S(X \cdot Y) = \phi \circ S(X) \cdot \phi \circ S(Y)$. By letting $F(X) := \phi \circ S(X)$, we then obtain that $F : [0, 1] \rightarrow [0, 1]$ is a continuous function by combining continuity of ϕ and of the function $S(\cdot)$. We thus obtain the functional equation

$$F(XY) = F(X) \cdot F(Y)$$

for all $X, Y \in \mathcal{M}_{GG}$.

The general solution of this functional equation is illustrated in chapter 4.2 of Aczél and Dhombres (1989) for all real valued square matrices and real valued functions F. It coincides (see Theorem 8 and Corollary 6, p. 41-42) with either (i) $F(X) = |\det X|^c$ or (ii) $F(X) = |\det X|^c \cdot \text{sign}(\det X)$ for any c arbitrary real constant. Note that by construction X is row stochastic and therefore $|\det X| \leq 1$, moreover the value of F ranges in $[0, 1]$, it then follows that $c \geq 0$. Moreover, by Property P4 the value of $F(X)$ is independent from permutations of the columns of X, this fact eliminates solution (ii) because any permutation of a matrix changes the sign of its determinant. In addition $F(X)$ is continuous for any $X \in \mathcal{M}_{GG}$ which rules out the case where $c = 0$, in fact in this case we have that either $F(X) = 1$ for all the cases where $|\det X| \neq 0$ or $F(X) = 0$ if $|\det X| = 0$ which is the case for instance for any matrix X^0.

We are then left with

$$F(X) = |\det X|^c$$

with $c > 0$ for all $X \in \mathcal{M}_{GG}$.

In order to recover the functional form for $S(\cdot)$ one has to consider that by definition $F(\mathcal{L}) := \phi \circ S(\mathcal{L})$ for all $\mathcal{L} \in \mathcal{M}_{GG}$. Recalling that ϕ is strictly increasing and therefore invertible one then obtains $S(\mathcal{L}) := \phi^{-1} \circ F(\mathcal{L})$, that is $S(\mathcal{L}) := \phi^{-1} \circ |\det X|^c$ for $c > 0$. Note that any strictly increasing transformation of $|\det X|^c$ for $c > 0$ can be written as a strictly increasing function of $|\det X|$, say function $\psi : [0, 1] \to [0, 1]$ that we can set such that $\psi(s) := \phi^{-1} \circ s^c$ for $c > 0$.

We can then conclude with the result where there exists function $\psi : [0, 1] \to [0, 1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ such that $S^G(\mathcal{L}) := \psi(|\det \mathcal{L}|)$ for all $\mathcal{L} \in \mathcal{M}_{GG}$. □

A.3 Proof of Theorem 1

Proof We already know from Proposition 1 that S^G when applied to matrices in $\mathcal{M}_{GG}$ satisfies axioms UN, LN, C and IT if and only if where there exists function $\psi : [0, 1] \to [0, 1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ such that $S^G(\mathcal{L}) := \psi(|\det \mathcal{L}|)$. The axioms apply exclusively to matrices in this class. We combine now this result with the restrictions imposed by IS and D in order to generalize the result to the case where $N > G$.

Necessity part. We illustrate first the restrictions on the function $\psi(\cdot)$ for the case where $G = 2$. Apply a split of a column vector of matrix $\mathcal{L} = (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2) \in \mathcal{M}_{22}$, say vector $\boldsymbol{\ell}_2$ that is split into vectors $\boldsymbol{\ell}_{2_1} = \alpha\boldsymbol{\ell}_2$, and $\boldsymbol{\ell}_{2_2} = (1-\alpha)\boldsymbol{\ell}_2$, for $0 < \alpha < 1$. The new matrix is $\mathcal{L}' = (\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2, (1-\alpha)\boldsymbol{\ell}_2)$. We consider first the matrix $\mathcal{L} = \begin{bmatrix} d & 1-d \\ 0 & 1 \end{bmatrix}$ with $1 \geq d > 0$ from which we derive matrix $\mathcal{L}' = \begin{bmatrix} d & \alpha(1-d) & (1-\alpha)(1-d) \\ 0 & \alpha & 1-\alpha \end{bmatrix}$. Making use of D, following the discussion in the proof of Proposition 1, and considering that $S^G(\mathcal{L}) = \psi(|\det \mathcal{L}|)$ with $\psi(0) = 0$ and $\psi(1) = 1$ for all $\mathcal{L} \in \mathcal{M}_{GG}$ one obtains

$$\begin{aligned} S^2(\mathcal{L}') = & \frac{1}{2!}\beta_1\psi(|\det(\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2)|/\beta_1) + \frac{1}{2!}\beta_1\psi(|\det(\alpha\boldsymbol{\ell}_2, \boldsymbol{\ell}_1)|/\beta_1) \\ & + \frac{1}{2!}\beta_2\psi(|\det(\boldsymbol{\ell}_1, (1-\alpha)\boldsymbol{\ell}_2)|/\beta_2) + \frac{1}{2!}\beta_2\psi(|\det((1-\alpha)\boldsymbol{\ell}_2, \boldsymbol{\ell}_1)|/\beta_2) \\ & + \frac{1}{2!}\beta_3\psi(|\det((1-\alpha)\boldsymbol{\ell}_2, \alpha\boldsymbol{\ell}_2)|/\beta_3) + \frac{1}{2!}\beta_3\psi(|\det(\alpha\boldsymbol{\ell}_2, (1-\alpha)\boldsymbol{\ell}_2)|/\beta_3) \end{aligned}$$

where $\beta_1 = \alpha \cdot [d + \alpha(1-d)]$, $\beta_2 = (1-\alpha) \cdot [d + (1-\alpha)(1-d)]$ and $\beta_3 = 1-d$. After simplifications the condition becomes

$$S^2(\mathcal{L}') = \beta_1\psi(|\det(\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2)|/\beta_1) + \beta_2\psi(|\det(\boldsymbol{\ell}_1, (1-\alpha)\boldsymbol{\ell}_2)|/\beta_2) + \beta_3\psi(0)$$

$$= \beta_1 \psi(\alpha d/\beta_1) + \beta_2 \psi((1-\alpha)d/\beta_2) + \beta_3 \psi(0).$$

Substituting and recalling that $\psi(0) = 0$ we obtain that

$$S^2(\mathcal{L}') = \alpha[d + \alpha(1-d)] \cdot \psi\left(\frac{d}{d + \alpha(1-d)}\right)$$
$$+(1-\alpha) \cdot [d + (1-\alpha)(1-d)] \cdot \psi\left(\frac{d}{d + (1-\alpha)(1-d)}\right)$$

for $0 < \alpha < 1$ and $0 < d \leq 1$. According to IS it should hold that $S^2(\mathcal{L}') = S^2(\mathcal{L})$ where $S^2(\mathcal{L}) = \psi(|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)|) = \psi(d)$. We need to verify the restrictions on $\psi(\cdot)$ that guarantee that $S^2(\mathcal{L}') = \psi(d)$ for all $0 < \alpha < 1$ and all $0 < d \leq 1$. We first assume that $\alpha = 1/2$ and derive the condition

$$S^2(\mathcal{L}') = \frac{1}{2}\frac{d+1}{2} \cdot \psi\left(\frac{2d}{d+1}\right) + \frac{1}{2}\frac{d+1}{2} \cdot \psi\left(\frac{2d}{d+1}\right) = \frac{d+1}{2} \cdot \psi\left(\frac{2d}{d+1}\right)$$
$$= \psi(d) = S^2(\mathcal{L})$$

for all $0 < d \leq 1$. That is $\frac{d+1}{2d} \cdot \psi\left(\frac{2d}{d+1}\right) = \psi(d)\frac{1}{d}$ for all $0 < d \leq 1$.

Let $h(t) := \psi(t)/t$ for all $0 < t \leq 1$. The above condition requires that the continuous function $h(\cdot)$ satisfies

$$h\left(\frac{2d}{d+1}\right) = h(d)$$

for all $0 < d \leq 1$.

Construct now a sequence of values d_k for $k \in \mathbb{N}\cup\{0\}$ starting from $0 < d_0 \leq 1$ such that $d_{k+1} = \frac{2d_k}{d_k+1}$. Considering the value of $h(\cdot)$ for all the points in the sequence, by construction we obtain that $h(d_k) = h(d_0)$ for all k.

If $d_0 = 1$ then $d_k = 1$ for all $k \in \mathbb{N}$. Therefore $h(d_k) = h(1) = \bar{h}$.

If $0 < d_0 < 1$, for each d_0 we have a sequence of terms $d_k(d_0)$. The sequence is increasing, moreover, by construction each sequence converges to 1, that is $\lim_{k\to\infty} d_k(d_0) = 1$ for all $0 < d_0 < 1$. Recalling that $h(d_k(d_0)) = h(d_0)$ for all $0 < d_0 < 1$ and $k \in \mathbb{N}$ and that the function $h(\cdot)$ is continuous, then $h(d_0) = \lim_{k\to\infty} h(d_k(d_0)) = h(1) = \bar{h}$ for all $0 < d_0 < 1$.

We have then obtained that $h(d) = \bar{h}$ for all $0 < d \leq 1$, leading to $\psi(d)/d = \bar{h}$. Letting $d = 1$ and recalling that $\psi(1) = 1$ we obtain $\bar{h} = 1$. Considering that $\psi(0) = 0$ we finalize the result as $\psi(d) = d$ for all $0 \leq d \leq 1$.

Note that no additional restrictions are required other than $\psi(d) = d$ in order to satisfy the initial condition, in fact one obtains that $S^2(\mathcal{L}') = d = S^2(\mathcal{L}')$.

As a result the transformation function $\psi(\cdot)$ is the identity function thus leading to $S^2(\mathcal{L}) = |\det \mathcal{L}|$.

Note now that the result in Proposition 1 on the specification of $S^G(\mathcal{L}) := \psi(|\det \mathcal{L}|)$ is obtained for a fixed value of $G \geq 2$ and therefore in general the transformation $\psi(\cdot)$ may depend on G. Following the same logic the above considerations related to the effect on $\psi(\cdot)$ of the additional axioms IS and D can be expanded to the case where $G > 2$. In all these cases the restrictions that guarantees that IS is satisfied require that $\psi(\cdot)$ is the identity function. In order to prove the result one can consider a starting matrix $G \times G$ with the same terms of the 2×2 matrix considered above placed in the corners of the matrix and all 1's in the remaining terms of the main diagonal and 0's in all other cells, that is a matrix with $\ell_{11} = d$, $\ell_{1G} = 1 - d$, $\ell_{ii} = 1$ for $i = 2, 3, ..., G$ and all other terms equal 0.

To conclude, note that we have now that $S^G(\mathcal{L}) = |\det \mathcal{L}|$ for all $\mathcal{L} \in M_{GG}$. By direct application of the definition of axiom D, note first that it is not necessary for the computation of matrices where at least one row is made of zeros, because in this case the associated $\left(\Pi_{g=1}^G \lambda_g^{\{i_1,i_2,...,i_G\}}\right) = 0$. We should therefore focus on matrices in $\mathcal{M}_{GN}$. For these matrices, we apply axiom D to obtain

$$S^N(\mathcal{L}) = \sum_{\{i_1,i_2,...,i_G\} \subseteq \mathcal{N}_{O(\mathcal{L})}} \left(\Pi_{g=1}^G \lambda_g^{\{i_1,i_2,...,i_G\}}\right) \cdot \left|\det(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G})\right|.$$

However, note that by construction

$$\left|\det(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G})\right| = \left|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})\right| \cdot \left(\Pi_{g=1}^G \lambda_g^{\{i_1,i_2,...,i_G\}}\right)^{-1}. \tag{4}$$

After simplifying in the $S^N(\mathcal{L})$ formula we obtain

$$S^N(\mathcal{L}) = \sum_{\{i_1,i_2,...,i_G\} \subseteq \mathcal{N}_{O(\mathcal{L})}} \left|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})\right|.$$

Here the ordered distribution of the G units in $\mathcal{N}_{O(\mathcal{L})}$ is taken into account. If we allow all possible permutations of these units as in $\mathcal{N}(\mathcal{L})$ then we obtain for each ordered set of units $G!$ times the same index $\left|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})\right|$ that, being expressed in absolute terms is not modified by permutation of the columns. These considerations lead to the final result.

Sufficiency part. The Gini Segregation index is consistent with D as a consequence of equation (4) where $S^G(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G}) = \left|\det(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G})\right|$. For the reasons noted above, the Gini Segregation index is also consistent with IS. The sufficiency part in the proof of Proposition 1 for the specific case where ψ is the identity guarantees consistency with axioms UN, LN, C and IT for matrices $(\tilde{\boldsymbol{\ell}}_{i_1}, \tilde{\boldsymbol{\ell}}_{i_2}, ..., \tilde{\boldsymbol{\ell}}_{i_G}) \in \mathcal{M}_{GG}$. □

A.4 Proof of Theorem 2

The proof of the Theorem relies on a Lemma, which provides a characterization of the Gini Segregation index based on axioms UN, UA, D and SM (and thus does not invoke axioms LN, C, IT and IS) but that is valid only within a restricted domain, that of matrices in $\hat{\mathcal{M}}_{GN} \subseteq \mathcal{M}_{GN}$ that could be obtained one from the other only by applying permutation of units or elementary merge of units transformations.

As we will show when $G = 2$ the two sets coincide, that is $\hat{\mathcal{M}}_{22} = \mathcal{M}_{22}$. However, $\hat{\mathcal{M}}_{GG}$ could be strictly included in $\mathcal{M}_{GG}$ for $G \geq 3$ and therefore the result holds for a subset of all distribution matrices. As shown in Andreoli and Zoli (2023), any matrix in $\mathcal{M}_{GG}$ can be obtained from I_G through a finite sequence of splitting of units, merge of units and permutation of units. Even though split and merge operations are the basis for the elementary mixture of units, there is no guarantee that any appropriate sequence could be decomposed combining all split and merge operations into elementary mixture operations so that the starting and arriving matrices are all square matrices of dimension G.

As stated $\hat{\mathcal{M}}_{GG}$ and $\mathcal{M}_{GG}$ could not coincide when $G \geq 3$. However, the degree of flexibility in expanding the set $\hat{\mathcal{M}}_{GG}$ so that it could approximate $\mathcal{M}_{GG}$, is “large”. In fact the construction of matrices in $\mathcal{M}_{GG}$ requires to identify $G \cdot (G-1)$ values, because matrices are row-stochastic. Moreover, the inequalities restrictions among all the values of the matrix are at most $\frac{1}{2}\left[(G^2-1)\cdot G^2\right]$. On the other hand, in order to construct $\hat{\mathcal{M}}_{GG}$ it is possible to use $G \cdot (G-1)$ transformation matrices as $T(\alpha, i, j)$ in (5) each one with a possibly different parameter α. Moreover, the product of these matrices could be permuted in $[G \cdot (G-1)]!$ configurations, and could be possibly integrated in the sequence by insertion of permutation matrices. This large flexibility in the number of parameters and operations behind the construction of $\hat{\mathcal{M}}_{GG}$ however does not allow to reach $\mathcal{M}_{GG}$ and in fact $\hat{\mathcal{M}}_{GG} \subset \mathcal{M}_{GG}$ whenever $G \geq 3$.[9]

Lemma 1 *Let $\mathcal{L} \in \hat{\mathcal{M}}_{GN}$, the segregation index $S^N : \hat{\mathcal{M}}_{GN} \to [0, 1]$ satisfies axioms UA, UN, SM, and D if and only if $S^N(\mathcal{L}) := GS(\mathcal{L})$.*

Proof *Necessity part.* Consider matrix $\mathcal{L} \in \hat{\mathcal{M}}_{GG}$. By construction, given that $\mathcal{L} = I_G\mathcal{L}$ and that $\mathcal{L} \in \hat{\mathcal{M}}_{GG}$, it follows that there is a finite sequence of elementary mixture of units/columns transformations and permutations of units/columns that allow to construct $\mathcal{L}$ starting from I_G. These transformations can be summarized in terms of matrices multiplications, by considering permutations matrices in $\mathcal{P}_G$ and matrices

[9] A counterexample can be constructed considering a matrix in $\mathcal{M}_{GG}$ whose elements on the main diagonal are all 0’s. If $G \geq 3$ it could be proved that there is no sequence of elementary mixture operations and permutation of columns that allow to reach such matrix starting from an identity matrix. Thus such a matrix is not in $\hat{\mathcal{M}}_{GG}$.

$T(\alpha, i, j) \in \mathcal{M}_{GG}$ such that

$$\begin{array}{rc} & \begin{matrix} 1 & \dots & i & \dots & \dots & j & \dots \end{matrix} \\ T(\alpha, i, j) = & \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \alpha & & & 1-\alpha & \\ & & & 1 & & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix} \end{array} \tag{5}$$

where all empty cells should be occupied by zeros. Let index by k the elements of the sequence of transformations $T(\alpha_k, i_k, j_k)$, where $\alpha_k \in (0, 1]$ denotes the mixture coefficient and i_k, j_k denote the columns involved in the mixture at stage k. It follows that either

$$\mathcal{L} = I_G \Pi_G^1 \cdot \prod_k T(\alpha_k, i_k, j_k) \cdot \Pi_G^2 \tag{6}$$

or matrix $\mathcal{L}$ can be decomposed similarly by also inserting some permutations matrices among the elements of the sequence of transformations $T(\alpha_k, i_k, j_k)$.

We consider first the case in (6). The first permutation matrix Π_G^1 in the sequence can be considered as a column permutation of the matrix I_G. Note that Π_G^1 could also coincide with an identity matrix and therefore leads to no effects on the sequence of operations. Thus, $S^G(\mathcal{L}) = S^G\left(I_G \Pi_G^1 \cdot \prod_k T(\alpha_k, i_k, j_k) \cdot \Pi_G^2\right)$. Making use of axiom UA it follows that $S^G(\mathcal{L}) = S^G\left(I_G \Pi_G^1 \cdot \prod_k T(\alpha_k, i_k, j_k)\right)$. Note that if we apply the transformation associated with $T(\alpha_k, i_k, j_k)$ to matrix $I_G \Pi_G^1$ we obtain $I_G \Pi_G^1 \cdot T(\alpha_k, i_k, j_k)$. Recall that according to axiom UA in combination with axiom UN, it follows that $S^G(I_G \Pi_G^1) = 1$. Then by applying axiom SM we obtain that $S^G[I_G \Pi_G^1 \cdot T(\alpha_k, i_k, j_k)] = \alpha_k \cdot S^G(I_G \Pi_G^1) = \alpha_k$. By repeated application of axiom SM one obtains that

$$S^G(\mathcal{L}) = S^G\left(\prod_k T(\alpha_k, i_k, j_k)\right) = \prod_k \alpha_k. \tag{7}$$

Note that by construction $\alpha_k = \det\left(T(\alpha_k, i_k, j_k)\right)$. Moreover, by making use of the property that the determinant of the product of square matrices is equal to the product of the determinants of these matrices, one obtains that

$$\prod_k \alpha_k = \det\left[\prod_k T(\alpha_k, i_k, j_k)\right]. \tag{8}$$

Note, however, that all elements α_k are positive, and thus the above relation holds also if we consider the absolute value of the determinant. This in general should be the case if we consider matrix $\mathcal{L}$ that could be obtained permuting either matrix I_G and/or matrix $\prod_k T(\alpha_k, i_k, j_k)$. These operations may invert the sign of the determinant and therefore we may have that combining (7) and (8) one obtains

$$|\det \mathcal{L}| = \det \left[\prod_k T(\alpha_k, i_k, j_k) \right] = \prod_k \alpha_k = S^G(\mathcal{L}).$$

The above considerations could be extended to the case where permutation matrices in $\mathcal{P}_G$ are inserted into the sequence of operations $T(\alpha_k, i_k, j_k)$. These insertions do not affect the final result, but allow to enrich the set of matrices that can be reached by the combinations of operations.

So far we have considered the case where $S^G(\mathcal{L}) > 0$, by continuity of S^G we can also approach the situations where $S^G(\mathcal{L}) = 0$, these can be obtained as limiting cases where some $\alpha_k \to 0$.

In order to complete the proof we need to verify the uniqueness of the index and the sufficiency part.

Suppose that the sequence of matrices $T(\alpha_k, i_k, j_k)$ and of the permutation matrices Π_G is not unique. By construction then either it leads to the same value $|\det \mathcal{L}|$ or it is not possible that the sequence leads to $\mathcal{L}$ because $|\det \mathcal{L}|$ is uniquely defined.

Sufficiency. The index $S^G(\mathcal{L}) = |\det \mathcal{L}|$ satisfies all the three axioms UA, UN and SM. In fact the absolute value of the determinant is not affected by permutation of the columns of a matrix (axiom UA) and it equals 1 if $\mathcal{L} = I_G$ (axiom UN). To prove that also axiom SM is satisfied, we need to combine some properties of the determinants.

Recall in the definition of axiom SM the notation for

$$\mathcal{L}(\alpha, i, j) = (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \alpha\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G).$$

First note that if one column of a matrix is multiplied by α also its determinant is multiplied by α. It then follows that

$$\begin{aligned} \det \mathcal{L}(\alpha, i, j) &= \det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \alpha\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G) \\ &= \alpha \det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G). \end{aligned}$$

Then recall that the determinants are multilinear functionals and therefore

$$\begin{aligned} &\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G) \\ &= \det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, , ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j, ..., \boldsymbol{\ell}_G) + (1-\alpha) \det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G). \end{aligned}$$

Note that the last determinant equals 0 because two columns are identical to $\boldsymbol{\ell}_i$. It then follows that

$$\begin{aligned} &\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j + (1-\alpha)\boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_G) \\ &= \det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, ..., \boldsymbol{\ell}_j, ..., \boldsymbol{\ell}_G) = \det \mathcal{L}. \end{aligned}$$

To summarize, $\det \mathcal{L}(\alpha, i, j) = \alpha \cdot \det \mathcal{L}$. The above considerations are not affected if one considers the absolute value of the determinant, thereby leading to $S^G(\mathcal{L}(\alpha, i, j)) = |\det \mathcal{L}(\alpha, i, j)| = \alpha \cdot |\det \mathcal{L}| = \alpha S^G(\mathcal{L})$ as required by axiom SM. $\square$

We combine the conclusions of Lemma 1 with the observation that $\hat{\mathcal{M}}_{2,2} = \mathcal{M}_{2,2}$ to prove Theorem 2 for the specific case where $G = 2$.

Proof Consider only situations in which $G = 2$, within this set it can be shown that $\hat{\mathcal{M}}_{2,2} = \mathcal{M}_{2,2}$. That is, for any $\mathcal{L} \in \mathcal{M}_{2,2}$, $S^2(\cdot)$ satisfies SM, UN and UA if and only if $S^2(\mathcal{L}) = |\det(\mathcal{L})|$. In order to prove this statement, consider matrix

$$L = \begin{bmatrix} p & 1-p \\ x & 1-x \end{bmatrix}$$

where $p > x$, we show that it can be obtained as the product of matrices

$$T(\alpha_1, 1, 2) = \begin{bmatrix} \alpha_1 & 1-\alpha_1 \\ 0 & 1 \end{bmatrix}, \text{ and } T(\alpha_2, 2, 1) = \begin{bmatrix} 1 & 0 \\ 1-\alpha_2 & \alpha_2 \end{bmatrix}$$

and eventually the permutation matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Consider the product

$$T(\alpha_1, 1, 2) \cdot T(\alpha_2, 2, 1) = \begin{bmatrix} \alpha_1 + (\alpha_1 - 1)(\alpha_2 - 1) & \alpha_2 (1-\alpha_1) \\ 1-\alpha_2 & \alpha_2 \end{bmatrix},$$

it follows that $\alpha_2 = 1 - x$ and $\alpha_2 (1 - \alpha_1) = 1 - p$ that is $\alpha_1 = \frac{p-x}{1-x}$. This latter term is consistent with the definition of $\alpha_1 > 0$. If this was not the case then one had to permute the columns of the matrix to get the result.

Recall that the value of the index could be obtained by multiplying $\alpha_1 \cdot \alpha_2$. It follows that $S^2(L) = \frac{p-x}{1-x} \cdot (1-x) = p - x = |\det(L)|$ that also coincides with the Gini index of the distribution.

The circumstance in which $p = x$ falls outside the domain of application of Axiom SM. This case can, however, be obtained as limiting situation where $\alpha_1 \to 0$, utilizing an analogous argument as in the proof of Lemma 1.

Using the results in Lemma 1, it is easy to demonstrate that the only index that is capable of ranking configurations in $\mathcal{M}_{2N}$ consistently with axioms UA, UN, SM and D is the Gini index in (2). $\square$

Funding Open access funding provided by Università degli Studi di Verona within the CRUI-CARE Agreement.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Aczel, J. (1966). *Lectures on functional equations and their applications*. New York: Academic Press.

Aczél, J. & Dhombres, J. (1989) . *Functional Equations in Several Variables*, Cambridge University Press.

Alonso-Villar, O., & del Rio, C. (2010). Local versus overall segregation measures. *Mathematical Social Sciences, 60*, 30–38.

Andreoli, F. (2014). The Gini-Exposure index for measuring segregation in networks, *Rivista Italiana degli Economisti/The Journal of the Italian Economic Association XIX* (1): 129–168.

Andreoli, F., & Zoli, C. (2020). From unidimensional to multidimensional inequality: a review. *METRON, 78*, 5–42.

Andreoli, F., & Zoli, C. (2023). Robust dissimilarity comparisons with categorical outcomes. *Social Choice and Welfare, 60*, 397–437.

Clark, W. A. V., Anderson, E., Östh, J., & Malmberg, B. (2015). A multiscalar analysis of neighborhood composition in Los Angeles, 2000–2010: A location-based approach to segregation and diversity. *Annals of the Association of American Geographers, 105*(6), 1260–1284. https://doi.org/10.1080/00045608.2015.1072790

Cutler, D., & Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics, 112*, 827–872.

Dahl, G. (1999). Matrix majorization, *Linear Algebra and its Applications 288*: 53 – 73. http://www.sciencedirect.com/science/article/B6V0R-3W6M0B4-6/2/6eec84941915b49acd7c8c9133747cb9

Duncan, O. D. & Duncan, B. (1955). A methodological analysis of segregation indexes, *American Sociological Review 20*(2): 210–217. http://www.jstor.org/stable/2088328

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

Echenique, F., Fryer, Roland G., J. & Kaufman, A. (2006). Is school segregation good or bad?, *American Economic Review 96*(2): 265–269. http://www.aeaweb.org/articles.php?doi=10.1257/000282806777212198

Flückiger, Y., & Silber, J. (1999). *The Measurement of Segregation in the Labour Force*. Heidelberg, Germany: Physica-Verlag.

Frankel, D. M. & Volij, O. (2011). Measuring school segregation, *Journal of Economic Theory 146*(1): 1 – 38.
http://www.sciencedirect.com/science/article/B6WJ3-5161PK7-7/2/f3c23c21aeeb1e2ba8377436e3fa937b

Gajdos, T., & Weymark, J. A. (2005). Multidimensional generalized gini indices. *Economic Theory, 26*(3), 471–496. https://doi.org/10.1007/s00199-004-0529-x

Humphreys, J. F., & Prest, M. Y. (1989). *Numbers*. Groups and Codes: Cambridge University Press.

Hutchens, R. M. (1991). Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations, *Mathematical Social Sciences* 21(1): 31 – 51.

http://www.sciencedirect.com/science/article/B6V88-4582FM1-1D/2/b1891c8954a6d40963da229a60acbffa
Hutchens, R. M. (2004). One measure of segregation, *International Economic Review 45*(2): 555–578. http://www.jstor.org/stable/3663531
Koshevoy, G. (1995). Multivariate Lorenz majorization. *Social Choice and Welfare, 12*, 93–102. https://doi.org/10.1007/BF00182196
Koshevoy, G. A. & Mosler, K. (1997) . Multivariate Gini indices, *Journal of Multivariate Analysis* **60**(2): 252–276. http://ideas.repec.org/a/eee/jmvana/v60y1997i2p252-276.html
Koshevoy, G. & Mosler, K. (1996). The Lorenz zonoid of a multivariate distribution, *Journal of the American Statistical Association* **91**(434): 873–882. http://www.jstor.org/stable/2291682
Lasso de la Vega, C., & Volij, O. (2014). Segregation, informativeness and Lorenz dominance. *Social Choice and Welfare, 43*(3), 547–564. https://doi.org/10.1007/s00355-014-0801-3
Massey, D. S. & Denton, N. A. (1988). The dimensions of residential segregation, *Social Forces 67*(2): 281–315.
http://www.jstor.org/stable/2579183
McMullen, P. (1971). On zonotopes. *Transactions of the American Mathematical Society, 159*, 91–109.
O'Sullivan, D., & Wong, D. W. S. (2007). A surface-based approach to measuring spatial segregation. *Geographical Analysis, 39*(2), 147–168. https://doi.org/10.1111/j.1538-4632.2007.00699.x
Puerta, C., & Urrutia, A. (2016). A characterization of the Gini segregation index. *Social Choice and Welfare, 47*(3), 519–529. https://ideas.repec.org/a/spr/sochwe/v47y2016i3d10.1007_s00355-016-0980-1.html
Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology,32*, 33–67. http://www.jstor.org/stable/3186153
Reardon, S. F., & O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological Methodology, 34*(1), 121–162. https://doi.org/10.1111/j.0081-1750.2004.00150.x
Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica, 46*(5), 1013–1024. http://www.jstor.org/stable/1911433.
Wong, D. W. S. (1993). Spatial indices of segregation. *Urban Studies, 30*(3), 559–572.