The interaction dimension of segregation: Measurement and evidence from US metro areas.*

Francesco Andreoli † Claudio Zoli‡

University of Verona and LISER, University of Verona

This draft: January 2019

Abstract

We study the heterogeneity of social interaction profiles among individuals and define the extent of the interaction dimension of segregation. An interaction profile quantifies the probabilities that one individual has to interact with different social groups. It can be inferred, for instance, from observation of social ties through networks data. Heterogeneity is minimal if everybody exhibit the same profile, and is maximal if everybody interacts with only one group. All the in-between configurations can be ordered on the bases of an intuitive principle based on operation that generate mixtures of interaction profiles. We proposes a characterization of the multi-group Gini-exposure index to assess heterogeneity in interaction patters in a society. We stress the empirical relevance of the index using information of exposure segregation along ethnic lines across US metro area over the last 35 years.

Keywords: Interaction, segregation, dissimilarity, Gini index, US census, American Community Survey.

JEL Codes: R23, D63, D85.

*An erlier version of this paper has benefitted from the comments and suggestions by Eugenio Peluso, Alain Trannoy, Dirk Van de gaer and participants to the SSCW conference, the IT10 Winter School and PET Conference. This paper forms part of the research project *The Measurement of Ordinal and Multidimensional Inequalities* (grant ANR-16-CE41-0005-01) of the French National Agency for Research and of the project EZOP (AFR grant 5932132) of the the Fonds National de la Recherche Luxembourg, whose respective financial support is gratefully acknowledged.

†Francesco Andreoli, Luxembourg Institute of Socio-Economic Research (LISER), 11 Porte des Sciences, L4366 Esch-sur-Alzette, Luxembourg and University of Verona, Department of Economics, Via Cantarane 24, 37129 Verona, Italy. E-mail: francesco.andreoli@univr.it.

‡Corresponding author. Claudio Zoli, University of Verona, Department of Economics, Via Cantarane 24, 37129 Verona, Italy. E-mail: claudio.zoli@univr.it.

1 Introduction

In this paper we focus on the mesurement of the exposure dimension of residential segregation. Following the seminal work of Massey and Denton (1988) (p.287): "Residential exposure refers to the degree of potential contact, or the possibility of interaction, between minority and majority group members within geographic areas of a city. Indices of exposure measure the extent to which minority and majority members physically confront one another by virtue of sharing a common residential area."

The measurement of this phenomenon requires to partition the population into groups and to know the distribution of these groups across *organizational units*, such as neighborhoods (Reardon and O'Sullivan 2004, Cutler and Glaeser 1997), school assignment (Frankel and Volij 2011, Echenique, Fryer and Kaufman 2006) or job types (Flückiger and Silber 1999, Hutchens 1991, Hutchens 2004).

We focus on multi-group measures of segregation in the exposure/interaction dimension by extending the original notion of exposture to the case where many groups can interact. We are interested in the distributional information that can be retained from a network of connections between units belonging to different groups, and not on the network's structure itself. In order to implement this approach with available date we consider *interaction profiles*, that correspond to vectors of probabilities that every unit in a network has to interact with each of the groups that compose the society.

We contribute to the literature on segregation measurement by characterizing a multi-group index of segregation, the *Gini Exposure index*, which measures segregation in a network as a form of inequality in the distribution of interaction profiles.

Following Massey and Denton (1988), exposure measures should capture the differences across groups in the likelihood that any randomly selected individual from one of these groups interacts with a person/unit from his own group or from another group. Segregation is zero when the chances that any two randomly selected individuals interact are made independent on their respective groups of origin. On the contrary, segregation is maximized whenever every individual interactions are limited to the members of the same group.

Segregation measurement (Massey and Denton 1988, Reardon and Firebaugh 2002, Reardon and O'Sullivan 2004, Frankel and Volij 2011) has mainly focused on the rankings produced by segregation indices for populations partitioned into two or many groups. None of these indices, however, has been designed to deal with problems of segregation that use individual level data and the axiomatization of these indices, where it exists (see for instance Hutchens 1991, Flückiger and Silber 1999, Reardon and O'Sullivan 2004, Frankel and Volij 2011), cannot be meaningfully adapted to capture segregation patterns across individuals interaction profiles.

We fill this gap by considering a framework to study segregation at individual level, conceptualized as a form of inequality in interaction profiles. We provide an axiomatic characterization of the Gini Exposure index that generalizes to the multi-group case the traditional Gini index of segregation. The index can be interpreted as the Gini volume index discussed in the multivariate inequality measurement literature (Koshevoy and

Mosler 1996, Koshevoy and Mosler 1997, Arnold 2005). To our knowledge this is the first characterization of an multivariate Gini volume index.

The axiomatic characterization of the Gini Exposure index is mainly based on operations defined on interaction profiles that preserve or decrease segregation. Our analysis is grounded on a simple principle: when the number of units equals the number of the groups, if a portion of a unit interaction profile is merged with another unit, this mixture operation should not increase the segregation. These considerations are in line with those adopted in (Andreoli and Zoli 2014) to characterize dissimilarity partial orders.

Our characterization of the multi-group Gini segregation index is illustrated in two steps, first we derive a partial result based on an axiom that formalizes the effect of an elementary transformation where a portion of a unit interaction profile is merged with another unit, then we characterize the full result by considering a generalization of the axiom where are taken into account transformations of the data that combine elementary transformations. Our preliminary key axiom requires that an elementary transformation-this operation should reduce the overall segregation in proportion of the quota of the initial unit that is merged. The more general axiom instead assumes that each combinaton of elementary transformation has a fixed proportional effect on the reduction of the value of segregation.

Consider, for instance, a population that is partitioned into two groups of equal size, the "Reds" and "Greens", where every individual interacts with half of the remaining individuals. If each individual interacts with half of the Reds and half of the Greens then all units exhibit the same interaction profile, there is therefore no segregation. If instead every individual of the Reds interacts with all the Reds and exclusively with them, and analogously for the Greens then the population is analogous to two units that collect all the individuals that interact with a specific groups thereby leading to maximal segregation.

If a proportion $1 - \alpha$ of the unit of Reds, is joining the unit of Greens and shares proportionally its interaction links then segregation is reduced. Moreover according to our preliminary key axiom this reduction occurs in proportion $1 - \alpha$. In fact as $1 - \alpha$ tends to 1 the overall segregation should be eliminated because all the individuals will share the same average interaction profile. Our preliminary result will show that this property plays a crucial role for the characterization of the Gini Exposure index for a large class of distribution matrices representing interactions profiles.

More generally, by combining portions of all the units interaction profiles it is possible to reduce segregation, these mixture transformations occur if the initial distribution of interactions profiles are multiplied by row stochastic matrices. Our main axiom will require that these general transformations do not affect the ranking in terms of segregation of two different configurations. Making use of a formalization of this condition and applying a decomposition property we derive a full characterization of the multi-group Gini Exposure index for all the matrices representing the interaction profiles in a society.

We use census tables and the American Community Survey to recover spatial interaction patterns between social groups in US metropolitan areas. We use a spatial model to identify interaction probabilities across neighborhoods of these cities within almost four decades. Our main assumption is that the chances for two individuals to interact decrease with the spatial distance between the area where the two individuals reside. Within the same urban context, spatial distance has been shown to be strongly correlated with other socio-economic variables that are relevant for defining interaction patterns (Conley and Topa 2002). Our focus will be on the ethnic interaction. We partition the resident population of the city in groups defined on the basis of observable attributes, and for each individual in each group we attach an interaction profile depending on the neighborhood where this people live. We then compute the Gini Exposure index for all American metro areas in 1980, 1990, 2000 and for each year in the 2007-2015 period. We use these estimates to assess the empirical correlations of various indices of segregation and the Gini Exposure segregation index. We then decompose segregation to assess how interaction profiles are unequally distributed among social groups defined along the dimensions of age, labor market attachment, profession, education and labor income. Finally, we merge our data with information on city attributes in (Glaeser, Resseger and Tobio 2009) to evaluate potential mechanisms explaining interaction dimension of segregation in major US cities.

The paper develops as follows. In Section 2 we provide the main notation. The Gini Exposure index is formally introduced in Section 3 and characterized in Section 4. Empirical results are discussed in Section 5. Section 6 concludes the paper. All proofs are collected in the appendix.

2 Notation

In this paper we consider the problem of ranking *configurations* $A, B \in \mathcal{C}(G)$ according to the level of segregation in the exposure dimensions that they exhibit.

Definition 1 (Configuration) *A configuration $A \in \mathcal{C}(G)$ is a triplet*

$$\Big[\mathcal{N}(A),\ \mathcal{G},\ ((\pi_{gi}(A))_{g\in\mathcal{G}}, \xi_i(A))_{i\in\mathcal{N}(A)}\Big]$$

where $\mathcal{N}(A)$ is a finite, non-empty set of units of cardinality $N(A)$, $\mathcal{G}$ is a finite, non-empty set of G population groups, with variable demographic size denoted by $N_g(A)$. For each unit $i \in \mathcal{N}(A)$ and group $g \in \mathcal{G}$, the variable $\pi_{gi} \in [0,1]$ denotes the probability that i interacts with a randomly selected individual from group g. Unit i's demographic weight is denoted by $\xi_i(A)$, with $\sum_{i\in\mathcal{N}(A)} \xi_i(A) = 1$.[1]

To avoid cumbersome notation, references to the configuration A are dropped in what follows, unless disambiguation is needed. Thus, we denoted π_{ig}, N, N_g and ξ_i for a generic configuration in $\mathcal{C}(G)$.

A configuration can be constructed, for instance, from empirical observation of the social connections between individuals, or from aggregate statistics of expected interaction patters. For a configuration the interaction profile of unit i is a column vector:

$$\boldsymbol{\pi}_{.i} := (\pi_{1i}, \dots, \pi_{Gi})^t \in [0,1]^G,$$

[1] One particular case is the uniform weighting scheme, where $\xi_i(A) = 1/N(A)$ for all $i \in \mathcal{N}(A)$.

such that $\sum_{g\in\mathcal{G}} \pi_{gi} = 1$ for any $i \in \mathcal{N}$. Hence, $\boldsymbol{\pi}_{.i}$ represents the social ties of unit i in terms of the probabilities that the individuals associated with this unit have to interact with members of each of the groups in $\mathcal{G}$. The $G \times N$ *interaction matrix* $\boldsymbol{\pi}$ represents a collection of the N interaction profiles (by column). The rows of the interaction matrix are denoted *group profiles* and are indicated with row vectors $\boldsymbol{\pi}_{g.} := (\pi_{g1}, \dots, \pi_{gN}) \in [0,1]^N$.

The *expected interaction profile* associated with group g is the expected probability that a randomly drawn individual interacts with group g:

$$\pi^e_g(A) = \sum_{i\in\mathcal{N}(A)} \xi_i(A)\,\pi_{gi}(A).$$

Again, we write π^e_g in shorthand notation. For configuration A, we make use of expected interaction profiles to normalize the entries of the interaction matrix $\boldsymbol{\pi}$. This leads to define a $G \times N$ interaction matrix $\mathbf{A}$ (always denoted with boldface letters) such that $\mathbf{A} := (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots \mathbf{a}_N)$ where $a_{gi} := \frac{\pi_{gi}(A)}{\pi^e_g(A)}$.

3 The Gini Exposure index

3.1 The index

The *Gini inequality* index of a univariate income distribution, represented by the N-dimensional vector $\mathbf{x}$, is defined as the average income gap between any pair of realizations in the income distribution $\mathbf{x}$, scaled by the overall average income:

$$G(\mathbf{x}) \ := \ \frac{1}{2N^2\left(\sum_i x_i/N\right)} \sum_{i=1}^{N}\sum_{j=1}^{N} |x_i - x_j|.$$

Alternatively, the Gini index can be related to the Lorenz curve: it is equal to twice the area between the Lorenz curve and the diagonal, representing the equal distribution.[2] As illustrated by Shephard (1974), the overall area delimited by the Lorenz curve can be represented as the sum of the areas spanned by every pair of vectors $(x_i, 1)$ and $(x_j, 1)$, corresponding to the determinant of a 2×2 matrix formed by these vectors. The gap $|x_i - x_j|$ corresponds, in fact, to the determinant of these matrices. It follows that the Gini inequality index rewrites:[3]

$$G(\mathbf{x}) \ := \ \frac{1}{2} \sum_{\forall\{i,j\}\subseteq\{1,\dots,N\}} \frac{1}{N}\frac{1}{N} \left| \det \begin{pmatrix} x_i/\left(\sum_i x_i/N\right) & x_j/\left(\sum_i x_i/N\right) \\ 1 & 1 \end{pmatrix} \right|$$

[2]The Lorenz Zonotope of distribution is defined as the area between the Lorenz curve and its *dual*. It can be written as a Minkowski sum of line segments, hence its area equals the sum of the areas spanned by each pair of bi-dimensional vectors $\left(\frac{x_i}{\sum x_i}, \frac{1}{N}\right)$ and $\left(\frac{x_j}{\sum x_i}, \frac{1}{N}\right)$, for all i, j. This area coincides with a parallelogram and it corresponds to a measure of inequality between incomes shares received by two individuals i, j equally weighted $\frac{1}{N}$ in the population.

[3]The terms $\frac{1}{N^2\sum_i x_i/N}$ disappears at it is incorporated in the determinant calculation. The comparison is now expressed in relative, rather than absolute, incomes. Moreover, the determinant is a measure of linear dependence, and therefore similarity, between oriented vectors.

In practice, the Gini inequality index can be conceptualized as a weighted average of the dissimilarity between the incomes of pairs of units and the two units' weights. The function measuring the intensity of this dissimilarity is the determinant, while the weights corresponds to the probability of drawing the pair of units i and j from the sample. Since every pair of incomes can be compared twice, the index must be standardized by two, so that its maximum is equal to one.

A similar logic can be adapted to the measurement of the degree of dissimilarity in interaction profiles, where income realizations have to be replaced by probabilities of interaction. Segregation assessments boil down to check how much dissimilar is each group's interaction probability deviation from the mean when compared to the value of 1. A configuration exhibiting no segregation corresponds to the case where i's interaction profile with group g is such that $\pi_{gi} = \pi_g^e$, for every $g \in \mathcal{G}$.

An obvious extension of the Gini inequality index is the *expected Gini* (EG) segregation index analyzed in Flückiger and Silber (1999) and Alonso-Villar and del Rio (2010). The EG index is an average of local Gini indices G_g, weighted by groups size:

$$EG(A) \ := \ \sum_{g\in G} s_g \, G_g(A),$$

where $s_g = \frac{N_g}{N}$ is the share of individuals in the network associated to group g. Each local Gini index is meant to capture the inequality in the distribution of interaction probability with a group, say g, across the population:

$$G_g(A) \ := \ \frac{1}{2} \sum_{\forall\{i_1,i_2\}\subseteq\mathcal{N}(A)} \xi_{i_1}(A)\xi_{i_2}(A) \left|\det\begin{pmatrix} \frac{\pi_{gi_1}(A)}{\pi_g^e(A)]} & \frac{\pi_{gi_2}(A)}{\pi_g^e(A)]} \\ 1 & 1 \end{pmatrix}\right|.$$

The expected Gini index assumes that evaluations of segregation can be separated across dimensions. This strong assumption leaves aside concerns about the composition of the interaction profiles. To overcome these limitations, following Andreoli (2014), we present a multi-group extension of the local Gini index presented above, denoted the *Gini Exposure* index of segregation.

The Gini Exposure index of segregation, $G_E : \mathcal{C}(G) \to [0,1]$ captures the dispersion in the normalized interaction profiles across units in the same configuration. The index is a weighted mean of a measure of dissimilarity between G-tuples of interaction profiles, as captured by the determinant of a square $G \times G$ matrix. The weight attached to each G-tuple corresponds to its probability of being observed. The index is standardized by $G!$, the overall number of possible G-tuples, so that the index maximum is equal to one.

Definition 2 (The Gini Exposure segregation index)

$$G_E(A) := \frac{1}{G!} \sum_{\forall\{i_1,\ldots,i_G\}\subseteq\mathcal{N}(A)} \xi_{i_1}(A)\cdot\ldots\cdot\xi_{i_G}(A) \left|\det\begin{pmatrix} \mathbf{a}_{i_1} & \ldots & \mathbf{a}_{i_G} \end{pmatrix}\right|.$$

3.2 A geometric illustration of the index

An equivalent way of assessing heterogeneity in interaction profiles consists in looking at the likelihood of any randomly chosen individual from group g to interact with unit i, given the original information about the distribution of interaction profiles across the population. For configuration A, we define the *interaction likelihood* $\ell_{gi}^A \in [0,1]$ as this probability. The sequence of probabilities $\ell_{g1}^A, \ldots, \ell_{gN}^A$ defines a distribution of *interaction likelihoods* of group g with all the units in the distribution, hence satisfying $\sum_{i\in\mathcal{N}} \ell_{gi}^A = 1$ for every $g \in \mathcal{G}$. The interaction likelihood is tied to interaction profiles and individual weights through the Bayes' rule:

$$\ell_{gi}^A \;:=\; a_{gi}\xi_i \;=\; \frac{\pi_{gi}}{\pi_g^e}\xi_i.$$

where the units' weights $\xi_i \in [0,1]$ are such that $\sum_i \xi_i = 1$. Heterogeneity in interaction profiles always implies that a form of dissimilarity between interaction likelihoods prevails. When all interaction profiles coincide, then $a_{gi} = 1$ and $\ell_{gi}^A = \xi_i$ for any i and g, meaning that the interaction likelihoods coincide across groups. This does not necessary imply, however, that the interaction likelihoods are constant across individuals. Conversely, when each individual interacts with exactly one group, say g, then the knowledge of the group allows to infer with certainty the individuals that will interact with it, because $\ell_{gi}^A > \ell_{g'i}^A = 0$ for all $g' \neq g$. All in-between situations display some form of dissimilarity between the rows of the interaction likelihood matrix $\mathcal{L}^A$ associated with configuration A and defined as:

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

where $\mathcal{L}^A$ is a *row stochastic matrix* (i.e. the entries add up to one by row, but not necessarily by column) of the type analyzed in Andreoli and Zoli (2014).

An illustration of this procedure is provided in the Appendix.

Andreoli and Zoli show that the dissimilarity between the rows of a $G\times N$ row stochastic matrix (depicting sets of G discrete probabilities distributions defines over n classes of realizations) can be visually represented through the *Zonotope set* Z of the interaction likelihood matrix $\mathcal{L}^A$, denoted $Z(\mathcal{L}^A)$. The Zonotope is a centrally symmetric polytope in the G-dimensional space representing the Minkowski sum of the matrix's columns (see Shephard 1974). More formally, it is defined as:

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

A Zonotope can be seen as a multi-group extension of the *segregation curve*[4], where the actual distributions of groups across organizational units are replaced by the interaction

[4]Segregation curves were introduced by (Duncan and Duncan 1955) and further studied by (Hutchens 1991) and generalized by (Carrington and Troske 1997) and (Silber 1989).

likelihoods these groups have with individuals. Since the Zonotope represents the extent of dissimilarity across interaction likelihoods, the order of distribution matrices produced by Zonotopes inclusion is always consistent with decreasing segregation.

Various works in linear algebra have studied the properties of the volume of the Zonotopes (McMullen 1971, Shephard 1974). It is shown, in particular, that the volume of any Zonotope of a $G \times N$ matrix can be written as the sum of the volumes of the Zonotope sets generated by every $G \times G$ (thus square) matrix obtained from the original one by considering distinct G-tuples of its columns. The volume of a square matrix is the absolute value of its determinant, as already noted in Koshevoy and Mosler (1997) and Arnold (2005). In our case, since the reference matrix is $\mathcal{L}^A$, the volume of $Z(\mathcal{L}^A)$ is the Gini-Exposure index of segregation. This immediately bears the following implication.

Remark 1 *For any* $A, B \in \mathcal{C}(G)$*,* $Z(\mathcal{L}^B) \subseteq Z(\mathcal{L}^A) \Rightarrow G_E(B) \leq G_E(A)$*.*

This remark is important for two reasons. First, it shows that the analysis of the exposure dimension of segregation is associated with the analysis of dissimilarity between distributions, in this case consisting in interaction likelihoods of interactions. It then follows that the exposure dimension of segregation can be studied by making use of methods developed in the context of dissimilarity analysis (Andreoli and Zoli 2014). The robust dissimilarity test based on Zonotopes inclusion, for instance, defines sufficient conditions for decreasing dissimilarity as picked up by the Gini-Exposure index.

Second, the result that characterize the Zonotope inclusion partial order in (Andreoli and Zoli 2014) can provide guidance on what axiom could be used as basic requirements in the characterization of the Exposure index.

4 Characterization

In this section, we study the minimal transformations of the data that allow to *characterize the Gini Exposure index* as a measure of segregation for a set of configurations of interactions. We use the term *segregation* to indicate any departure from the situation where interaction profiles are equalized within the network.

For expositional purposes we consider likelihood matrices of dimension $G \times N$ where $N \geq G \geq 2$. Each matrix is by construction row-stochastic, and represents configurations where in each column there exists at least a positive element. The set of such matrices in denoted $\mathcal{M}_{GN}$, where $\mathcal{M}_{GG}$ is the subset of $\mathcal{M}_{GN}$ containing only square row-stochastic matrices of dimension G that do not include empty units (i.e. columns with all elements equal to 0). For technical purposes we will consider also an extended set $\mathcal{M}^0_{GG}$ that will include matrices belonging to $\mathcal{M}_{GG}$ and those 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.

Let $\mathcal{L}^A \in \mathcal{M}_{GN}$, denote a *likelihood matrix* obtained from configuration A, where its generic element $\ell_{gj} \in [0, 1]$ represents the probability that an individual in group g interacts with individuals associated with unit j.

Given matrix $\mathcal{L}\in\mathcal{M}_{GN}$, we measure its exposure dimension segregation through the index $E^N(\mathcal{L})$, where $E^N : \mathcal{M}_{GN} \to [0,1]$ denotes a sequence of *continuous functions* from the set $\mathcal{M}_{GN}$ to the interval $[0,1]$. The index is increasing in the degree of segregation exhibited by a likelihood matrix and reaches its maximum value at 1.

We illustrate here some properties that should be satisfied by the E^N index.

Let Π_N denote a $N \times N$ permutation matrix. The set of all these matrices is $\mathcal{P}_N$. The property of *Units Anonymity* requires that the index is invariant with respect to permutations of the units (columns) of matrix $\mathcal{L}$.

Axiom 1 (UA: Units Anonymity) $E^N(\mathcal{L}) = E^N(\mathcal{L}\Pi_N)$ *for all* $\mathcal{L}\in\mathcal{M}_{GN}$, *all* $\Pi_N\in\mathcal{P}_N$.

The UA axiom can be interpreted equivalently also in terms of interaction matrices $\boldsymbol{\pi}$. It requires that segregation in exposure is not affected by columns permutations of $\boldsymbol{\pi}$. In this case also the weights of the units ξ should be permuted accordingly.

Next, the *Normalization* axiom identifies the reference case of maximal segregation. It is specified only for matrices in $\mathcal{M}_{GG}$. Let I_G denote the identity matrix of dimension G. When each unit is associated with only a group then the segregation is maximal and the index reaches the value of 1.

Axiom 2 (N: Normalization) $E^G(I_G) = 1$.

In terms of interaction matrices the maximal segregation is also associated with the case where the matrix $\boldsymbol{\pi}$ is an identity matrix.

In order to make segregation comparisons for matrices where $N > G$ we adopt a decomposition property. This property assumes that overall segregation evaluations can be based on a weighted combination of evaluations applied to square likelihood matrices in $\mathcal{M}_{GG}$ that are obtained by focussing only on G units ordered as in $\mathcal{L}$. According to this view the set $\mathcal{M}_{GG}$ is the minimal set of matrices that allow to "fully" express segregation evaluations. This is the smaller set of matrices where each unit could interact with only one group and all groups interact with at least one unit. This view is consistent with the N axiom that specify the reference case for maximal segregation in terms of matrices in $\mathcal{M}_{GG}$.

Let $\mathcal{N}_O(\mathcal{L})$ denote the set of *units ordered* according to the ranking in $\mathcal{L}$. Any G dimensional subset 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 $\mathcal{M}^0_{GG}$. 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 probability that an individual from group g interacts with one of the units in the set $\{i_1, i_2, ..., i_G\}$. The joint probability of interaction obtained taking into account all G groups 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 $\mathcal{M}_{GG}^0$ weighted according to the joint probability of interaction.

Axiom 3 (D: Decomposition)

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

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

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

Next axiom will allow to quantify changes in segregation moving from one matrix to another. To introduce it we consider a combination of two operations that should reduce the segregation for matrices in $\mathcal{M}_{GG}$ and we also quantify the cardinal level of this reduction. We first provide an intuition of these operations for interaction matrices $\boldsymbol{\pi}$ and then we express them in terms of transformations applied to likelihood matrices in $\mathcal{M}_{GG}$.

Consider a generic interaction matrix $\boldsymbol{\pi}$ and two units characterized by the column vectors of interactions $\boldsymbol{\pi}_{.i}$ and $\boldsymbol{\pi}_{.j}$, with demographic weights ξ_i and ξ_j respectively. Assume now that a proportion $0 \leq (1-\alpha) < 1$ of the individuals in unit i is joining unit j and shares its interaction probabilities. The new demographic weights then become respectively $\alpha\xi_i$ and $\xi_j + (1-\alpha)\xi_i$, the column vectors of interactions $\boldsymbol{\pi}_{.i}$ is unaffected but the one of group j is modified and is given by the weighted "mixture" of interaction probabilities of the merged proportions of units, it becomes

$$\boldsymbol{\pi}'_{.j} = (\xi_j\boldsymbol{\pi}_{.j} + (1-\alpha)\xi_i\boldsymbol{\pi}_{.i})\left(\frac{1}{\xi_j + (1-\alpha)\xi_i}\right).$$

The axiom of "Exposure segregation reduction through Mixtures of units" postulates that these operations should not increase segregation, and more precisely also assumes that segregation should be reduced proportionally according to the coefficient $(1-\alpha)$. Thus, if the original configuration is associated to a positive level of segregation these operations should strictly reduce it proportion $(1-\alpha)$.

We formalize now the axiom in terms of matrices in $\mathcal{M}_{GG}$. This notation will allow us also to highlight more directly the connections of the property and the Gini measures of inequality.

Consider a likelihood matrix $\mathcal{L} = (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_G) \in \mathcal{M}_{GG}$ represented by making explicit the G column vectors $\boldsymbol{\ell}_i$. This matrix can be transformed by taking a portion $(1-\alpha)$ where $0 < \alpha \leq 1$ of unit i and transferring it so to merge it with unit j, such that the resulting matrix $\mathcal{L}(\alpha, i, j)$ can be written as

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

We call such transformations *elementary "mixture of units"*. Note that the matrix $\mathcal{L}(\alpha, i, j)$ is still in $\mathcal{M}_{GG}$ by construction. If $\boldsymbol{\ell}_j$ and $\boldsymbol{\ell}_i$ are not linearly dependent this operation should not increase the degree of segregation. The fact that $E^G(\mathcal{L}(\alpha, i, j)) \leq E^G(\mathcal{L})$ can be explained by combining two operations [split and merge of columns] discussed in Andreoli and Zoli (2014) and in the literature on dissimilarity measurement with permutable columns. On one hand the split of a unit into two units is supposed to keep dissimilarity/segregation unchanged, on the other hand the merge of two units that are linearly independent is supposed not to increase the dissimilarity/segregation, by leveling the disparities between these two units. Next axiom is also quantifying the reduction in segregation.

Axiom 4 (EM: Exposure segregation reduction through Mixtures of units) $E^G(\mathcal{L}(\alpha, i, j)) = \alpha E^G(\mathcal{L})$ *for all* $\mathcal{L} \in \mathcal{M}_{GG}$*,* $i, j \in \mathcal{N}$*,* $0 < \alpha \leq 1$.

According to axiom EM any elementary mixture of units that involves a transfer of a portion $(1-\alpha)$ of a unit connection links is reducing segregation by the proportion $(1-\alpha)$.

Axiom EM is a multidimensional generalization of a property satisfied by the Gini index when $G = 2$.

Consider for instance the Gini index derived from the matrix

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

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 $p - x$ as could be calculated by computing the Lorenz 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 L by $\begin{bmatrix} \alpha & 1-\alpha \\ 0 & 1 \end{bmatrix}$ the obtained new matrix $L(\alpha, 1, 2)$ will be

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

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

An analogous result could be obtained by post multiplying 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

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

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

Moreover, the EM property can be seen as a result of a combination of two basic properties used in (Andreoli and Zoli 2014) to characterize the dissimilarity partial orders.

The first property is the invariance with respect to operations of splitting of units. This property postulates that if one unit say unit i with likelihood $\boldsymbol{\ell}_i$ is split into two units i_1 and i_2 with respectively likelihood profiles $\alpha\boldsymbol{\ell}_i$ and $(1-\alpha)\boldsymbol{\ell}_i$ then dissimilarity is not affected. This implies that segregation is neutral with respect to partition of units as long as the units likelihood profiles are proportional to each other.

Consider matrix $\mathcal{L}\in\mathcal{M}_{GN}$ such that $\mathcal{L} = (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, ..., \boldsymbol{\ell}_i, .., \boldsymbol{\ell}_j, ..., \boldsymbol{\ell}_N)$ formally this condition requires that:

Axiom 5 (EIS: Exposure segregation Invariance from Splitting of units)

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

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

Note that the new obtained likelihood matrix is of dimension $G \times N + 1$.

The formulation of Axiom EM requires that one of the split units obtained from unit i is merged with unit j, thereby leading to matrix $\mathcal{L}(\alpha, i, j)$. If the original matrix $\mathcal{L}$ is in $\mathcal{M}_{GG}$ also matrix $\mathcal{L}(\alpha, i, j)$ is in $\mathcal{M}_{GG}$. This merge operations between units is supposed to not increase dissimilarity (and therefore not increase segregation). The reason is that by combining together two units likelihood vector this operation reduces the disparities among them and therefore reduces the dissimilarity.

Axiom EM also quantifies the effect of this segregation/dissimilarity reduction.

We are now ready to prove the first characterization result for the Gini Exposure index.

4.1 A first characterization result

We apply our result to a set $\hat{\mathcal{M}}_{GG} \subseteq \mathcal{M}_{GG}$ of square matrices of dimension G that could be obtained as a combination of elementary mixture of units/columns operations and permutations of units/columns. 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 likelihood matrices. Making use of an equivalence result in Theorem 1 in Andreoli and Zoli (2014) it can be shown that 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.[5] 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.

Proposition 1 *Let $\mathcal{L}\in\hat{\mathcal{M}}_{GG}$, the exposure segregation index $E^G : \mathcal{M}_{GG} \to [0,1]$ satisfies axioms UA, N and EM if and only if it is the absolute value of the determinant of $\mathcal{L}$, that is*

$$E^G(\mathcal{L}) := |\det \mathcal{L}| .$$

Proof. See Appendix B.1. ■

Before moving to the extension of the result to matrices where $N > G$ it is important to highlight the relevance of the restrictions applied in the result in Proposition 1, that holds for matrices in $\hat{\mathcal{M}}_{GG}$. As already 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 (1) 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 $G \geq 3$.[6]

The result in Proposition 1 goes beyond the simple proof of the sufficiency part, it shows also the necessity condition for the characterization that holds for a potentially large set of admissible matrices of interest.

Note that $E^G(\mathcal{L}) = |\det \mathcal{L}|$ is the *volume of the Zonotope* associated with matrix $\mathcal{L}\in\hat{\mathcal{M}}_{GG}$. By construction the index 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

[5] This statement is obtained by considering the equivalence between statements (iv) and (i) in Theorem 1 in Andreoli and Zoli (2014) and adapting it to the notation adopted here. The result is more general as the one restated here, in fact it holds for matrices B and A where $B = AX$ for a generic row stochastic matrix X such that A and B could exhibit a different number of columns. It involves the possibility of adding or eliminating empty classes, the use of operations of splitting of columns and merging of columns and columns permutations. In the current setting it suffices to consider $A = I_G$.

[6] 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 elelmentary mixture operations and permutation of colums that allow to reach such matrix starting from an identity matrix. Thus such a matrix is not in $\hat{\mathcal{M}}_{GG}$.

if one restricts attention to problems with G groups and G units.[7] However, as we are going to show by making use of axiom D, when $N > G$ the overall segregation measure boils down to 0 only if there are not 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 two rows of $\mathcal{L}$ are identical.

We derive now the general formula for the exposure segregation index for matrices where $N > G$. Given the construction of Proposition 2, we consider the set of matrices $\hat{\mathcal{M}}_{GN} \subseteq \mathcal{M}_{GN}$, such for any matrix in $\hat{\mathcal{M}}_{GN}$ any of its square submatrices, where all rows have at least one positive element, that are obtained after eliminating $N - G$ columns is in $\hat{\mathcal{M}}_{GG}$. By direct application of axiom D in conjunction with the result of Proposition 1 it follows that:

Corollary 1 *Let $\mathcal{L}\in\hat{\mathcal{M}}_{GN}$, the exposure segregation index $E^N : \mathcal{M}_{GN} \to [0,1]$ satisfies axioms UA, N, EM and D if and only if it is the Gini Exposure index, that is*

$$E^N(\mathcal{L}) := \frac{1}{G!} \sum_{\{i_1,i_2,\ldots,i_G\}\subseteq\mathcal{N}(\mathcal{L})} \left|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, \ldots, \boldsymbol{\ell}_{i_G})\right|.$$

Proof. See Appendix B.2. ■

Note that the obtained index satisfies as a by-product also axiom EIS. This claim can be verified directly considering that the Gini exposure index can be computed as the volume of the Zonotope associated with the likelihood matrix $\mathcal{L}$, and that the Zonotope is not affected by splitting of columns of $\mathcal{L}$ as done in the definition of EISU.

Next remark formalizes the fact that when two groups are considered the result applies to all matrices in $\mathcal{M}_{2N}$. In this case Proposition 1 and the related corollary provide a full characterization of the Gini segregation index in the *two groups case*.

Remark 2 $\hat{\mathcal{M}}_{22} = \mathcal{M}_{22}$.

Proof. See Appendix B.3. ■

4.2 A more general characterization result

Despite the intuitive appeal of the result in Proposition 1 and Corollary 1, the full application is limited to the two groups case or to a subset of interaction matrices for the case with more groups. Making use of the same basic concepts it is however possible to derive a more general result that lead to the characterization of the index that holds for all matrices in $\mathcal{M}_{GN}$.

The key axiom considered here is a generalization of EM that takes into account more general segregation reducing transformations of matrix $\mathcal{L}$ than the *elementary "mixture of units"*.

[7] These concerns are in line with those expressed in Shorrocks (1978) for the measures of intergenerational mobility that take into account the determinant of the square transition mobility matrix.

Consider matrix $\mathcal{L}\in\mathcal{M}_{GG}$ and assume that each unit i of the matrix is split into G units with splitting coefficients $x_{ij} \in [0,1]$ such that $\sum_j x_{ij} = 1$, with $j = 1, 2, ..., G$. For instance the vector $\boldsymbol{\ell}_i$ is split into G vectors $\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G}$ where $\boldsymbol{\ell}_{i_j} = x_{ij} \cdot \boldsymbol{\ell}_i$. Then merge all these new vectors combining for all $j = 1, 2, ..., G$ each vector $\boldsymbol{\ell}_{i_j}$ that is obtaining a new vector $\boldsymbol{\ell}'_j = \boldsymbol{\ell}_{1_j} + \boldsymbol{\ell}_{2_j} + ... + \boldsymbol{\ell}_{G_j}$. This operation of merges of vectors leads to a new matrix $\mathcal{L}'\in\mathcal{M}_{GG}$ such that $\boldsymbol{\ell}'_j = \sum_i \boldsymbol{\ell}_{i_j} = \sum_i x_{ij} \cdot \boldsymbol{\ell}_i$. In compact notation we obtain

$$\mathcal{L}' = \mathcal{L}X$$

where X is a row stochastic matrix in $\mathcal{M}_{GG}$ with generic element x_{ij}. Note that an *elementary "mixture of units"* can be formalized making use of a specific matrix X.[8] Each one of the split operations used to construct matrix X is not affecting the overall level of segregation (see EIS), while each of the merge operations that combine all obtained split vectors is not-increasing the level of segregation. As a result $\mathcal{L}'$ should be at most as segregated/dissimilar as $\mathcal{L}$, that is $E^G(\mathcal{L}') \leq E^G(\mathcal{L})$.

The next general version of axiom EM formalizes the change in segregation generated by the transformation X. The axiom requires that any segregation reduction generated by the transformation X induces a proportional decrease in the overall level of segregation that is independent from $\mathcal{L}$ and formalized by a function $F^G(X)$ that could range in $[0,1]$. Here we formalize the GEM (Generalized Exposure segregation reduction through Mixtures of units) axiom.

Axiom 6 (GEM: Generalized EM) *For all $\mathcal{L}$,$X\in\mathcal{M}_{GG}$, there exists $F^G : \mathcal{M}_{GG} \to [0,1]$ such that*

$$E^G(\mathcal{L}X) = E^G(\mathcal{L}) \cdot F^G(X).$$

An implication of the GEM property is that if two distributions $\mathcal{L}$ and $\mathcal{L}''$ exhibit the same level of exposure segregation, then this should be the case also if they are both transformed through the same segregation levelling transformation X. That is it is always the case that for all $\mathcal{L}$,$\mathcal{L}''$,$X\in\mathcal{M}_{GG}$ then $E^G(\mathcal{L}) \leq E^G(\mathcal{L}'')$ is equivalent to $E^G(\mathcal{L}X) \leq E^G(\mathcal{L}''X)$.

Axiom GEM generalizes EM in two directions, first the *elementary "mixture of units"* transformations are special cases of those obtained making use of matrix X, and then secondly axiom EM requires that $F^G(X) = \alpha$ whenever the matrix X formalizes an *elementary "mixture of units".* The next theorem formalizes the general characterization of the Gini exposure segregation index for all distributions in $\mathcal{M}_{GN}$.

Theorem 1 *Let $\mathcal{L}\in\mathcal{M}_{GN}$, the exposure segregation index $E^N : \mathcal{M}_{GN} \to [0,1]$ satisfies axioms N, UA, GEM, EIS, and D if and only if it is the Gini Exposure index, that is*

$$E^N(\mathcal{L}) := \frac{1}{G!} \sum_{\{i_1,i_2,...,i_G\}\subseteq\mathcal{N}(\mathcal{L})} \left|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})\right|.$$

Proof. See Appendix B.4. ■

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

5 Empirical analysis of the exposure dimension of segregation among US metropolitan areas: 1980-2015

5.1 Data

We explore information on demographic composition within U.S. metropolitan areas over four decades, drawing on the census files of the U.S. Census Bureau for 1980, 1990 and 2000, to investigate patters in the ethnic dimension of segregation. We consider pattern of individual interaction with four ethnic groups, as identified in the American census: White, Black, Asiatic and Hispanic origin. These groups have different demographic weights in different cities, implying differences in the underlying expected interaction profiles. The implications of these differences are born out using the Gini Exposure index, which is a relative measure.

Data on the ethnic composition are from the decennial census Summary Tape File 3A.[9] Due to anonymization issues, the STF 3A data are given in the form of statistical tables representative at the census tract level. Each table reports, for a given geography, information about population counts by ethnic origin. After 2000, the count statistics reported on the STF 3A files have been replaced with survey-based estimates from the American Community Survey (ACS), which runs annually since 2005 on representative samples of the U.S. resident population. We focus on the 2005-2015 5-years Estimates ACS modules. Sampling rates in ACS vary independently at the census tract level according to 2010 census population counts, covering on average 2% of the U.S. population over the 5 years. As far as we know, the ACS 2011-2015 wave has not yet been used for empirical analysis of urban inequality.

We consider all 381 US Metropolitan Statistical Areas (MSA) identified by the Census Bureau and explore the census tract division of their territory provided in 2010. We reconstruct the neighborhood composition across the ACS and census waves using appropriate cross-walk files. All tracts are georeferenced, and measures of distance between the tracts centroids can therefore be constructed. All observations within the same block group are assumed to occur on its centroid. We use the census tracts as the reference partition to identify the interaction probabilities π_{gi}, under the assumption that all residents in the same tract share the same vector of interaction probabilities. Interaction probabilities and *expected* (by city and year) probabilities are calculated according to their definitions. We use the relative demographic size of a tract i in a given MSA to infer ξ_i.
The distribution of interaction profiles across tracts within the same MSA in each year defines the object of our study.

[9]The Census STF 3A provides cross-sectional data for all U.S. States and their subareas in hierarchical sequence down to the census tract level. The geography of the census tracts partition changes over the decades to keep track with demographic changes within the Counties of each State.

5.2 Descriptives, trends and patterns

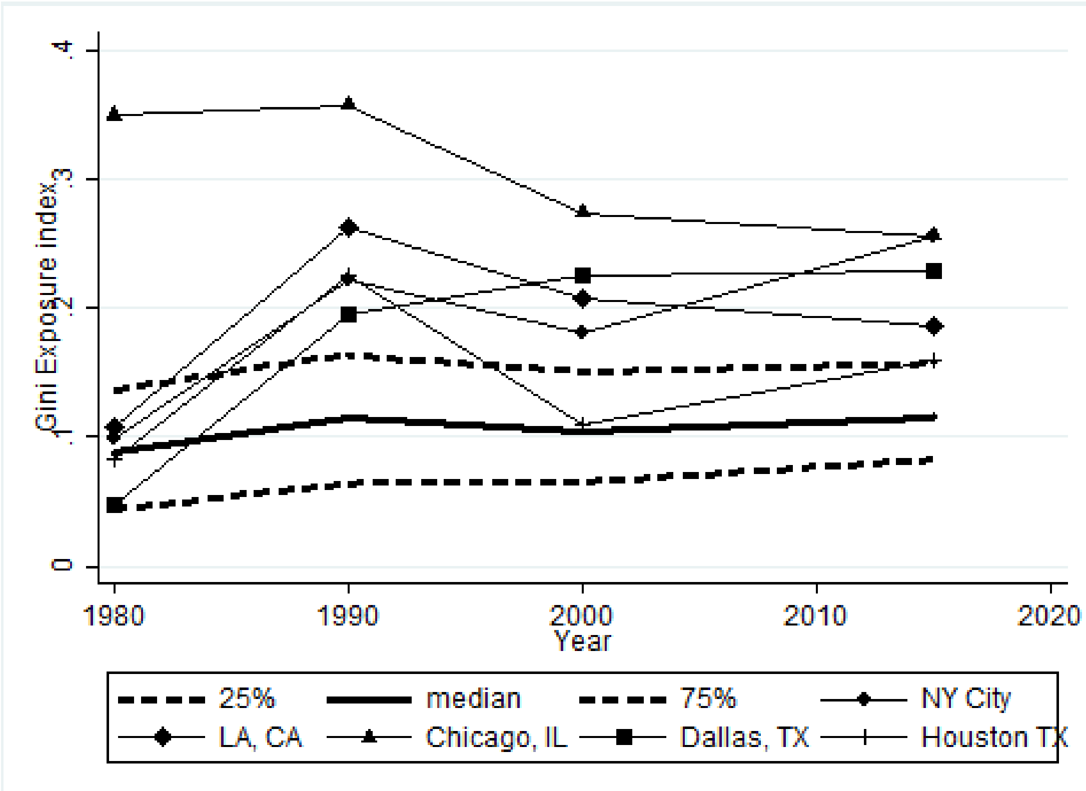
We produce estimates of the Gini Exposure index for all American metro areas in 1980, 1990, 2000 and 2015. Overall, we consider 381 metropolitan statistical areas in each year. The average MSA in 1980 has 0.651 million residents, with more than half of the sample of MSA displaying more than 0.250 millions residents. While the size of the median city has remained stable over 35 years, the size of the average metro area in the US has grown to 0.701 million inhabitants, mostly driven by increasing demographic growth of the largest cities.[10]

In 1980, on average, almost 77% of American residents in large urban areas are White, with just 5% of the American cities displaying more than 50% of non-White residents, and larger cities displaying a more mixed environment. These numbers have changed dramatically over the last four decades: in 2015, the White population has shrunk to 60% of urban residents on averages, and more than 25% of American metro areas now display a clear non-White majority of residents. This change is not associated with an expansion of the Black population (the average city in 1980 displayed 11.1% of Black residents versus 13.5% in 2015) or of the Asian population (whose share in American metro areas has grown from 4% to 6% over 1980-2015), but with a substantial increase in the Hispanic and Latino population (the Hispanic proportion of residents grew from 7% in 1980 to almost 20% in 2015). Along with this change, we also record from the data a major shift in the variability of the distribution of Hispanic population across the census tracts of the American cities, which is only matched, to a smaller extent, by a increase in variability in the Asian population. The patterns are however less strong in largest American metro areas: in these cities, the spatial concentration of the Hispanic and Asian population over the last four decades has substantially increased.

These figures suggest a changing panorama of social interaction based on place of residence in urban America. Nevertheless, they are scarcely informative about the patterns of interactions of residents in major cities with other White, Black, Asian and Hispanic residents. In fact, increasing concentration of residents of Hispanic origin might foster segregation in the exposure dimension if these people concentrate in areas that are isolated from the other groups, or it might lead to a more equitable distribution of interaction profiles if the patterns of concentration converge with historic patterns observed for other groups. We highlight patterns and trends of segregation in the exposure dimension through the Gini Exposure index computed for all 381 American metro areas in the years considered in this study. In Figure 1 we report patterns of the Gini Exposure index for the median, the bottom and the top quartile city in each year as ranked by the level of the Gini Exposure index in that year (patterns of these numbers are defined by bold solid and dashed lines in the graph). These numbers provide information on the distribution of exposure segregation in each year, while the lines provide information about the change in the distribution of exposure segregation across cities.

[10] Census tracts define the basic partition of the urban space. On average, a census tract gathers 4.417 residents in the sample of cities. These areas are constantly reshaped from the Census Bureau to guarantee a balanced demographic composition through time.

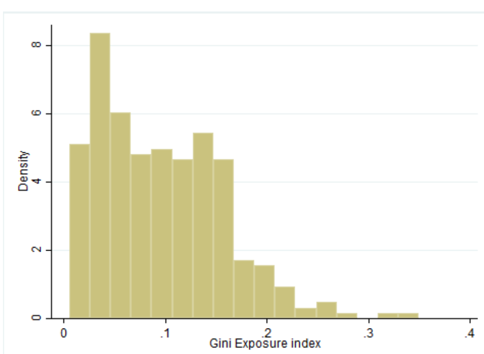
Figure 1: Gini Exposure dynamics



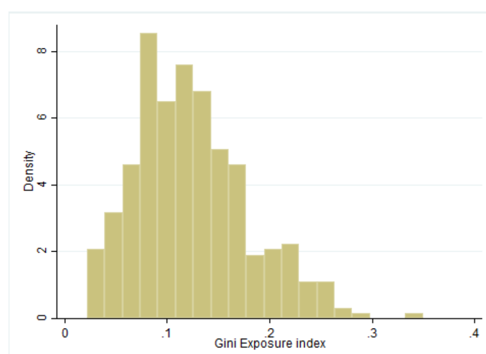
Note: Authors analysis of U.S. census and ACS data.

Figure 2: Gini Exposure across cities

[1980]



[2015]



Note: Authors analysis of U.S. census and ACS data.

We find evidence of a pattern of slow growth in the Gini Exposure over the period 1980-

2015. The increments in exposure segregation have been stronger during the Nineties, and have stabilized thereafter. The actual change in the patterns of exposure segregation across American cities is well illustrated in Figure 2, which reflects the starting and ending points of the lines in Figure 1. Despite changes in exposure segregation have been modest over the period considered, the figures highlight an increase in frequencies (i.e., in the number of metro areas) of values of the Gini Exposure index that are above 0.2, almost two times larger than the Country median level.

Most of the cities that have experienced a growth in the Gini Exposure index are among the largest metro areas in the U.S. In Figure 1 we report the dynamics of the Gini Exposure index for the five largest cities in 2015. Aside from Chicago, IL, which displays historically high levels of ethnic segregation that are well documented in the literature (White 1983, Reardon and Firebaugh 2002, Reardon and O'Sullivan 2004), the data highlight that exposure segregation has grown considerably in the other four cities. In 1980, NY City, Los Angeles, Dallas and Houston all displayed a level of exposure segregation matching the median segregation in the Country. In all these cities, the level of segregation peaked in 1990 and remained high afterwards. In 2015, all these cities can be located among the top quartile of American cities as ranked by the Gini Exposure index.

These patters are suggestive that, across the years, populations size and density, along with other characteristics, correlate with both measured segregation and with the changes in segregation. Figure 3 reports a cross section of American cities where levels and changes in the Gini Exposure index are reported for two periods: 1980-2015 in panel a) and 2000-2015 in panel b). Concerning the whole period under analysis (panel a)), we find evidence of convergence in exposure segregation: cities that display larger values of Gini Exposure in 1980 are those that, on average, display smaller changes across time. Patterns of growth in exposure segregation among cities that display little segregation in 1980, however, are substantially heterogeneous.

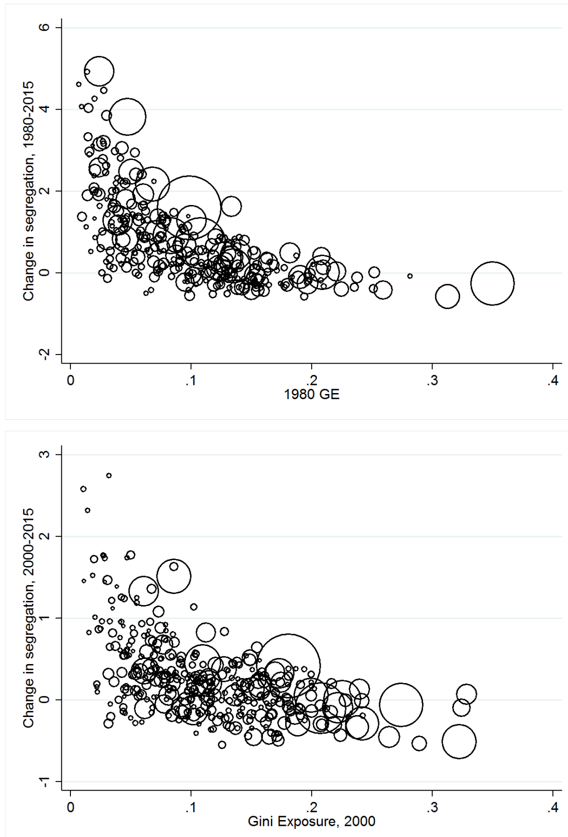
We find similar patterns of convergence when we focus on the period 2000-2015. This time frame is interesting because we can study how exposure segregation and its changes covary with demographic and economic characteristics of the inhabitant of the cities, and we can single out the partial correlation with components defining the individuals' decisions to sort across and within cities, thus determining their profiles of interactions with the ethnic groups we consider. The next section describes these correlations.

5.3 Covariates of the interaction dimension of segregation

We use reduced form models to highlight the implications of covariates describing differences across cities in demographics (A) public goods spending (B) and educational quality (C). These observable are related to sorting decisions across cities. We also consider covariates that are instead connected to the sorting behavior of inhabitants in each city (D), that likely affect patterns of formations of networks of interactions across neighborhoods of the city. The covariates we consider are taken from Andreoli and Peluso (2017), where

Figure 3: Convergence in exposure

[1980-2015] [2000-2015]



Note: Authors analysis of U.S. census and ACS data.

one can find an in-depth description of the sources and the mothods used to construct such covariates.

In Table 1 (see appendix) we report the marginal effect of these covariates on the level of the Gini Exposure index in 2000, obtained from cross-MSA regressions. The scope is to assess the simultaneous effect of demographic, economic and local public good spending conditions on exposure segregation. In model (1) and (2) we show that the demographic conditions of the city (racial composition, population density) are all strongly correlated with variability of the Gini Exposure index across American cities. In particular, the size of the Asian population has a negative effect on exposure segregation, suggesting that patterns of interactions with all the ethnic groups are more equally distributed in those cities with larger fractions of the asian population. Large populations of Blacks and Hispanic contribute positively to rising exposure segregation. Segregation is also larger in cities displaying larger population density, where interactions are more likely to occur even at smaller geographic scales. The significance and sign of these correlations survive even when we additionally control for other covariates.

In model (3) we highlight that characteristics of local tax schedule and spending in public goods display little correlation with the Gini Exposure index, suggesting a minor role of local fiscal policies in smoothing exposure segregation of urban residents. Conversely, spending in education seems to robustly and significantly correlate with the Gini Exposure (model (4)). While public school budget seems to be positively associated with segregation (possibly reflecting a reverse effect of segregation on local spending that is activated via mechanisms related to ethnic voting and political participation), it seems that the way in which cities with similar budget decide to spend it on different items matters more for exposure segregation.

Cities with better public schools and larger spending in public kindergarten experience smaller exposure segregation, the effect being robust even after controlling for potential sources of sorting within the city. The effect survives even after accounting for heterogeneity in spending within the city, this being a possible source of sorting of different ethnic groups across neighborhoods and thus determining uneven patterns of interactions with other residents. In this case, we find that public schools of better quality which are more inclusive for the black population are associated with smaller exposure segregation. These effects are, however, offset by other dimensions that are strongly connected to the features of the urban income distribution (model (5)).

Variables indicating affordability of the city's neighborhoods, as well as income segregation, are positively associated with the Gini Exposure index, as it is the presence of violent crimes in the city, a measure connected to the poverty status and lack of economic opportunities of the inhabitants. Income inequality has also a significant positive partial effect on exposure segregation: income is an underlying dimension of heterogeneity across individuals that contributes positively to inequalities in the prospects of interactions with other racial and ethnic groups in the city. Surprisingly, we find evidence that inequality in the neighborhood as measured in Andreoli and Peluso (2017), rather than inequality in the city, has a powerful role in explaining this exposure segregation levels in 2000.

We correlate the same variables used in Table 1 with changes in the Gini Exposure index during the last decade. The marginal effects of these variables are displayed in models (1)-(6) of Table 2 (see appendix). We focus for simplicity on the most comprehensive model (6).

First, we consider the convergence dynamics of the Gini Exposure index by regressing changes in exposure segregation on the level of segregation measured in 2000. The effect is negative and significant even after introducing other controls. This is evidence that changes in exposure segregation strongly depend on historical patterns of exposure segregation in the city. Almost all demographic variables (A) have negative and significant effects on changes in segregation. We highlight, however, that most of their effects are offset by changes in migration flows: cities experiencing larger migration are also more likely to see exposure segregation rising. This is in line with migrant network theories, claiming that new immigrants tend to interact with well-established local communities of same-group individuals, developing networks that are advantageous for their employability chances but that contribute to raise inequalities in the interaction probabilities among long term residents.

Educational variables, related to sorting across and within metro areas, have negative, scarcely significant effects on changes in exposure segregation, and thus do not contribute substantially to its changes. We find that some features of the income distribution are highly correlated with the changes in exposure segregation. First, we find that cities that display larger poverty and higher average incomes, hinting towards an accessibility crisis in the housing market, have experienced larger increases in exposure segregation. The sign of the effect is hence consistent with these changes. Second, expanding inequality in the city, rather than at the neighborhood level, is associated with reductions in exposure segregation. This effect, however, largely reflects the implications of simulataneity biases and deserve a dedicated treatment.

6 Concluding remarks

We have proposed and analysed a new measure of multi-group segregation in networks: the Gini Exposure index. The index is designed to evaluate across the individuals in a network the inequality in the distribution of their interaction profiles with social groups. It can be interpreted as the volume of the Zonotope of the matrix of the likelihood probabilities of interaction with the social groups in analogy with the generalization of the inequality Gini index in the multidimensional setting provided by the volume Gini index. In order to highlight the properties of the Gini Exposure index we have presented an axiomatic characterization of the index that holds for a large set of interactions configurations.

We highlight the relevance of the index with an empirical application to the American Metropolitan Statistical Areas. We find high heterogeneity in the Gini Exposure index across American metro areas, although exposure segregation in most cities has remained substantially stable over the last 35 years. Larger cities have experienced larger growth in exposure segregation, which peaked in the Nineties and remained well above the level of exposure segregation of the median American city ever since.

We also study the determinants of exposure segregation. Although our reduced form estimates are far from identifying causal effects, we are able to detect interesting partial correlations that deserve further investigation: First, demographic characteristics of a city strongly and robustly correlate with exposure segregation and with its changes, immigration being an important driver; Second, the quality and widespread provision of public education significantly correlates with exposure segregation; Third, the shape of the local income distribution matters for exposure segregation. While inequality in the neighborhood strongly correlates with simultaneous segregation, past citywide poverty and inequality seem, instead, to play a dominant role in predicting its decennial changes.

References

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

Alonso-Villar, O. and 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. and Peluso, E. (2017). So close yet so unequal: Spatial inequality in American cities, *LISER, 2017, Working Papers n2017-11* .

Andreoli, F. and Zoli, C. (2014). Measuring dissimilarity, *Working Papers Series, Department of Economics, Univeristy of Verona, WP23* .

Arnold, B. C. (2005). Inequality measures for multivariate distributions, *Metron - International Journal of Statistics* **LXIII**(3): 317–327.

Carrington, W. J. and Troske, K. R. (1997). On measuring segregation in samples with small units, *Journal of Business & Economic Statistics* **15**(4): 402–409.
URL: `http://www.jstor.org/stable/1392486`

Conley, T. G. and Topa, G. (2002). Socio-economic distance and spatial patterns in unemployment, *Journal of Applied Econometrics* **17**(4): 303–327.
URL: `http://dx.doi.org/10.1002/jae.670`

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

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

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

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

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

Glaeser, E., Resseger, M. and Tobio, K. (2009). Inequality in cities, *Journal of Regional Science* **49**(4): 617–646.

Hutchens, R. M. (1991). Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations, *Mathematical Social Sciences* **21**(1): 31 – 51.
URL: 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.
URL: http://www.jstor.org/stable/3663531

Koshevoy, G. A. and Mosler, K. (1997). Multivariate Gini indices, *Journal of Multivariate Analysis* **60**(2): 252–276.
URL: http://ideas.repec.org/a/eee/jmvana/v60y1997i2p252-276.html

Koshevoy, G. and Mosler, K. (1996). The Lorenz zonoid of a multivariate distribution, *Journal of the American Statistical Association* **91**(434): 873–882.
URL: http://www.jstor.org/stable/2291682

Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation, *Social Forces* **67**(2): 281–315.
URL: http://www.jstor.org/stable/2579183

McMullen, P. (1971). On zonotopes, *Transactions of the American Mathematical Society* **159**: 91–109.

Reardon, S. F. and Firebaugh, G. (2002). Measures of multigroup segregation, *Sociological Methodology* **32**: 33–67.
URL: http://www.jstor.org/stable/3186153

Reardon, S. F. and O'Sullivan, D. (2004). Measures of spatial segregation, *Sociological Methodology* **34**(1): 121–162.
URL: http://dx.doi.org/10.1111/j.0081-1750.2004.00150.x

Shephard, G. C. (1974). Combinatorial properties of associated zonotopes, *Canadian Journal of Mathematics* **26**: 302–321.
URL: http://cms.math.ca/cjm/v26/cjm1974v26.0302-0321.pdf

Shorrocks, A. F. (1978). The measurement of mobility, *Econometrica* **46**(5): 1013–1024.
URL: http://www.jstor.org/stable/1911433

Silber, J. G. (1989). On the measurement of employment segregation, *Economics Letters* **30**(3): 237 – 243.
URL: http://www.sciencedirect.com/science/article/B6V84-45DMS77-7B/2/b62c3ef322470f5af58ce78f5d1df60c

White, M. J. (1983). The measurement of spatial segregation, *The American Journal of Sociology* **88**(5): pp. 1008–1018.
URL: http://www.jstor.org/stable/2779449

A Illustration

We introduce, with a simple example, the representation of the data that we use and the type of transformations involved in our analysis.

An interaction profile defines the conditional probability that a given population unit[11], denoted by i, interacts with each of the social groups, denoted by g, in which the population is partitioned. This probability is denoted by π_{gi}. Each unit is associated with its own interaction profile, that may depend on her network, location, or demographic attributes. In the example, we consider four individuals $l_1,\ l_2,\ j$ and k. Two of them belong to group g_1 and the remaining to group g_2. We assume that profiles of interactions are estimated on the individual bases, so that the units of analysis coincide with the individuals. The interaction profile of individual i specifies the probability s/he has to interact with groups g_1 and g_2. Let assume for simplicity that individuals l_1 and l_2 share the same interaction profile, which is marked with an l. We can reduce the analysis to three profiles. We use the following data to fix ideas:

$$\begin{pmatrix} \pi_{g_1 l} \\ \pi_{g_2 l} \end{pmatrix} = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}, \quad \begin{pmatrix} \pi_{g_1 j} \\ \pi_{g_2 j} \end{pmatrix} = \begin{pmatrix} 1/8 \\ 7/8 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \pi_{g_1 k} \\ \pi_{g_2 k} \end{pmatrix} = \begin{pmatrix} 3/8 \\ 5/8 \end{pmatrix}.$$

According to the first profile, the chances that l_1 or l_2 interact with a person of group g_1 is 25%, while 75% of the times they interact with members of group g_2.

To normalize the data and eliminate any form of heterogeneity *within* interaction profiles we use the vector of expected interaction probabilities π_g as the endogenously determined reference interaction profile.

It turns out that segregation can be measured as a form of *dissimilarity* (see Andreoli and Zoli 2014) between the likelihood that any randomly drawn individual of group g interacts with the demographic unit i, for any group g and any unit i. This likelihood, denoted ℓ_{gj}, should ideally equate the probability of interacting with unit i, namely $\Pr[i]$ if interaction profiles are equally distributed in the population. That is, $\ell_{gj} = \ell_{g'j}$ for all i's and all groups $g \neq g'$. Any departure from this configuration leads to a form of segregation in the exposure dimension.

The Bayes rule ties interaction probabilities to the likelihood of interaction in the following way:

$$\ell_{gj} = \frac{\Pr[i] \cdot \pi_{gi}}{\pi_g}.$$

In our example, suppose that weights are defined as follows: $\Pr[l] = 2/4$, $\Pr[j] = 1/4$

[11] A unit can be an individual, in which case it receives a weight equal to the inverse of the overall population size. It can also represent the minimum statistical unit used to empirically construct interaction profiles, for instance a class of student in a school, a neighborhood or a family unit. In this case, the weight of the unit is proportional to the group of individuals attached to that unit, and thus experiencing the same interaction profile

and $\Pr[k] = 1/4$. The expected interaction profile can be computed as follows:

$$\begin{pmatrix} \pi_{g_1} \\ \pi_{g_2} \end{pmatrix} = \frac{2}{4} \begin{pmatrix} \pi_{g_1 l} \\ \pi_{g_2 l} \end{pmatrix} + \frac{1}{4} \begin{pmatrix} \pi_{g_1 j} \\ \pi_{g_2 j} \end{pmatrix} + \frac{1}{4} \begin{pmatrix} \pi_{g_1 k} \\ \pi_{g_2 k} \end{pmatrix} = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}.$$

The interaction profiles are not equally distributed. In fact, one obtains:

$$\begin{pmatrix} \ell_{g_1 l} \\ \ell_{g_2 l} \end{pmatrix} = \begin{pmatrix} 2/4 \\ 2/4 \end{pmatrix}, \quad \begin{pmatrix} \ell_{g_1 j} \\ \ell_{g_2 j} \end{pmatrix} = \begin{pmatrix} 1/8 \\ 7/24 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \ell_{g_1 k} \\ \ell_{g_2 k} \end{pmatrix} = \begin{pmatrix} 3/8 \\ 5/24 \end{pmatrix},$$

or, in matrix notation,

$$\mathcal{L} := \begin{pmatrix} 2/4 & 1/8 & 3/8 \\ 2/4 & 7/24 & 5/24 \end{pmatrix},$$

which shows that the sources of exposure are units j and k, given that $\ell_{g_1 l} = \ell_{g_2 l}$ holds. In fact, unit l_1 and l_2 interaction profiles coincide with the expected profile.

B Proofs

B.1 Proof of Proposition 1

Proof. *Necessity part.* Consider matrix $\mathcal{L} \in \hat{\mathcal{M}}_{GG}$, by construction it could be obtained from I_G applying a finite sequence of elementary mixture of units/columns operations and permutation of units/columns. Note that $\mathcal{L} = I_G \mathcal{L}$, 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 $T(\alpha, i, j) \in \mathcal{M}_{GG}$ such that

$$T(\alpha, i, j) = \begin{array}{c} \begin{array}{cccccc} 1 & \ldots & i\ldots & \ldots & j & \ldots \end{array} \\ \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \alpha & & 1-\alpha & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \end{array} \tag{1}$$

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

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

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 (2). 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 lead to no effect on the sequence of operations. Thus, $E^G(\mathcal{L}) = E^G\left(I_G \Pi_G^1 \cdot \prod_k T(\alpha_k, i_k, j_k) \cdot \Pi_G\right)$. Making use of axiom UA it follows that $E^G(\mathcal{L}) = E^G\left(I_G \Pi_G^1 \cdot \prod_k T(\alpha_k, i_k, j_k)\right)$. 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 N, it follows that $E^G(I_G \Pi_G^1) = 1$. Then by applying axiom EM we obtain that $E^G[I_G \Pi_G^1 \cdot T(\alpha_k, i_k, j_k)] = \alpha_k \cdot E^G(I_G \Pi_G^1) = \alpha_k$. By repeated application of axiom EM one obtains that

$$E^G(\mathcal{L}) = E^G\left(\prod_k T(\alpha_k, i_k, j_k)\right) = \prod_k \alpha_k. \tag{3}$$

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 determinant of the matrices, one obtains that

$$\prod_k \alpha_k = \det\left[\prod_k T(\alpha_k, i_k, j_k)\right]. \tag{4}$$

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 operation may invert the sign of the determinant and therefore we may have that combining (3) and (4) one obtains

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

The above consideration 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 $E^G(\mathcal{L}) > 0$, by continuity of E^G we can also approach the situations where $E^G(\mathcal{L}) = 0$, these can be obtained as limiting cases where some $\alpha_k \to 0$.

In order to complete the proof it is left 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 not possible that the sequence leads to $\mathcal{L}$ because $|\det \mathcal{L}|$ is uniquely defined.

Sufficiency. The index $E^G(\mathcal{L}) = |\det \mathcal{L}|$ satisfies all the three axioms UA, N and EM. 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 N). To prove that also axiom EM is satisfied, we need to combine some properties of the determinants.

Recall in the definition of axiom EM 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 $E^G(\mathcal{L}(\alpha, i, j)) = |\det \mathcal{L}(\alpha, i, j)| = \alpha \cdot |\det \mathcal{L}| = \alpha E^G(\mathcal{L})$ as required by axiom EM. ■

B.2 Proof of Corollary 1

Proof. By direct application of the definition of axiom D to the result in Proposition 1, note first that is not necessary to consider 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 only focus on matrices in $\hat{\mathcal{M}}_{GG}$. For these matrices by combining axiom D with the result in Proposition 1 we obtain

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

However, note that by construction

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

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

$$E^N(\mathcal{L}) = \sum_{\{i_1, i_2, ..., i_G\} \subseteq \mathcal{N}_O} |\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})|.$$

Here the ordered distribution of the G units in $\mathcal{N}_O$ 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 $|\det(\boldsymbol{\ell}_{i_1}, \boldsymbol{\ell}_{i_2}, ..., \boldsymbol{\ell}_{i_G})|$ that, being expressed in absolute terms is not modified by permutation of the columns. These considerations lead to the final result. ■

B.3 Proof of Remark 2

Proof. In order to prove the remark, 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 has 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 $E^2(L) = \frac{p-x}{1-x} \cdot (1-x) = p - x = |\det(L)|$ that also coincides with the Gini index of the distribution. ■

B.4 Proof of Theorem 1

Proof. *Necessity part.* Consider matrices $\mathcal{L}, X \in \mathcal{M}_{GG}$ and axiom GEM. Suppose that $\mathcal{L}$ coincide with the identity matrix I_G, then GEM requires that $E^G(I_G X) = E^G(I_G) \cdot$

$F^G(X)$ for all $X \in \mathcal{M}_{GG}$, making use of N follows that $E^G(I_G) = 1$. Then we obtain that $E^G(I_G X) = E^G(X) = F^G(X)$, thus $F^G(X) = E^G(X)$ for all $X \in \mathcal{M}_{GG}$. Moreover note that the set of all matrices X coincides with $\mathcal{M}_{GG}$, thus starting from the identity matrix I_G it is possible to reach any matrix $\mathcal{L}$ in $\mathcal{M}_{GG}$.

Thus GEM in combination with N leads to the following functional equation

$$E^G(\mathcal{L}X) = E^G(\mathcal{L}) \cdot E^G(X).$$

or all $\mathcal{L}, X \in \mathcal{M}_{GG}$.

The general solution of this functional equation is illustrated in Aczél and Dhombres (1989) Ch. 4.2 for all real valued square matrices and real valued functions E^G. It coincides (see Theorem 8 and Corollary 6, p. 41-42) with either (i) $E^G(\mathcal{L}) = |\det \mathcal{L}|^c$ or (ii) $E^G(\mathcal{L}) = |\det \mathcal{L}|^c \cdot \text{sign}(\det \mathcal{L})$ for any c arbitrary real constant. Note that by construction $\mathcal{L}$ is row stochastic and therefore $|\det \mathcal{L}| \leq 1$, moreover the value of E^G ranges in $[0, 1]$, it then follows that $c \geq 0$. Moreover, by UA the value of $E^G(\mathcal{L})$ is independent from permutations of the columns of $\mathcal{L}$, this fact eliminates solution (ii) because any permutation of a matrix changes the sign of its determinant.

We are then left with $E^G(\mathcal{L}) = |\det \mathcal{L}|^c$ with $c \geq 0$ for all $\mathcal{L} \in \mathcal{M}_{GG}$.

We combine now this result with the restrictions from EIS and D. We illustrate the restrictions on c 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 vector $\boldsymbol{\ell}_{2_1} = \alpha \boldsymbol{\ell}_2$, and $\boldsymbol{\ell}_{2_2} = (1-\alpha)\boldsymbol{\ell}_2$, for $0 < \alpha < 1$. The new matrix is $\mathcal{L}' = (\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2, (1-\alpha)\boldsymbol{\ell}_2)$. Making use of D, following the discussion in the proof of Corollary 1, and considering that $E^G(\mathcal{L}) = |\det \mathcal{L}|^c$ with $c \geq 0$ for all $\mathcal{L} \in \mathcal{M}_{GG}$ one obtains

$$\begin{aligned}
E^2(\mathcal{L}') &= \frac{1}{2!}\left|\det(\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2)\right|^c + \frac{1}{2!}\left|\det(\alpha\boldsymbol{\ell}_2, \boldsymbol{\ell}_1)\right|^c \\
&\quad + \frac{1}{2!}\left|\det(\boldsymbol{\ell}_1, (1-\alpha)\boldsymbol{\ell}_2)\right|^c + \frac{1}{2!}\left|\det((1-\alpha)\boldsymbol{\ell}_2, \boldsymbol{\ell}_1)\right|^c \\
&\quad + \frac{1}{2!}\left|\det((1-\alpha)\boldsymbol{\ell}_2, \alpha^c\boldsymbol{\ell}_2)\right| + \frac{1}{2!}\left|\det(\alpha\boldsymbol{\ell}_2, (1-\alpha)\boldsymbol{\ell}_2)\right|^c \\
&= \left|\det(\boldsymbol{\ell}_1, \alpha\boldsymbol{\ell}_2)\right|^c + \left|\det(\boldsymbol{\ell}_1, (1-\alpha)\boldsymbol{\ell}_2)\right|^c \\
&= \alpha^c\left|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)\right|^c + (1-\alpha)^c\left|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)\right|^c \\
&= \left[\alpha^c + (1-\alpha)^c\right] \cdot \left|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)\right|^c
\end{aligned}$$

According to EIS it should hold that $E^2(\mathcal{L}') = E^2(\mathcal{L})$ where $E^2(\mathcal{L}) = |\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)|^c$. It then follows that

$$E^2(\mathcal{L}') = \left[\alpha^c + (1-\alpha)^c\right] \cdot \left|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)\right|^c = \left|\det(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)\right|^c = E^2(\mathcal{L}),$$

as a result it should hold that

$$\alpha^c + (1-\alpha)^c = 1$$

for all $0 < \alpha < 1$, that is it should be that $c = 1$. Following the same logic the above considerations can be expanded to the case where $G > 2$. In all these cases the restriction that guarantees that EIS is satisfied is that $c = 1$.

To conclude note that we have now that $E^G(\mathcal{L}) = |\det \mathcal{L}|$ for all $\mathcal{L} \in \mathcal{M}_{GG}$. This is the starting point in the proof of Corollary 1 that could be applied directly to derive the final result. Note however that here the result holds for all $\mathcal{L} \in \mathcal{M}_{GN}$.

Sufficiency part. See the sufficiency part of Proposition 1 and Corollary 1 for the axioms N, UA and D. Note that as already argued that the Gini segregation index satisfies EIS. It is only left to prove GEM for all matrices in $\mathcal{M}_{GG}$. If $F^G(X) = E^G(X) = |\det X|$, then GEM holds because

$$E^G(\mathcal{L}X) = |\det \mathcal{L}X| = E^G(\mathcal{L}) \cdot F^G(X) = E^G(\mathcal{L}) \cdot E^G(X) = |\det \mathcal{L}| \cdot |\det X| .$$

This condition is always satisfied because $(\det \mathcal{L}X) = (\det \mathcal{L}) \cdot (\det X)$ and therefore this is the case also if we consider the absolute values. ■

Table 1: Determinants of exposure segregation: 2000)

	OLS				
	(1)	(2)	(3)	(4)	(5)
A) Black pop	0.000**	0.000**	0.000**	0.000**	0.000**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Asiatic pop	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Hispanic pop	0.000**	0.000**	0.000**	0.000**	0.000**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Pct black		-0.004	-0.029	-0.131**	0.218^{+}
		(0.03)	(0.03)	(0.04)	(0.13)
A) Racial segregation		0.184**	0.172**	0.022	-0.119
		(0.05)	(0.05)	(0.06)	(0.09)
A) Pop density (log)		0.012**	0.021**	0.025**	0.015*
		(0.00)	(0.01)	(0.01)	(0.01)
A) Pct foreign born		-0.306**	-0.317**	-0.411**	-0.617**
		(0.05)	(0.06)	(0.07)	(0.15)
A) Migration flow		-1.114**	-1.004**	-1.100**	-2.016**
		(0.29)	(0.30)	(0.33)	(0.64)
B) Avg tax rate			0.706	1.325	1.863
			(1.09)	(1.23)	(3.28)
B) Fiscal revenues pc			-0.068**	-0.051*	-0.109
			(0.03)	(0.03)	(0.09)
B) Expenditure pc			0.000	-0.000	-0.000
			(0.00)	(0.00)	(0.00)
B) Avg EITC exposure			-0.001	-0.002**	-0.002*
			(0.00)	(0.00)	(0.00)
C) Pub. school budget				0.019**	0.015**
				(0.00)	(0.01)
C) Students/teachers (pub.)				-0.006**	-0.006**
				(0.00)	(0.00)
C) Avg pub. school score				-0.004**	-0.002^{+}
				(0.00)	(0.00)
C) Avg dropout rate (pub.)				0.025	0.242
				(0.20)	(0.25)
C) Pub. schoold pc				-0.012	1.003^{+}
				(0.43)	(0.63)
C) Avg tuition				0.000	0.000
				(0.00)	(0.00)
C) Kindergartens(pub.)				-0.142**	-0.171**
				(0.04)	(0.05)
C) Students/teachers (priv.)				0.002	0.006**
				(0.00)	(0.00)
D)Sd of students/teachers (priv.)					-0.002
					(0.00)
D) Sd of students/teachers (pub.)					-0.001
					(0.00)
D) Pct black students (pub.)					-0.296**
					(0.09)
D) Sd of pct black students (pub.)					-0.009
					(0.09)
D) Pct violent crime					13.745*
					(7.43)
D) Crimes pc					-4.565*
					(2.32)
D) Median rent					0.000**
					(0.00)
D) Pct poors					0.307
					(0.27)
D) Income segregation					0.537**
					(0.22)
D) Avg income					0.000
					(0.00)
D) Gini index					-0.037
					(0.10)
D) Spatial inequality					0.015**
					(0.01)
Constant	0.148**	0.103**	0.066**	0.124**	0.009
	(0.00)	(0.02)	(0.03)	(0.06)	(0.13)
R-squared	0.187	0.376	0.400	0.602	0.671
MSA	436	436	436	320	262
F-test	33.150	32.107	23.483	22.583	14.622

Note: Based on authors' elaboration of data from U.S. Census, CCD, PSS. All controls are for year 2000. Significance levels: $^{+}$ = 15%, * = 10% and ** = 5%.

C Tables of results

Table 2: Determinants of exposure segregation convergence: 2000-2015

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
Gini Exposure 2000 (sd)	-0.256**	-0.305**	-0.296**	-0.293**	-0.356**	-0.346**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
A) Black pop		0.000**	0.000**	0.000**	0.000**	0.000**
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Asiatic pop		0.000	0.000*	0.000**	0.000	0.000**
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Hispanic pop		-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
A) Pct black			-0.346*	-0.355*	-0.864**	-1.142^{+}
			(0.19)	(0.19)	(0.22)	(0.74)
A) Racial segregation			0.869**	0.910**	1.148**	2.475**
			(0.28)	(0.29)	(0.36)	(0.52)
A) Pop density (log)			-0.041^{+}	-0.062**	-0.104**	-0.096*
			(0.03)	(0.03)	(0.04)	(0.05)
A) Pct foreign born			-0.101	0.213	1.222**	1.319^{+}
			(0.29)	(0.34)	(0.41)	(0.84)
A) Migration flow			6.878**	6.516**	2.813^{+}	15.240**
			(1.65)	(1.71)	(1.93)	(3.61)
B) Avg tax rate				-3.073	-20.257**	-18.521
				(6.21)	(7.19)	(18.18)
B) Fiscal revenues pc				0.203	0.674**	0.346
				(0.15)	(0.17)	(0.51)
B) Expenditure pc				-0.000**	-0.000**	-0.000
				(0.00)	(0.00)	(0.00)
B) Avg EITC exposure				-0.005	-0.011**	-0.016**
				(0.00)	(0.00)	(0.01)
C) Pub. school budget					-0.033	-0.071*
					(0.03)	(0.04)
C) Students/teachers (pub.)					-0.046**	-0.050**
					(0.01)	(0.02)
C) Avg pub. school score					-0.006	0.002
					(0.00)	(0.01)
C) Avg dropout rate (pub.)					-3.771**	-1.332
					(1.17)	(1.40)
C) Pub. schoold pc					0.056	0.803
					(2.49)	(3.52)
C) Avg tuition					0.000	-0.000
					(0.00)	(0.00)
C) Kindergartens(pub.)					-0.060	-0.005
					(0.26)	(0.29)
C) Students/teachers (priv.)					0.014*	0.004
					(0.01)	(0.01)
D)Sd of students/teachers (priv.)						0.001
						(0.01)
D) Sd of students/teachers (pub.)						-0.018
						(0.01)
D) Pct black students (pub.)						0.493
						(0.51)
D) Sd of pct black students (pub.)						-0.552
						(0.50)
D) Pct violent crime						-66.896^{+}
						(41.44)
D) Crimes pc						-11.439
						(12.97)
D) Median rent						-0.001
						(0.00)
D) Pct poors						5.069**
						(1.51)
D) Income segregation						-1.203
						(1.24)
D) Avg income						0.000**
						(0.00)
D) Gini index						-1.701**
						(0.54)
D) Spatial inequality						0.030
						(0.04)
Constant	0.426**	0.293**	0.219**	0.456**	1.590**	0.325
	(0.02)	(0.02)	(0.11)	(0.16)	(0.32)	(0.74)
R-squared	0.365	0.527	0.550	0.561	0.679	0.766
MSA	436	436	436	436	320	262
F-test	249.630	119.844	57.780	41.478	29.949	22.572

Note: Based on authors' elaboration of data from U.S. Census, ACS, CCD, PSS. All controls are for year 2000. Significance levels: $^{+}$ = 15%, * = 10% and ** = 5%.