SEGREGATION, INFORMATIVENESS AND LORENZ DOMINANCE

Casilda Lasso de la Vega
and Oscar Volij

Discussion Paper No. 13-12

November 2013

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

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

Segregation, Informativeness and Lorenz Dominance*

Casilda Lasso de la Vega †

University of the Basque Country

Oscar Volij ‡

Ben-Gurion University of the Negev

August 8, 2013

Abstract

It is possible to partially order cities according to the informativeness of neighborhoods about their ethnic groups. It is also possible to partially order cities with two ethnic groups according to the Lorenz criterion. We show that a segregation order satisfies four basic axioms if and only if it is consistent with the informativeness criterion. We then use this result to show that for the two-group case, the Lorenz and the informativeness criteria are equivalent.

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

Keywords: Segregation, informativeness, Lorenz dominance, Blackwell's theorem.

*We thank Sergiu Hart and Mikel Bilbao for their valuable comments. Sergiu Hart helped us shorten the proof of Theorem 2 by pointing to Sherman's [14] version of the Blackwell-Sherman-Stein theorem. We also thank the Spanish Ministerio de Educación y Ciencia (project ECO2012-31346) and the Gobierno Vasco (project IT568-13) for research support.

†casilda.lassodelavega@ehu.es

‡ovolij@bgu.ac.il

1 Introduction

Sociologists and economists have long been interested in how to adequately measure segregation. While early studies restricted attention to segregation between two groups, i.e., blacks and whites, or men and women, later ones developed measures for multi-group cases.[1] One of the difficulties of measuring segregation is that it is not clear what segregation actually means. Massey and Denton [11] identified five dimensions of segregation: evenness, exposure, concentration, centralization and clustering. Each of these dimensions captures some aspect of the idea of segregation. Evenness refers to the similarity among distributions of members of different groups across locations. The more similar these distributions are, the less is the degree of segregation. Exposure, on the other hand, refers to the degree of contact among members of the different groups. For the two-group case, concentration refers to the relative amount of space occupied by the minority group, and centralization to the tendency of the minority group to be located in the center of an urban area. Finally, clustering refers to the tendency of the areas populated by the minority group to be clustered together.

Not only do the various dimensions of segregation relate to different concepts, but their concrete measurement requires different kinds of data. Indeed, while measures of concentration, centralization and clustering require some sort of geographical data, evenness and exposure require information only on the numbers of members of the different groups found in the existing locations. Furthermore, for the measurement of the evenness dimension of segregation only data on the different relative distributions of individuals across locations are necessary. Although the number of segregation indices is very large, it is safe to say that most of the segregation literature, both theoretical and empirical, focuses on the evenness dimension, as does the present paper. [2]

[1] See Reardon and Firebaugh [12] for an enumeration and analysis of various multigroup segregation measures. For the two-group case, Massey and Denton [11] provide a comprehensive survey.

[2] For papers that model segregation differently, see Echenique and Fryer [6] and Ballester and

For the two-group case, the literature on segregation borrowed the device of the Lorenz curve from the income inequality literature and applied it to partially order cities. A segregation curve in the context of segregation is the analogue of the Lorenz curve in the context of income inequality. Indeed, recall that for each fraction p, the Lorenz curve depicts the proportion of total income that is owned by the poorest proportion p of the population. A segregation curve is essentially a Lorenz curve where one group, say blacks, is treated as a population, and the other group, say whites, is treated as income. With this convention, the lower the proportion of whites that live in a neighborhood, the "poorer" is a black individual residing there. Thus, for each fraction p, a segregation curve describes the proportion of the total number of whites that share their neighborhoods with the "poorest" fraction p of blacks. Segregation curves appear in the literature as early as in Duncan and Duncan [5]. The early literature on segregation took advantage of segregation curves to partially order cities. Specifically, given two cities, their corresponding segregation curves may or may not cross. If they do not cross then the city whose segregation curve lies below that of the other one is deemed, according to the Lorenz criterion, the more segregated one.

One can also borrow, this time from the literature on the value of information, another device in order to partially order cities, even for the multigroup case. Indeed, given a city, the location of a randomly selected individual is a signal that provides information about the ethnic group he belongs to. In that sense, the collection of distributions of the various ethnic groups across locations can be seen as an experiment in the sense of Blackwell [3, 4], one in which locations play the role of signals and ethnic groups play the role of states of nature. We can then borrow Blackwell's partial order of experiments according to their informativeness and apply it to partially order cities. Specifically, a city whose locations are more informative than another city's locations will be considered more segregated than the latter.

In this paper we show that any segregation partial order of cities that satisfies four

---

Vorsatz[2].

basic axioms must be consistent with the segregation order induced by the informativeness of their neighborhoods. We next use this characterization to show that when restricted to cities with only two groups, the partial order derived from the segregation curves coincides with the partial order derived from the informativenes of the city's neighborhoods. In that sense, not only is the latter partial order applicable to the multigroup case, but it is also a generalization of the standard order based on segregation curves.

The fact that any partial order that satisfies the four axioms must be consistent with the partial order derived from the segregation curves was stated without proof by James and Taeuber [10]. Later, a proof of this result for the case where all locations contain the same number of members of one group (e.g., all occupations contain the same number of women), was proved by Hutchens [9]. Frankel and Volij [7] noted that any order that satisfies three of the four axioms and weak form of the fourth one must be consistent with the partial order associated with the informativeness of the cities' experiments restricted to the class of cities with the same ethnic distribution.[3] We prove this result for the case of all cities, independently of their ethnic distribution.

2 Notation

The basic model of segregation measurement consists of a list of locations containing different numbers of members of various groups. Papers that focus on residential racial segregation refer to the locations as neighborhoods, and to the groups as ethnic groups. Papers dealing with occupational gender segregation usually use occupations as locations and classify the groups by gender. We will use the language of racial residential segregation, and refer to the list of neighborhoods as cities.

Let G be a finite set of ethnic groups. This set will remain fixed for the whole analysis until Section 4 where it will be restricted to contain two groups. A *neighborhood*

[3]See also Grant, Kajii and Polak [8], and Andreoli and Zoli [1] for related results.

n is characterized by its racial composition, which is a vector $(T_n^g)_{g\in G}$ of non-negative numbers, at least one of which is positive. The number T_n^g is the number of residents of n that belong to ethnic group g. A *city* is a finite collection of neighborhoods such that, for each ethnic group g, at least one neighborhood has a positive number of residents of that group. Formally, a city is a system $\langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ such that N is the set of neighborhoods, for each ethnic group $g \in G$, $\sum_{n\in N} T_n^g > 0$, and for each $n \in N$, $\sum_{g\in G} T_n^g > 0$.

Given a city $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$, we denote by $T^g(X)$ the total number of residents of group g: $T^g(X) = \sum_{n\in N} T_n^g$. When it is clear to which city we are referring, we will write simply T^g. We will denote by t_n^g the proportion of individuals of ethnic group g that reside in neighborhood n. Formally, $t_n^g = T_n^g/T^g$. Similarly, $p_n^g = T_n^g/\sum_{g\in G} T_n^g$ is the proportion of residents of n that belong to ethnic group g. The ethnic distribution of a neighborhood n is given by $(p_n^g)_{g\in G} = (T_n^g)_{g\in G}/\sum_{g\in G} T_n^g$, and the ethnic distribution of a city X is given by $(T^g)_{g\in G}/\sum_{g\in G} T^g$.

For any positive integer k, I_k denotes the $k \times k$ identity matrix. We will sometimes apply certain operations on matrices by postmultiplying them with special Markov matrices. A *splitting matrix* is one that is obtained from an identity matrix by splitting some of its columns into several columns. Permuting the columns of a splitting matrix also results in a splitting matrix. When a matrix is postmultiplied by a splitting matrix, some of its columns are split into several proportional columns. A *merging matrix* is one that is obtained from an identity matrix by replacing some of its columns by their sum and then possibly permuting the columns. A product of merging matrices is also a merging matrix. When a matrix is postmultiplied by a merging matrix, some of its columns are replaced by their sum.

3 The Blackwell partial order

Given a set of states of nature $\Omega = \{1, \ldots, I\}$, an *experiment* provides information about the realized state. Specifically, when the realized state is i, the experiment issues a signal with a distribution that depends on i. An experiment on Ω can be described by a Markov matrix (m_{ij}), whose rows represent the possible states of nature, and whose columns represent the possible signals, the entry m_{ij} being the probability that the signal j is sent when the realized state is i. Conversely, every Markov matrix with I rows can be interpreted as an experiment for Ω. Blackwell [4] partially ordered experiments according to their informativeness, and showed that this partial order has a convenient description in terms of the corresponding matrices.

In this section, we will make use of Blackwell's informativeness order on experiments in order to define a (segregation) partial order on cities. The idea is to consider a city as an experiment where neighborhoods play the role of signals and ethnic groups the role of states of nature, and say that city X is more segregated than city Y if the neighborhoods of X are more informative about the ethnic group of its residents than the neighborhoods of Y.

Let $X = \langle N, ((T_n^g)_{g \in G})_{n \in N} \rangle$ be a city. Also let $\phi : \{1, 2, \ldots, |N|\} \to N$ be an ordering of the neighborhoods. The *experiment matrix* of X with respect to ϕ is the $|G| \times |N|$ matrix

$$M(X, \phi) = (m_{ij})$$

where $m_{ij} = t^i_{\phi(j)}$ is the proportion of individuals of group i that reside in neighborhood $\phi(j)$. Note that $M(X, \phi)$ is a Markov matrix. It represents an experiment in the sense of Blackwell. Its generic entry m_{ij} is the probability that a randomly chosen individual belongs to ethnic group i given that he resides in neighborhood $\phi(j)$.

Let $\mathcal{M}$ be the set of Markov matrices with $|G|$ rows. These matrices can be partially ordered according to their informativeness (Blackwell [4]). Given two matrices $A_{|G| \times |N_A|}, B_{|G| \times |N_B|} \in \mathcal{M}$, we say that A *is at least as informative as* B if there is an

$|N_A| \times |N_B|$ Markov matrix Π such that

$$B = A \cdot \Pi.$$

If A is at least as informative as B, it will remain so even after we permute each of the matrices columns in any arbitrary way. Indeed, let P_B be a $|N_B| \times |N_B|$ permutation matrix and let P_A be a $|N_A| \times |N_A|$ permutation matrix. If

$$B = A \cdot \Pi$$

then

$$B \cdot P_B = A \cdot P_A \cdot P_A^T \cdot \Pi \cdot P_B$$

Since $P_A^T \cdot \Pi \cdot P_B$ is a Markov matrix, we conclude that if A is at least as informative as B then $A \cdot P_A$ is at least as informative as $B \cdot P_B$.

We now define a partial order on cities based on the informativeness of their respective experiment matrices.

Definition 1 Let $X = \langle N_X, ((T_n^g)_{g \in G})_{n \in N_X} \rangle$ and $Y = \left\langle N_Y, ((T_{n'}^{\prime g})_{g \in G})_{n' \in N_Y} \right\rangle$ be two cities. We say that X is at least as segregated as Y according to Blackwell's criterion, denoted $X \succcurlyeq_I Y$, if $M(X, \phi)$ is at least as informative as $M(Y, \psi)$ for some orderings $\phi : \{1, 2, \ldots, |N_X|\} \to N_X$ and $\psi : \{1, 2, \ldots, |N_Y|\} \to N_Y$ of the neighborhoods of X and Y, respectively.

Note that segregation according to Blackwell's criterion is well-defined since the informativeness relation on $\mathcal{M}$ is invariant to permutations of columns. Since for most of the analysis the particular ordering of neighborhoods ϕ that is chosen is not important as long as it remains fixed, in what follows we will keep ϕ tacit and write, with some abuse of notation, $M(X)$ instead $M(X, \phi)$.

3.1 Properties of the Blackwell partial order

Let $\mathcal{C}$ be the set of all cities. A *segregation order* is a partial order on $\mathcal{C}$. For any X and $Y \in \mathcal{C}$, $X \succcurlyeq Y$ means that X is as least as segregated as Y according to $\succcurlyeq$.[4] Blackwell's relation $\succcurlyeq_I$ defined above is an example of a segregation order. We will now inquire into the properties that this particular segregation order satisfies.

We say that two cities, $X = \langle N_X, ((T_n^g)_{g\in G})_{n\in N_X}\rangle$ and $Y = \left\langle N_Y, ((T_{n'}^{\prime g})_{g\in G})_{n'\in N_Y}\right\rangle$, are *equivalent* if there is a one-to-one mapping $\varphi : N_X \to N_Y$ such that for all $n \in N_X$, $(T_n^g)_{g\in G} = (T_{\varphi(n)}^{\prime g})_{g\in G}$.

Equivalent cities differ only in the names of their neighborhoods. It is clear that two equivalent cities have the same experiment matrices, up to permutation of columns. Therefore, Blackwell's order satisfies the following axiom.

Anonymity (ANON) A segregation order $\succcurlyeq$ satisfies *anonymity* if for any two equivalent cities X and Y we have $X \sim Y$.

Consider now the city $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ and the city $Y = \langle N, ((\alpha_g T_n^g)_{g\in G})_{n\in N}\rangle$ that is obtained from X by multiplying the number of group g individuals by $\alpha_g > 0$, for $g \in G$. Since both cities have the same proportions t_n^g, we have that $M(X, \phi) = M(Y, \phi)$ for any ordering ϕ of N. Therefore, Blackwell's order satisfies the following axiom.

Composition Invariance (CI) Let $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ be a city and let $Y = \langle N, ((\alpha_g T_n^g)_{g\in G})_{n\in N}\rangle$ be the city that is obtained from X by multiplying the number of agents of a group g, for $g \in G$, by the same nonzero factor $\alpha_g > 0$, in all neighborhoods. A segregation order $\succcurlyeq$ satisfies *composition invariance* if for any such cities we have $Y \sim X$.

Composition invariance requires that only the relative distributions of members of the various ethnic groups across neighborhoods affect segregation. In particular, the city's ethnic distribution does not affect segregation.

[4]Given $\succcurlyeq$, the associated relations $\succ$ and $\sim$ are defined as usual. $X \succ Y \Leftrightarrow X \succcurlyeq Y$ and not $Y \succcurlyeq X$, and $X \sim Y \Leftrightarrow X \succcurlyeq Y$ and $Y \succcurlyeq X$.

Let $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ be a city and consider the city Y that is obtained from X by splitting a particular neighborhood $(T_n^g)_{g\in G}$ into two neighborhoods, n_1 and n_2, with the same ethnic distribution, namely, $(T_{n_1}^g)_{g\in G} = (\alpha T_n^g)_{g\in G}$ and $(T_{n_2}^g)_{g\in G} = ((1-\alpha)T_n^g)_{g\in G}$ for some $\alpha \in (0,1)$. Then, their experiment matrices satisfy

$$M(Y) = M(X) \cdot I(n, \alpha) \tag{1}$$

where $I(n, \alpha)$ is the splitting matrix that is obtained from the identity matrix $I_{|N|}$ by splitting the column that corresponds to neighborhood n into two columns, according to the proportions α and $(1-\alpha)$. Furthermore,

$$M(X) = M(Y) \cdot I(n_1, n_2) \tag{2}$$

where $I(n_1, n_2)$ is the merging matrix that is obtained from the identity matrix $I_{|N|+1}$ by merging the two columns that correspond to n_1 and n_2 into one. Equations (1) and (2) imply that $M(X)$ and $M(Y)$ are equally informative, and therefore $X \sim_I Y$. Consequently, Blackwell's order satisfies the following axiom.

Organizational Equivalence (OE) Let $X \in \mathcal{C}$ be a city and let $(T_n^g)_{g\in G}$ be one of its neighborhoods. Let Y be the city that results from dividing $(T_n^g)_{g\in G}$ into two neighborhoods, $(T_{n_1}^g)_{g\in G}$ and $(T_{n_2}^g)_{g\in G}$, with the same ethnic distribution. Namely, $(T_{n_1}^g)_{g\in G} = (\alpha T_n^g)_{g\in G}$ and $(T_{n_2}^g)_{g\in G} = ((1-\alpha)T_n^g)_{g\in G}$ for some $\alpha \in (0,1)$. A segregation order $\succcurlyeq$ satisfies *organizational equivalence* if for any such cities we have $Y \sim X$.

Let $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ be a city and consider the city Y that is obtained from X by splitting a particular neighborhood $(T_n^g)_{g\in G}$ into 2 neighborhoods, n_1 and n_2, but now with different ethnic distributions. Then

$$M(X) = M(Y) \cdot I(n_1, n_2)$$

where, as before, $I(n_1, n_2)$ is the merging matrix that is obtained from the identity matrix $I_{|N|+1}$ by merging the two columns that correspond to n_1 and n_2 into one. Therefore, $Y \succcurlyeq_I X$.

On the other hand, as the following lemma states, there is no $|N| \times (|N| + 1)$ Markov matrix Π such that $M(Y) = M(X) \cdot \Pi$. Hence, $Y \succ_I X$.

Lemma 1 *Let A be an $n \times m$ Markov matrix and let B be an $n \times (m+1)$ Markov matrix that is obtained from A by splitting one of A's columns into two, but not proportionally. Then, there is no Markov matrix Π such that $B = A \cdot \Pi$.*

Proof. See appendix. ■

Therefore, Blackwell's order satisfies the following axiom.

Neighborhood Division Property (NDP) Let $X \in \mathcal{C}$ be a city and let $(T^g_n)_{g \in G}$ be a neighborhood of X. Let Y be the city that results from dividing $(T^g_n)_{g \in G}$ into 2 neighborhoods, $(T^g_{n_1})_{g \in G}$ and $(T^g_{n_2})_{g \in G}$ with different ethnic distributions. Namely, $(T^g_{n_1})_{g \in G} \neq (\alpha T^g_n)_{g \in G}$ for any $\alpha \in [0, 1]$. A segregation order $\succcurlyeq$ satisfies the *neighborhood division property* if for any such cities we have $Y \succ X$.

We summarize the above observations in the following Proposition.

Proposition 1 *The Blackwell segregation order $\succcurlyeq_I$ satisfies ANON, CI, OE, and NDP.*

We can now state our first result.

Theorem 1 *Let $\succcurlyeq$ be a segregation order on $\mathcal{C}$. It satisfies ANON, CI, OE and NDP if and only if for all two cities $X, Y \in \mathcal{C}$,*

$$Y \succ_I X \;\; \Rightarrow \;\; Y \succ X \tag{3}$$

$$Y \sim_I X \;\; \Rightarrow \;\; Y \sim X \tag{4}$$

Theorem 1 states that all segregation orders that satisfy ANON, CI, OE and NDP are consistent with Blackwell's order. Namely, whenever Blackwell's order ranks two cities, any segregation order that satisfies the above four axioms must rank them in the same way. And conversely, any segregation order that is consistent with Blackwell's order must satisfy the four axioms.

Proof. Let $\succcurlyeq$ be a segregation order that satisfies (3) and (4). We will show that it satisfies the four axioms.

ANON: Let X and Y be two equivalent cities. Then, by Proposition 1, $X \sim_I Y$. By (4), $X \sim Y$.

CI: Let $X = \langle N, ((T_n^g)_{g\in G})_{n\in N}\rangle$ be a city and let $Y = \langle N, ((\alpha_g T_n^g)_{g\in G})_{n\in N}\rangle$ be the city that is obtained by multiplying the number of agents of a group g by the same nonzero factor $\alpha_g > 0$, for $g \in G$ in all neighborhoods. Then, by Proposition 1, $Y \sim_I X$. By (4), $Y \sim X$.

OE: Let $X \in \mathcal{C}$ be a city and let $(T_n^g)_{g\in G}$ be a neighborhood of X. Let Y be the city that results from dividing $(T_n^g)_{g\in G}$ into 2 neighborhoods, $(T_{n_1}^g)_{g\in G}$ and $(T_{n_2}^g)_{g\in G}$ with the same ethnic distribution. Then, by Proposition 1, $Y \sim_I X$. By (4), $Y \sim X$.

NDP: Let $X \in \mathcal{C}$ be a city and let $(T_n^g)_{g\in G}$ be a neighborhood of X. Let Y be the city that results from dividing $(T_n^g)_{g\in G}$ into 2 neighborhoods, $(T_{n_1}^g)_{g\in G}$ and $(T_{n_2}^g)_{g\in G}$ with different ethnic distributions. Then, by Proposition 1, $Y \succ_I X$. By (3), $Y \succ X$.

We now show that any partial order that satisfies the four axioms must be consistent with Blackwell's criterion. Let now $\succcurlyeq$ be a segregation order that satisfies ANON, CI, OE and NDP. Also, let $X = \langle N_X, ((T_n^g)_{g\in G})_{n\in N_Y}\rangle$ and $Y = \left\langle N_Y, ((T_{n'}^{\prime g})_{g\in G})_{n'\in N_Y}\right\rangle$ be two cities such that $Y \succcurlyeq_I X$. We need to show that (3) and (4) hold. Since $Y \succcurlyeq_I X$, there is a Markov matrix $\Pi = \left((\pi_{ij})_{i=1}^{|N_Y|}\right)_{j=1}^{|N_X|}$ such that

$$M(X) = M(Y) \cdot \Pi.$$

But Π can be written as a product of two matrices

$$\Pi = \beta \cdot \gamma$$

where

$$\beta = \begin{pmatrix} \pi_{11} & \cdots & \pi_{1|N_X|} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \pi_{21} & \cdots & \pi_{2|N_X|} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \pi_{|N_Y|1} & \cdots & \pi_{|N_Y||N_X|} \end{pmatrix}$$

and

$$\gamma = \begin{pmatrix} I_{|N_X|} \\ I_{|N_X|} \\ \vdots \\ I_{|N_X|} \end{pmatrix}$$

Therefore,

$$M(X) = M(Y) \cdot \beta \cdot \gamma \tag{5}$$

Note that $M(Y) \cdot \beta$ is the matrix that is obtained from $M(Y)$ by splitting its ith column, $i = 1, \ldots |N_Y|$, into $|N_X|$ columns, in the proportions π_{ij}, $j = 1, \ldots |N_X|$. Also note that $M(Y) \cdot \beta \cdot \gamma$ is obtained from $M(Y) \cdot \beta$ by merging all the ith (mod($|N_X|$)) columns together. Therefore, (5) says that $M(X)$ is obtained from $M(Y)$ by successively splitting columns proportionally and then merging columns, which may or may not be proportional to each other. Alternatively, matrix $M(Y)$ is obtained from $M(X)$ by splitting its columns, not necessarily in a proportional way, and then merging some proportional columns. Consequently, by OE and NDP,

$$Y \succcurlyeq X.$$

An analogous argument shows that if $X \succcurlyeq_I Y$ we must also have $X \succcurlyeq Y$. Consequently, if $Y \sim_I X$ then $Y \sim X$, which is implication (4).

In order to show implication (3) assume that $Y \succ_I X$. We already know that matrix $M(Y)$ is obtained from $M(X)$ by splitting its columns, not necessarily in a proportional way, and then merging some proportional columns. We now argue that $Y \succ_I X$ implies that at least one of the columns is split *not* in a proportional way. Indeed, if all the columns of $M(X)$ were split proportionally, we would have

$$M(Y) = M(X) \cdot \beta' \cdot \gamma'$$

for some splitting matrix β' and merging matrix γ'. Since, as a product of Markov matrices $\beta' \cdot \gamma'$ is a Markov matrix, this would imply that $X \succcurlyeq_I Y$, contradicting $Y \succ_I X$. Therefore, Y is obtained from X by splitting some neighborhoods into smaller neighborhoods with different ethnic distributions, and then merging some neighborhoods with the same ethnic distributions. By NDP and OE, $Y \succ X$, which shows the implication in (3). ■

4 Two groups: The Lorenz partial order

There is another partial order defined on the class of cities with only two groups. It is known as the Lorenz partial order and is based on what is known as segregation curves. See Duncan and Duncan [5], James and Taeuber [10, 15] and Hutchens [9].

Let G be a set of two ethnic groups and denote by $\mathcal{C}_2$ the set of cities with these two groups. For ease of exposition, we refer to members of the two ethnic groups as blacks and whites, respectively. Let $X = \langle N, (B_n, W_n)_{n \in N} \rangle \in \mathcal{C}_2$ be a city, where for each neighborhood $n \in N$, B_n and W_n are the numbers of blacks and whites, respectively, that reside in n. For each $n \in N$, denote by p_n the proportion of whites in neighborhood n. That is, $p_n = W_n/(B_n + W_n)$. Also, b_n and w_n denote the proportion of the city's blacks and whites, respectively that reside in neighborhood n. Formally, $b_n = B_n / \sum_{n' \in N} B_n$ and $w_n = W_n / \sum_{n' \in N} W_n$. We will now build the segregation curve associated with the city X. Segregation curves, as experiment matrices in the $|G|$-group case analyzed in Section 3, will allow us to define a partial order on the set of two-group cities. Segregation curves, analogously to experiment matrices, are objects that do not depend on the cities' ethnic distribution. That is, city $X = \langle N, (B_n, W_n)_{n \in N} \rangle$ and city $\widehat{X} = \langle N, (b_n, w_n)_{n \in N} \rangle$, which is obtained from X by normalizing the groups' populations so that each group is of size one, will have the same segregation curve. In order to build the segregation curve, let $\phi : \{1, 2, \ldots, |N|\} \to N$ be an ordering of the neighborhoods such that $i \leq j \Rightarrow p_{\phi(i)} \leq p_{\phi(j)}$. Namely, ϕ orders neighborhoods in a non-decreasing way according to their proportion whites. Note that

$$p_n \leq p_m \Leftrightarrow w_n/(b_n + w_n) \leq w_m/(b_m + w_m). \tag{6}$$

That is, ordering the neighborhoods in N in non-decreasing order of the proportion of whites in X or in its normalized version $\widehat{X}$ results in the same order. Let $\beta_0 = \omega_0 = 0$, and for $i = 1, 2, \ldots, |N|$, and let $\beta_i = \beta_{i-1} + b_{\phi(i)}$ and $\omega_i = \omega_{i-1} + w_{\phi(i)}$. That is, β_i is the proportion of blacks that reside in the i neighborhoods with the lowest proportions of whites. Similarly, ω_i is the proportion of whites that reside in these same neighborhoods. *The Lorenz segregation curve of* X is the graph that is obtained by plotting the points $(\beta_i, \omega_i)_{i=0}^{|N|}$ and connecting the dots. Formally, it is the union of the line segments $seg[(\beta_{i-1}, \omega_{i-1}), (\beta_i, \omega_i)]$, $i = 1, 2, \ldots, |N|$, where for any two points $x, y \in \mathbb{R}^2$, $seg[x, y] = \{\alpha x + (1 - \alpha)\, y : \alpha \in [0, 1]\}$. Note that the line segment that connects the points $\left(\beta_{i-1}, \omega_{i-1}\right)$ and (β_i, ω_i) has a slope of $w_{\phi(i)}/b_{\phi(i)}$. Therefore, given (6), this slope is non-decreasing in i. Furthermore, the segregation curve is invariant to the choice of ordering ϕ as long as it satisfies $i \leq j \Rightarrow p_{\phi(i)} \leq p_{\phi(j)}$.

We can use the segregation curves to define a segregation order.

Definition 2 Let X and Y be two cities. We say that Y is at least as segregated as X according to the Lorenz criterion, denoted $Y \succcurlyeq_L X$, if the Lorenz curve of Y is nowhere above the Lorenz curve of X.

We can now state the main result of this section.

Theorem 2 *The Blackwell and the Lorenz orders on $\mathcal{C}_2$ are the same.*

Proof. It can be checked that the Lorenz order $\succcurlyeq_L$ satisfies ANON, CI, OE and NDP. Therefore, by Theorem 1, $Y \succcurlyeq_I X \Longrightarrow Y \succcurlyeq_L X$.

In order to show the converse implication, let $X = \langle N, (B_n, W_n)_{n\in N}\rangle$ and $Y = \langle N', (B'_{n'}, W'_{n'})_{n'\in N'}\rangle$ be two cities in $\mathcal{C}_2$. Since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy CI, we can assume without loss of generality that $\sum_{n\in N} B_n = \sum_{n'\in N'} B'_{n'} = \sum_{n\in N} W_n = \sum_{n'\in N'} W'_{n'} = 1$. Since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy ANON we can also assume that $N = \{1, \ldots, I\}$ and $N' = \{1, \ldots, I'\}$ and that the neighborhoods are ordered in a non-decreasing order of proportion of whites. Therefore, we can denote X by $(b_n, w_n)_{n=1}^{I}$ and Y by $(b'_{n'}, w'_{n'})_{n'=1}^{I'}$ with $w_1/(b_1+w_1) \leq \cdots \leq w_I/(b_I+w_I)$ and $w'_1/(b'_1+w'_1) \leq \cdots \leq w'_{I'}/(b'_{I'}+w'_{I'})$.

Case 1: For each $n \in N$ and $n' \in N'$, $b_n > 0$ and $b'_{n'} > 0$.

Let's build the following random variables: For each $n \in N$ the random variable x takes the value w_n/b_n with probability b_n. For each $n' \in N'$ the random variable y takes the value $w'_{n'}/b'_{n'}$ with probability $b'_{n'}$. Note that $E[x] = E[y] = 1$.

Denote by F_x the cumulative distribution function of x and by F_y the cumulative distribution function of y. Also denote their generalized inverses by F_x^{-1} and F_y^{-1}, respectively.[5]

By Theorem 3.A.5 of Shaked and Shanthikumar [13],

$$\begin{array}{lcll} \int_0^p F_x^{-1}(t)dt & \geq & \int_0^p F_y^{-1}(t)dt & \text{for all } p \in [0,1] \\ & \iff & & \\ \sum_{n\in N} b_n \phi(w_n/b_n) & \leq & \sum_{n'\in N'} b'_n \phi(w'_n/b'_n) & \text{for all convex functions } \phi : \mathbb{R} \to \mathbb{R}. \end{array} \quad (7)$$

By Sherman's [14] theorem, (7) holds if and only if there is a $I' \times I$ Markov matrix $\Pi = \{\pi_{n'n}\}$ such that

$$\begin{array}{rcl} b_n\left(w_n/b_n\right) & = & \sum_{n'\in N'} \pi_{n'n} b'_{n'}\left(w'_{n'}/b'_{n'}\right) \\ \sum_{n'\in N'} \pi_{n'n} b'_{n'} & = & b_n \end{array}$$

Or, in matrix notation,

$$M(X) = M(Y) \cdot \Pi$$

[5]The generalized inverse of a distribution function $F : \mathbb{R}_+ \to [0,1]$ is defined as $F^{-1} : [0,1] \to \mathbb{R}_+$ such that $F^{-1}(p) = \inf_s \{s > 0 : F(s) > p\}$.

Consequently, $\int_0^p F_x^{-1}(t)dt \geq \int_0^p F_y^{-1}(t)dt$ for all $p \in [0,1]$ if and only if $Y \succcurlyeq_I X$. Since the graphs of $\int_0^p F_x^{-1}(t)dt$ and $\int_0^p F_y^{-1}(t)dt$ are none other than the segregation curves of X and Y respectively, we obtain the desired result, i.e., $Y \succcurlyeq_L X \Leftrightarrow Y \succcurlyeq_I X$.

Case 2: For each $n \in N$, $b_n > 0$, and there is $n' \in N'$ with $b'_{n'} = 0$. Since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy OE, we can assume without loss of generality that in Y, there is only one neighborhood with no blacks. Furthermore, by OE we can assume without loss of generality that $|N'| = |N| + 1$. That is, $I' = I + 1$. Lastly, since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy OE, we can assume without loss of generality that $b_n = b'_n$ for all $n \in \{1, 2, \ldots I\}$. Therefore, we can denote X by $(b_n, w_n)_{n=1}^{I}$ and Y by $(b_n, w'_n)_{n=1}^{I+1}$ (where $b_{I+1} = 0$ and $w'_{I+1} > 0$). In this case, $X \succcurlyeq_L Y$ is impossible. Assume, therefore, that $Y \succ_L X$.

For each $t = 1, \ldots$, let $\varepsilon_t = \frac{1}{t} b_I \frac{w'_{I+1}}{w'_I + w'_{I+1}}$, and let $Y_t = (b_n^t, w'_n)_{n=1}^{I+1}$ be the city that is obtained from Y by relocating ε_t blacks from neighborhood I to neighborhood $I+1$. That is, $b_n^t = b_n$ for $n = 1, \ldots I-1$, $(b_I^t, w'_I) = (b_I - \varepsilon_t, w'_I)$ and $\left(b_{I+1}^t, w'_{I+1}\right) = \left(\varepsilon_t, w'_{I+1}\right)$. See Figure 1. Note that since $\varepsilon_t \leq b_I \frac{w'_{I+1}}{w'_I + w'_{I+1}}$, the proportion of whites in neighborhood

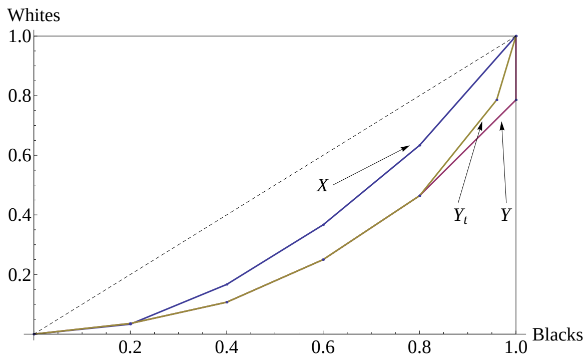

Figure 1: The segregation curves of X, Y and Y_t

I is less than or equal to the proportion of whites in neighborhood $I+1$. As a result, Y_t's neighborhoods are ordered in a non-decreasing order of the proportion of whites. Furthermore, it can be seen that $Y_t \succcurlyeq_L X$. By construction, Y_t has no neighborhoods with 0 blacks. By Case 1, $Y_t \succcurlyeq_I X$. That is, there is a $I' \times I$ Markov matrix Π_t such that

$$M(X) = M(Y_t) \cdot \Pi_t$$

Since the set of $I' \times I$ Markov matrices is compact, there is a subsequence $\{\Pi_{t_\ell}\}$ that converges to a Markov matrix Π. Since, $M(Y_t) \to M(Y)$, we obtain that

$$M(X) = M(Y) \cdot \Pi$$

which means that $Y \succcurlyeq_I X$.

Case 3: There is $n \in N$, and $n' \in N'$ such that $b_n = b'_{n'} = 0$.

Since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy OE, we can also assume without loss of generality that $|N| = |N'| = I$. By OE, we can assume without loss of generality that both in X and in Y, there is only one neighborhood with no blacks. Lastly, since both $\succcurlyeq_I$ and $\succcurlyeq_L$ satisfy OE, we can assume without loss of generality that $b_n = b'_n$ for all $n \in \{1, 2, \ldots I\}$.

Assume that $Y \succcurlyeq_L X$. For each $t = 1, \ldots$, let $\varepsilon_t = \frac{1}{t} b_I \frac{w_I}{w_{I-1}+w_I}$, and let $X_t = (b^t_n, w_n)_{n=1}^I$ be the city that is obtained from X by relocating ε_t blacks from neighborhood $I-1$ to neighborhood I. That is, $b^t_n = b_n$ for $n = 1, \ldots I-2$, $(b^t_I, w_I) = (b_{I-1} - \varepsilon_t, w_{I-1})$ and $(b^t_I, w_I) = (\varepsilon_t, w_I)$. See Figure 2. Note that since $\varepsilon_t \leq b_I \frac{w_I}{w_{I-1}+w_I}$ the proportion

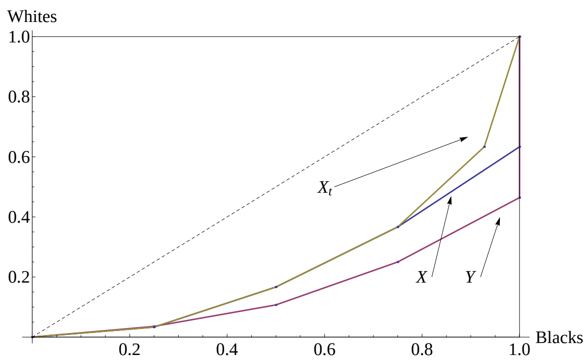

Figure 2: The segregation curves of X, Y and X_t

of whites in neighborhood $I-1$ is less than or equal to the proportion of whites in neighborhood I. As a result, X_t's neighborhoods are ordered in a non-decreasing order of the proportion of whites. Furthermore, by construction, $Y \succcurlyeq_L X_t$. Also by construction,

X_t has no neighborhoods with 0 blacks. By Case 2, $Y \succcurlyeq_I X_t$. That is, there is a $I' \times I$ Markov matrix Π_t such that

$$M(X_t) = M(Y) \cdot \Pi_t$$

Since the set of $I' \times I$ Markov matrices is compact, there is a subsequence $\{\Pi_{t_\ell}\}$ that converges to a Markov matrix Π. Since, $M(X_t) \to M(X)$, we obtain that

$$M(X) = M(Y) \cdot \Pi$$

which means that $Y \succcurlyeq_I X$. ■

5 Appendix

Proof of Lemma 1: Let A be an $n \times m$ Markov matrix and let B be an $n \times (m+1)$ Markov matrix that is obtained from A by splitting one of A's columns into two. Assume that A's kth column is the one that is split. Alternatively, A is obtained from B by replacing B's kth and $(k+1)$th columns by their sum. Consequently,

$$A = B \cdot S_k \tag{8}$$

where

$$S_k = \begin{pmatrix} I_{k-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{m-k} \end{pmatrix}.$$

Let us now assume that there is an $m \times (m+1)$ Markov matrix Π such that

$$B = A \cdot \Pi. \tag{9}$$

We will show that B is necessarily obtained from A by splitting A's kth column proportionally.

Let Π' be the matrtix that is obtained from Π by replacing Π's kth and $(k+1)$th columns by their sum. That is,

$$\Pi' = \Pi \cdot S_k. \tag{10}$$

Note that Π' is a square $m \times m$ Markov matrix. Moreover, by (10), (9) and (8),

$$A \cdot \Pi' = A \tag{11}$$

which means that each row of A is an invariant distribution of the matrix Π'.

Since Π' is a Markov matrix, there exists $r \geq 1$ and a permutation matrix P such that $P^T \cdot \Pi' \cdot P$ can be written in the following (almost block diagonal) form:

$$\left(\begin{array}{ccccc|c} R'_1 & & & & & 0 \\ & R'_2 & & 0 & & 0 \\ & & R'_3 & & & 0 \\ & 0 & & \ddots & & \vdots \\ & & & & R'_r & 0 \\ \hline S'_{r+1,1} & S'_{r+1,2} & S'_{r+1,3} & \cdots & S'_{r+1,r} & Q' \end{array}\right)$$

where for all $j = 1, ..., r$, R'_j are square $(m_j \times m_j)$ irreducible Markov matrices and Q' is an $\left(n - \sum_{j=1}^r m_j\right) \times \left(n - \sum_{j=1}^r m_j\right)$ reducible matrix. We can assume without loss of generality that P is the identity matrix and thus that Π' has the above form.[6]

Since R'_j, for $j = 1, ..., r$, is an irreducible Markov matrix, it has unique invariant distribution $q^j = (q^j_1, \ldots, q^j_{m_j})$, i.e., q^j is the unique probability vector q that satisfies $q = qB'_j$. Furthermore, any invariant distribution of Π' can be written as

$$(\alpha_1 q^1, \alpha_2 q^2, \ldots, \alpha_r q^r, \underbrace{0, \ldots, 0}_{n-\sum_{j=1}^r m_j})$$

for some $\alpha_1, \ldots \alpha_r \geq 0$ and $\sum_{j=1}^r \alpha_j = 1$ (see, for instance, Lucas and Stokey 1989 (Theorem 11.1, pages 326-330)). Therefore, since each row of A is an invariant distribution

[6]Otherwise, the whole analysis can be done using and $A \cdot P$ instead of A, $S_k \cdot P$ instead of S_k, $P^T \cdot \Pi$ instead of Π and $P^T \cdot \Pi' \cdot P$ instead of Π'.

of Π', it can be written as

$$A = \begin{pmatrix} \alpha_{11}q^1 & \alpha_{12}q^2 & \cdots & \alpha_{1r}q^r & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n1}q^1 & \alpha_{n2}q^2 & \cdots & \alpha_{nr}q^r & 0 & \cdots & 0 \end{pmatrix}, \tag{12}$$

where for each $i = 1, \ldots, n$ and $j = 1, \ldots, r$, $\alpha_{ij} \geq 0$ and $\sum_{j=1}^{r} \alpha_{ij} = 1$. If B was obtained from A by splitting column k in a disproportional way, it ought to be the case that this column is one that has at least one positive entry.

Assume that column k corresponds to the hth block of Π'. Therefore we can write

$$R'_h = \left(R'_{h_1}, v'_{*k}, R'_{h_2}\right)$$

where $v'_{*k} = (v'_{1k}, \ldots, v'_{m_h k})^T$ is the column of block B'_h that corresponds to the kth column of Π'. Since Π is obtained from Π' by splitting the kth column into two, Π can be written as

$$\Pi = \left(\begin{array}{ccccccccc|c} R'_1 & 0 & & 0 & 0 & & \cdots & 0 & 0 \\ 0 & \ddots & & 0 & 0 & & \cdots & \vdots & \vdots \\ & & R'_{h_1} & v_{*k} & v_{*k+1} & R'_{h_2} & & 0 & 0 \\ \cdots & \cdots & & 0 & 0 & & \ddots & 0 & 0 \\ 0 & \cdots & & 0 & 0 & & 0 & R'_r & 0 \\ \hline S'_{r+1,1} & \cdots & S'_{r+1,h_1} & s_{*k} & s_{*k+1} & S'_{r+1,h_2} & \cdots & S'_{r+1,r} & Q' \end{array}\right) \tag{13}$$

where v_{*k} and v_{*k+1} are column vectors such that $v_{*k} + v_{*k+1} = v'_{*k}$. Consequently, since $B = A \cdot \Pi$, using (12) and (13) we obtain that B's kth column is $\left(\alpha_{1h}q^h v_{*k}, ..., \alpha_{nh}q^h v_{*k}\right)^T$ and, B's $(k+1)$th column is $\left(\alpha_{1h}q^h v_{*k+1}, ..., \alpha_{nh}q^h v_{*k+1}\right)^T$, which are proportional to each other (the proportion is $q^h v_{*k}/q^h v_{*k+1}$). ■

References

[1] F. Andreoli and C. Zoli, On the Measurement of Dissimilarity and Related Orders, 2013, mimeo.

[2] C. Ballester and M. Vorsatz, Random-Walk-Based Segregation Measures. Review of Economics and Statistics (2013), forthcoming.

[3] D. Blackwell, Comparison of Experiments, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press (1951), 93-102.

[4] D. Blackwell, Equivalent Comparisons of Experiments, Annals of Mathematical Statistics 24 (1953), 265–72.

[5] O. Duncan and B. Duncan, A Methodological Analysis of Segregation Indexes, American Sociological Review, Vol. 20, No. 2 (Apr., 1955), 210-217.

[6] F. Echenique, R. G. Fryer, Jr., A Measure of Segregation Based on Social Interactions, Quarterly Journal of Economics 122 (2007), 441–85.

[7] D. Frankel and O. Volij, Measuring School Segregation, Journal of Economic Theory 146 (2011), 1–38.

[8] S. Grant, A. Kajii, and B. Polak, Intrinsic Preference for Information, Journal of Economic Theory 83 (1998), 233–59.

[9] R. Hutchens, Segregation Curves, Lorenz Curves, and Inequality in the Distribution of People across Ocuppations, Mathematical Social Sciences 21 (1991), 31-51.

[10] D. R. James and K. E. Taeuber, Measures of Segregation, Sociological Methodology 15 (1985), 1–32.

[11] D. S. Massey and N. A. Denton, The Dimensions of Racial Segregation, Social Forces 67 (1988), 281–315.

[12] S. F. Reardon and G. Firebaugh, Measures of Multigroup Segregation, Sociological Methodology 32 (2002), 33–67.

[13] M. Shaked and J. G. Shanthikumar, Stochastic Orders, Springer Series in Statistics, Springer, (2007).

[14] S. Sherman, On a Theorem of Hardy, Littlewood , Polya, and Blackwell, PNAS 37 (1951), 826–31.

[15] K. E. Taeuber and D. R. James, Racial Segregation among Public and Private Schools, Sociology of Education 55 (1982), 133–43.