Mathematical Social Sciences 73 (2015) 63–68

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/econbase

Symmetric measures of segregation, segregation curves, and Blackwell's criterion

Robert Hutchens

Cornell University, United States

HIGHLIGHTS

- A new approach is proposed for using segregation curves to measure segregation.
- The resulting incomplete segregation order incorporates a "symmetry of types" property.
- Using Blackwell's criterion, the approach is extended to more than two types of people.
- The new approach results in a more complete segregation order.

ARTICLE INFO

Article history:
Received 25 February 2014
Received in revised form
11 October 2014
Accepted 27 November 2014
Available online 8 December 2014

ABSTRACT

This paper first proposes a new way to use segregation curves to examine whether one distribution of people across groups (e.g., occupations or neighborhoods) is more segregated than another. It then uses Blackwell's criterion to extend the argument to more than two types of people. The basic idea is that by introducing additional assumptions about the nature of segregation, one obtains a more complete ranking of distributions. The paper demonstrates that the assumption of "symmetry in types" – an assumption that appears frequently in the literature on segregation measurement – has implications for both segregation curves and Blackwell's criterion.

A large and growing empirical literature seeks to measure segregation in school systems, occupational networks, and city neighborhoods.[1] The topic is important because ethnic and gender segregation is intimately tied to issues of access and opportunity—issues that are of fundamental interest to social scientists and the larger society. Standing alongside this empirical literature is a second more theoretical literature that grapples with the problem of how to properly measure segregation. When can we say that one distribution of people across groups (e.g., occupations or neighborhoods) is more segregated than another? This paper contributes to that second literature.

In particular, the paper proposes a new way to use segregation curves to examine whether one distribution is more segregated than another. Segregation curves were introduced into the literature in a seminal paper by Otis and Beverly Duncan.[2] In their analysis of racial segregation in US cities, the Duncans assumed two types of people (specifically, whites and non-whites) and plotted curves that indicate whether segregation in city A is greater than that in city B. The resulting segregation curves are similar to the Lorenz curves used to assess income inequality. Like Lorenz curves, segregation curves have the advantage of resting on a few weak and plausible assumptions about the nature of segregation. Also like Lorenz curves, segregation curves yield an incomplete order of distributions; when segregation curves cross, they yield no information about whether one distribution of people across groups can be ranked as more segregated than another.

Whether such an incomplete order is an advantage or disadvantage is in the eye of the beholder. On the one hand it is awkward to announce that because of intersecting segregation curves we do not know whether distribution X is more segregated than Y. There exist several numerical measures of segregation that yield a complete order of distributions; why not use one of them? On the other hand, intersecting segregation curves force us to say, "it is only by making additional assumptions about the nature of segregations – assumptions that are often neither explicit nor explained – that we can claim that X is more segregated than Y". There can be virtue in such candor.

This paper first shows that by introducing an additional plausible assumption about the nature of segregation, one can increase the number of alternatives that can be ranked by segregation

E-mail address: rmh2@cornell.edu.

[1] For example, Massey and Denton (1988), Flückiger and Silber (1999), Hutchens (2001, 2004), Weeden (2004) and Frankel and Volij (2011).

[2] Otis D. Duncan and Beverly Duncan (1955).

http://dx.doi.org/10.1016/j.mathsocsci.2014.11.007
0165-4896/

curves. It then extends that result to a more general segregation order that is based on Blackwell's criterion. In the case of segregation curves, the practical implication is as follows: if the segregation curves for X and Y cross, then plot a new segregation curve denoted X^T. This new segregation curve is closely related to the segregation curve for X, and is thus called the "twin" of X. If X^T lies beneath Y – more precisely, if X^T lies at some point below and at no point above Y – then despite the fact that the segregation curves for X and Y cross, distribution X can be declared more segregated than distribution Y, and that increases the number of alternatives that can be ranked.

But since that result pertains to segregation curves, it is restricted to situations where one can reasonably assume two types of people. A recent paper by Lasso de la Vega and Volij (2014) applies Blackwell's criterion (Blackwell, 1951, 1953) to the problem of measuring segregation. They not only obtain an incomplete order that can be used to assess segregation when there are more than two types of people (e.g., white Hispanic, black Hispanic, white non-Hispanic, black non-Hispanic), but they also establish an intimate link between that incomplete order and segregation curves. Their work raises the possibility of extending this paper's result to the general problem of two or more types of people. The final section of the paper shows that the result does, indeed, generalize. Thus, by making an additional plausible assumption about the nature of segregation, one can increase the number of alternatives that can be ranked by Blackwell's criterion, and that result applies to two or more types of people.

1. The problem, properties, segregation curves, and previous results

Consider I types of people distributed over J occupations. Let x_{ij} denote the number of type i people in occupation j ($i = 1, \ldots, I;\ j = 1, \ldots, J$), let $N_i(X) = \sum_j x_{ij},\ i = 1, \ldots, I$ denote the total number of type i people over all occupations, and let X be a data matrix of the x_{ij}. For example, type 1 could be women, type 2 could be men, and occupation j could be one of J occupations. To insure a meaningful problem, assume that J is greater than one, that the x_{ij} are non-negative real numbers,[3] and that the number of type i people is positive ($N_i(X) > 0,\ i = 1, \ldots, I$). Thus, with I types of people and J occupations, the data matrix takes the form,

$$X = \left[\boldsymbol{x}_1,\ \boldsymbol{x}_2, \ldots, \boldsymbol{x}_J\right] = \begin{bmatrix} x_{11} & \cdots & x_{1J} \\ \vdots & \ddots & \vdots \\ x_{I1} & \cdots & x_{IJ} \end{bmatrix}$$

where $\boldsymbol{x}_j$ is a column vector of length I that contains data on the types of people in occupation j. Denoting the vector space of all $I \times J$ real matrices with non-negative elements by $\mathbf{R}_+^{I\times J}$, the domain of X shall be $D_I = \bigcup_{J=2}^{\infty} D_{IJ}$ where $D_{IJ} = \left\{X \in \mathbf{R}_+^{I\times J} : N_i(X) > 0,\ i = 1, \ldots, I\right\}$.

For example, with two types of people and four occupations the data matrix could take the form,

$$X = \begin{matrix} \\ \textit{type 1} \\ \textit{type 2} \end{matrix} \overset{\textit{Occupation}}{\begin{bmatrix} \underline{1} & \underline{2} & \underline{3} & \underline{4} \\ 3 & 1 & 1 & 4 \\ 8 & 5 & 7 & 4 \end{bmatrix}}. \tag{1}$$

Thus, letting type 1 people be women, the distribution in (1) has three women and eight men in the first occupation, and the total number of men and women is respectively 24 and 9 ($N_1(X) = 3 + 1 + 1 + 4 = 9$ and $N_2(X) = 8 + 5 + 7 + 4 = 24$). Letting Y be another matrix of two types of people over four occupations, e.g.,

$$Y = \begin{bmatrix} 3 & 6 & 2 & 4 \\ 4 & 5 & 4 & 5 \end{bmatrix}, \tag{2}$$

this paper seeks to evaluate whether the X distribution is more segregated than Y.

To that end, let $\succcurlyeq$ denote a segregation order on D_I.[4] If $X \succ Y$ then X is ranked as more segregated than Y, and if $X \backsim Y$ then X has the same level of segregation as Y. Of course, the order may be partial; if the order does not rank some X and Y, then the result is an incomplete order.

Like any measure of inequality, a "good" measure of segregation should have properties that accord with perceptions of segregation. These properties are ultimately value judgments about what it means to say that one social state is more segregated than another. As such, the properties should be as unrestrictive as possible. Restrictive properties imply strong value judgments, which in turn lead to more controversial conclusions about whether X is more segregated than Y.

The literature on measuring segregation advances several properties for a good measure. Four of these properties are particularly important not only because they are unrestrictive, but also because they define a broad class of measures that are linked to segregation curves. While these properties are often stated for two types of people (e.g., men and women), because this paper ultimately addresses the more general problem of I types, it is useful to state the properties in their more general form. Finally, since the four properties are discussed extensively elsewhere, the following is a brief summary.

P1. **Scale invariance**. Let Y be obtained from X by multiplying the number of type i people in every occupation of X by a positive scalar β_i. Then $Y \sim X$.

P2. **Symmetry in occupations**. Let $(j_1, \ldots, j_J)$ be any permutation of $1, \ldots, J$, $X = \left[\boldsymbol{x}_1,\ \boldsymbol{x}_2, \ldots, \boldsymbol{x}_J\right]$ and $Y = \left[\boldsymbol{x}_{j_1},\ \boldsymbol{x}_{j_2}, \ldots, \boldsymbol{x}_{j_J}\right]$. Then $Y \sim X$.[5]

P3. **Organizational equivalence**. Let Y be obtained from X by dividing occupation J into two occupations such that for some $\alpha \in (0, 1)$,

$$y_{ij} = x_{ij}, \quad i = 1, \ldots, I;\ j = 1, \ldots, J-1$$
$$y_{iJ} = x_{iJ}\alpha, \quad i = 1, \ldots, I$$
$$y_{iJ+1} = x_{iJ}(1-\alpha), \quad i = 1, \ldots, I.$$

Then $Y \sim X$.

P4. **Neighborhood division property**. Let Y be obtained from X by dividing occupation J into two occupations such that,

$$y_{ij} = x_{ij}, \quad i = 1, \ldots, I;\ j = 1, \ldots, J-1$$
$$y_{iJ} + y_{iJ+1} = x_{iJ}, \quad i = 1, \ldots, I$$

and for **no** $\alpha \in [0, 1]$,

$$y_{iJ} = x_{iJ}\alpha, \quad i = 1, \ldots, I$$
$$y_{iJ+1} = x_{iJ}(1-\alpha), \quad i = 1, \ldots, I.$$

Then $Y \succ X$.

[3] Not only does this assumption address the general case, but it is also plausible. Part-time workers could be treated as fractional workers. Irrational numbers are also theoretically admissible (e.g., a full time worker is counted as "1" and a part-time worker as the square-root of 0.3).

[4] See Foster (1985, p. 42) for a particularly clear definition of a measure of inequality and an inequality order.

[5] The property could be equivalently stated in terms of a permutation matrix. Specifically, let P be a $J \times J$ permutation matrix with only one coefficient in each row equal to 1 and only one coefficient in each column equal to 1. If $Y = XP$ then $Y \sim X$.