Universidade Vigo

Departamento de Economía Aplicada
http://webx06.webs4.uvigo.es

Working Paper: 19/04. September 2019

On Measuring Segregation in a Multigroup Context: Standardized Versus Unstandardized Indices

Coral del Río

Olga Alonso-Villar

Coordinator: José María Chamorro Rivas
chamorro@uvigo.es

On Measuring Segregation in a Multigroup Context: Standardized Versus Unstandardized Indices*

Coral del Río and Olga Alonso-Villar#

Universidade de Vigo (ECOSOT-ECOBAS and EQUALITAS)

Abstract

There has been little discussion in the literature about the consequences of using standardized (versus unstandardized) segregation measures when comparing societies with different demographic compositions. To measure the segregation of a group in a multigroup setting, this paper develops standardized local segregation indices, which show a maximum value of 1 when the group is fully segregated, and links these measures with existing standardized overall segregation measures. Our research not only allows for enhancement of the local segregation approach—offering new measures and evaluating them against basic properties—but also provides a better understanding of existing standardized overall measures. To illustrate its value, this paper offers estimates of the occupational segregation of white women in the largest U.S. metropolitan areas using standardized and unstandardized segregation measures. This permits us to identify metropolitan areas that would have gone unnoticed if only one of these two approaches had been employed.

JEL Classification: D63; J15; J16; J71

Keywords: Multigroup segregation; Standardized segregation indices; Local segregation curves; Local segregation indices

---

* We gratefully acknowledge financial support from the *Ministerio de Economía, Industria y Competitividad,* the *Agencia Estatal de Investigación,* and the *Fondo Europeo de Desarrollo Regional* (ECO2017-82241-R) and *Xunta de Galicia* (ED431B2019/34).

Correspondence address: Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. Tel.: +34 986812507; fax: +34 986812401; e-mail: ovillar@uvigo.es.

1. Introduction

As societies grow more diverse—whether in terms of race, ethnicity, or other characteristics of individuals—there is an increasing need to measure segregation in a framework that involves three or more groups. Since the 1990's, there have been several indicators developed to quantify overall multigroup segregation, mainly according to an evenness perspective that focuses on differences in the sorting of the groups across organizational units (Silber, 1992; Boisso et al., 1994; Reardon and Firebaugh, 2002; Frankel and Volij, 2011).

In a multigroup setting, apart from overall segregation, it is also possible to measure the segregation of a group (called local segregation, to distinguish it from overall segregation). The framework put forward by Alonso-Villar and Del Río (2010) allows for the measurement of local segregation in a way that is consistent with the measurement of overall segregation. This approach allows for identification of each group's situation and contribution to overall segregation, which depends not only on the group's segregation level but also on its demographic size.[1] The indices proposed by these authors satisfy the property of scale invariance, according to which if the size of a group is multiplied by a positive number, the segregation level of that group remains unaffected provided that there is no change to its distribution across units (nor to the relative size of each organizational unit). Thus, for example, if the economy has 5 units, each of which accounts for 40 people, the segregation of a group of size 40 distributed as (10,10,10,10,0) should be equal to that of another group of size 80 distributed as (20,20,20,20,0). Both groups experience the same segregation level because they account for equal shares in each unit (a quarter of the corresponding group's population in the first four units and zero in the last one).

The property of scale invariance may result in the belief that the segregation level of a group is independent of the size of the group (in the above example, the segregation level is the same for groups of size 40 and 80). However, one must remember that a group's segregation and its size are not completely independent dimensions. As we will discuss in more detail later on, the demographic share of a group impacts the highest segregation level that the group can attain. Pursuant to our above example (an economy consisting of

[1] This approach also allows for the measurement of the consequences of segregation for each group in monetary terms (Del Río and Alonso-Villar, 2015) and in terms of objective well-being (Alonso-Villar and Del Río, 2017a).

200 people and 5 organizational units of equal size), a group consisting of 40 individuals is fully segregated if it is concentrated in units with no workers from other groups, i.e., (40, 0, 0, 0, 0), which implies that this group has no presence in units accounting for 80% of the total population. This scenario is impossible for a group of size 80 because, for such a group to be fully segregated, no group members may be found in units representing 60% of the total population, i.e., (40, 40, 0, 0, 0). In other words, this group is missing from a smaller part of the economy (60% vs. 80%).

The maximum segregation that a group can experience depends on the group's size, which is thereby an important feature to account for in empirical studies; this is particularly true when comparing segregation levels of groups of very different (relative) sizes, exploring the segregation of a growing group over time, or in international comparisons when analyzing a group whose size varies significantly among countries.

Given that some of the most well-known overall segregation indices can be expressed as weighted averages of the (local) segregation of the groups involved (Alonso-Villar and Del Río, 2010), it is reasonable to expect that the (relative) size of the groups may also determine the maximum value attainable by overall indices. In fact, many of these overall indices are not equal to 1 when there is full segregation. This is the case for the I_p index (Silber, 1992), the (unstandardized) Gini index (Alonso-Villar and Del Rio, 2010), and the mutual information index (Theil and Finizza, 1971; Frankel and Volij, 2011).[2] For this reason, Reardon and Firebaugh (2002) opted for standardized (or normalized) indices between 0 and 1. In particular, making use of disproportionality functions that compare the presence of each group in each unit with its share in the economy, these authors derived the generalized dissimilarity index, the generalized Gini index, and the Theil information theory index. These three indices result from dividing each of the abovementioned unstandardized indicators by its maximum value, which is a function of the groups' shares (Reardon and Firebaugh, 2002).[3]

However, as far as we know, there has been little discussion since that time of the consequences of using standardized (versus unstandardized) measures (Mora and Ruiz-

[2] Consider a simple economy with 4 groups of the same size and 4 equally-sized units. In case of maximum segregation, both the I_p index and the Gini index are equal to 0.75, whereas the mutual information index is equal to 1.39. If instead we had 4 groups (and 4 units) of sizes 10, 10, 10, and 70, the maximum values of the indices would be 0.48 and 0.79, respectively.

[3] These authors developed another overall segregation index, based on the squared coefficient of variation, the maximum value of which depends not on the group's weights but on the number of groups.

Castillo, 2011), especially when comparing societies with different demographic compositions. This paper aims to: a) develop standardized local segregation indices, which have a maximum value of 1 when the group is fully segregated; and b) link these measures with existing (standardized) overall segregation measures. Our research not only facilitates enhancement of the local segregation approach—offering new measures and evaluating them against basic properties—but a better understanding of some of the standardized overall segregation measures assessed in Reardon and Firebaugh (2002).

Importantly, our analysis shows that when a segregation indicator is standardized by dividing it by its maximum conditional on the groups' weights, instead of the absolute maximum, there is a different maximum depending on the case.[4] This implies that the resulting indices quantify segregation from an angle significantly different from the perspective assumed by unstandardized segregation indices. Further, this is the case whether we use local or overall measures. Unstandardized measures associated with disproportional functions account for the extent of the distance between the distribution of the groups (across units) and the "egalitarian" distribution—according to which the presence of each group in each unit must equal the expected value assigned by its weight in the economy. On the contrary, standardized measures quantify the proximity of the former distribution to the distribution of maximum segregation, which (per the three indices mentioned earlier) depends on the groups' weights. A standardized segregation approaching a score of 1 does not necessarily imply a high level of unstandardized segregation. Reciprocally, a high value of unstandardized segregation may be regarded as a low level when accounting for the demographic structure. By standardizing segregation using a maximum that depends on demographic composition, the resulting index attains a value of 1 when there is full segregation, although that scenario may be associated with a high or low unstandardized segregation level.[5]

In undertaking this research, we use the local segregation curve proposed by Alonso-Villar and Del Río (2010), which helps to interpret the relationship that exists between a group's size and its maximum segregation. We then define standardized local segregation

[4] The absolute maximum of an index refers to the highest segregation reached if the shares of the groups (and units) are not given. Standardizing indices using this maximum would ensure a common reference, which would facilitate comparability. In empirical analyses, however, it seems more sensible to take the size of the groups as given.

[5] Note that the standardized overall index *C* proposed by Reardon and Firebaugh (2002) using the squared coefficient of variation does not have this problem provided that one compares situations with the same number of groups.

indices—in line with the standardization measurement proposed by Reardon and Firebaugh (2002) in the case of overall segregation—and evaluate them against a set of properties. In addition, we establish the conditions under which the ranking provided by the local segregation curves is consistent with that of the standardized local indices. Finally, drawing on the 2012-16 American Community Survey (ACS), we conduct an empirical examination of the occupational segregation of white women in the largest metropolitan areas in the U.S., using standardized and unstandardized indices. We identify metropolitan areas that would have gone unnoticed if we had pursued only one of these two approaches.

2. Measuring the Segregation of a Group: A New Proposal

As mentioned earlier, the maximum segregation level of a group is not independent of the group's size. The reason is that when a group is small, it can be absent from organizational units that account for a large share of total population, whereas this situation is impossible for large groups. In other words, the larger the group, the lower the maximum segregation. How, therefore, is it possible to compare two groups that differ in terms of size but are distributed across units in the same way in relative terms? Here we explore a procedure that measures the segregation of a group while accounting not merely for how the group is distributed across units, but also the maximum segregation attainable by the group. In other words, we quantify the distance of a group's distribution from that of its maximum segregation.

2.1 Relationship Between Group's Size and Group's Maximum Segregation

To illustrate the effect of a group's size on its maximum segregation level, we use the local segregation curve proposed by Alonso-Villar and Del Río (2010). This curve shows the underrepresentation of a group across organizational units vis-à-vis the distribution of the population across these same units.

Let g be one of the N mutually exclusive groups of society (g=1,…,N). c_j^g denotes the number of individuals of group g in unit j (j=1,…,J), t_j is the number of total individuals in that unit ($c_j^g \le t_j$), $C^g = \sum_j c_j^g$ is the group's size, and $T = \sum_j t_j$ is total population.

To build the local segregation curve of group g, the demographic share of which is $\frac{C^g}{T}$, first, we must rank the organizational units in ascending order of the ratio $\frac{c_j^g}{t_j}$. Then, the cumulative proportion of total individuals is plotted on the horizontal axis, while the cumulative proportion of group's g individuals is plotted on the vertical axis. Namely, if we denote by $\tau_j \equiv \sum_{i \le j} \frac{t_i}{T}$ the proportion of individuals who are in the first j organizational units, the segregation curve at point τ_j is

$$S^g(\tau_j) = \frac{\sum_{i \le j} c_i^g}{C^g}.$$

The curve at intermediate points is determined by linear interpolation (see Figure 1). This tool can be used to compare different scenarios. Thus, if one curve dominates another (i.e., no point of the former curve lies below the latter curve and does at some point lie above, as is the case of S^g relative to S^{g*} in Figure 1), we can say that the group is less segregated in the first case than in the second.

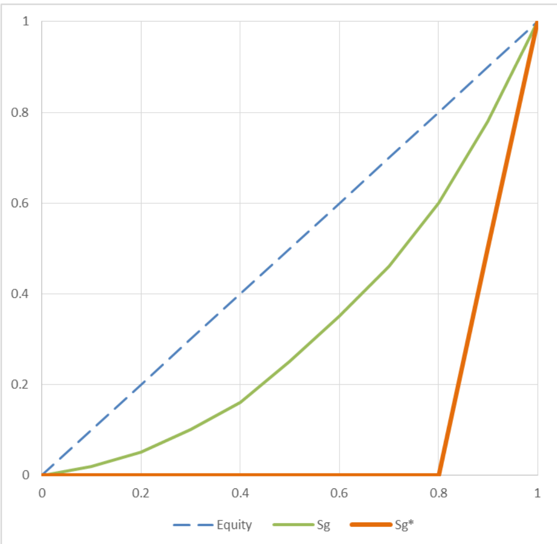

Figure 1. Two examples of local segregation curves

Alonso-Villar and Del Río (2010) proposed several local segregation indices—adapted from well-known inequality measures—to quantify the extent to which a local segregation curve diverges from an even distribution of the group across organizational units (i.e., if $\frac{c_j^g}{t_j} = \frac{C^g}{T}$) and, therefore, the level of that group's segregation. These indices are G^g, D^g, and Φ_α^g (see Table 1).

Table 1. Unstandardized and standardized local segregation indices

Local Segregation Indices	**Maximum Value of the Index**	**Standardized Local Segregation Indices**
$G^g = \dfrac{\sum_{i,j} \frac{t_i}{T}\frac{t_j}{T}\left\lvert \frac{c_i^g}{t_i} - \frac{c_j^g}{t_j} \right\rvert}{2\frac{C^g}{T}}$	$G^{g*} = 1 - \dfrac{C^g}{T}$	$\tilde{G}^g = \dfrac{\sum_{i,j} \frac{t_i}{T}\frac{t_j}{T}\left\lvert \frac{c_i^g}{t_i} - \frac{c_j^g}{t_j} \right\rvert}{2\frac{C^g}{T}\left(1-\frac{C^g}{T}\right)}$
$D^g = \dfrac{1}{2}\sum_j \left\lvert \dfrac{c_j^g}{C^g} - \dfrac{t_j}{T} \right\rvert$	$D^{g*} = 1 - \dfrac{C^g}{T}$	$\tilde{D}^g = \dfrac{\frac{1}{2}\sum_j \left\lvert \frac{c_j^g}{C^g} - \frac{t_j}{T} \right\rvert}{1-\frac{C^g}{T}}$
$\Phi_1^g = \sum_j \dfrac{c_j^g}{C^g} \ln\left(\dfrac{c_j^g/C^g}{t_j/T}\right)$	$\Phi_1^{g*} = \ln\left(\dfrac{T}{C^g}\right)$	$\tilde{\Phi}_1^g = \dfrac{\sum_j \frac{c_j^g}{C^g}\ln\left(\frac{c_j^g/C^g}{t_j/T}\right)}{\ln\left(\frac{T}{C^g}\right)}$
$\Phi_\alpha^g = \dfrac{1}{\alpha(\alpha-1)}\sum_j \dfrac{t_j}{T}\left[\left(\dfrac{c_j^g/C^g}{t_j/T}\right)^{\alpha} - 1\right]$	$\Phi_\alpha^{g*} = \dfrac{1}{\alpha(\alpha-1)}\left[\left(\dfrac{C^g}{T}\right)^{1-\alpha} - 1\right]$	$\tilde{\Phi}_\alpha^g = \dfrac{\sum_j \frac{t_j}{T}\left[\left(\frac{c_j^g/C^g}{t_j/T}\right)^{\alpha} - 1\right]}{\left(\frac{C^g}{T}\right)^{1-\alpha} - 1}$

Note: The expression for Φ_α^g is valid for $\alpha \neq 0,1$.

The G^g index is equal to twice the area between the local segregation curve and the 45º line. The index D^g measures the highest vertical distance of the curve to the 45º line. Along with its graphical interpretation, this index has a very intuitive meaning: when multiplied by 100, it represents the percentage of group g individuals that would have to switch organizational units for the group to have a segregation level of zero (Alonso-Villar and Del Río, 2017).[6] The generalized entropy family offers a different index depending on a parameter, α. This family accounts for both the group's underrepresentation in organizational units (i.e., the lower part of the local segregation curve) and its overrepresentation (the upper part of the curve), although the lower the value of α, the more sensitive the index is to the group's underrepresentation.[7] The α values that are more commonly employed are 0.5, 1, and 2.

Several of these local indices are related to overall segregation indices already extant in the literature (Alonso-Villar and Del Río, 2010). Thus, the weighted average of local indices G^g, D^g, Φ_1^g and Φ_2^g (with weights equal to the groups' shares) are, respectively, equal to the (unstandardized) Gini index (Alonso-Villar and Del Rio, 2010), the I_p index (Silber, 1992), the mutual information index M (Theil and Finizza, 1971; Frankel and Volij, 2011), and the unstandardized version of Reardon and Firebaugh's (2002) C index (divided by 2). It is important to note that, although these overall indices can be decomposed by groups in several ways, the components of such decompositions may not necessarily be good measures of the groups' segregation. For example, the mutual information index can be written as the weighted average (with weights equal to the groups' shares) of the difference between the entropy of the distribution of the population across units and the entropy of each group (Frankel and Volij, 2011). However, the difference between entropies is not a sensible local segregation indicator because its minimum value is not attained when the group is distributed across units in the same manner as the total population is—the difference can take negative values—nor does it

[6] This index was initially proposed by Moir and Shelby Smith (1979) in a binary context, although its properties in a multigroup context, together with its relation to the local segregation curve, were later explored in Alonso-Villar and Del Río (2010). In the case of two groups, labeled 1 and 2, the sum of the segregation of the groups, $D^1 + D^2$, is equal to the dissimilarity index.

[7] Index $\Phi_0^g = \sum_j \frac{t_j}{T} \ln\left(\frac{t_j / T}{c_j^g / C^g}\right)$ can only be used if the group appears in all organizational units, i.e., if $c_j^g > 0$ for all j. For this reason, we will not define a standardized version of this index.

satisfy the property of *insensibility to proportional* subdivisions—the entropy is sensitive to the number of organizational units. On the contrary, the indices proposed by Alonso-Villar and Del Río (2010) are truly local segregation measures because they satisfy a wide range of properties.

As mentioned above, the maximum segregation of a group is attained when it is fully concentrated in units with no members of other groups.[8] Let us assume, without loss of generality, that a group is fully concentrated in one organizational unit with no members of other groups.[9] Figure 1 illustrates this situation as the case of a group that accounts for 20% of the population. The curve of maximum segregation, denoted by S^{g*}, is equal to 0 up to the unit in which the group is fully concentrated (i.e., at point $1-\frac{C^g}{T}$) and jumps to 1 when that unit is aggregated with the previous ones (i.e., when the cumulative proportion of population is 100%), thereby rendering a straight line between these two points. Table 1 shows the maximum values of the indices G^g, D^g, Φ_1^g and Φ_α^g, labelled, respectively, G^{g*}, D^{g*}, Φ_1^{g*} and Φ_α^{g*} (the proofs are shown in Appendix A).

2.2 Standardized Local Segregation Measures

Building on the framework put forward by Alonso-Villar and Del Río (2010), we here develop several indicators with which to quantify the segregation of a group while accounting for that group's maximum segregation. These indices, globally denoted by $\tilde{\Theta}^g\left(c^g,t\right)$, are defined as the quotient between a local segregation index, $\Theta^g\left(c^g,t\right)$, and the value of that index when the group is fully segregated, Θ^{g*}, where c^g is the vector representing the number of individuals of group g in each unit j (i.e., c_j^g) and t is the vector indicating the number of individuals in each unit j (i.e., t_j). Namely, $\tilde{\Theta}^g\left(c^g,t\right)=\frac{\Theta^g\left(c^g,t\right)}{\Theta^{g*}}$. This approach squares with the measurement of overall

[8] Note that, in the real world, full segregation may not be possible because the number and size of the units cannot fit with the group's size. This is why when an index is standardized being divided by its maximum, it is accomplished using a theoretical maximum that does not account for the units but instead approximates the maximum existing in each empirical case. Consequently, the actual distribution of maximum segregation may differ among indices.

[9] The property of *insensibility to proportional subdivisions* (see next subsection) ensures that we can focus on cases in which the group is concentrated in one unit of size equal to that of the group because distributions of maximum segregation across several units would be equivalent to this.

segregation put forward by Reardon and Firebaugh (2002) in that we divide the index by the maximum segregation level, although in our case segregation refers to a group (say, white women) rather than to overall segregation (say, by gender and race).

To measure the segregation of a group we propose using the indices $\tilde{G}^g$, $\tilde{D}^g$, $\tilde{\Phi}_1^g$, and $\tilde{\Phi}_\alpha^g$, shown in Table 1, which are obtained dividing the indices G^g, D^g, Φ_1^g, and Φ_α^g, respectively, by their maximum values (G^{g^*}, D^{g^*}, $\Phi_1^{g^*}$, and $\Phi_\alpha^{g^*}$). Imposing this standardization yields a maximum value of indices $\tilde{G}^g$, $\tilde{D}^g$, $\tilde{\Phi}_1^g$, and $\tilde{\Phi}_\alpha^g$ that is always 1, which facilitates comparisons among different groups or a group across time and space.

Making use of the interpretation of D^g mentioned above, $\tilde{D}^g$ (multiplied by 100) may be thought of as the proportion of group g individuals who must transfer among units to attain 0 segregation divided by the proportion who must move if the group were fully segregated.

To determine whether these indices are suitable to measure the segregation of group, we list some basic properties proposed in the literature for that purpose, put forth new properties, and check whether our measures satisfy them.

Alonso-Villar and Del Río (2010) established the following properties for measuring the segregation of a group:

a) *Size Invariance*, which signifies that if we multiply both the number of individuals of the group and the number of total workers in each organizational unit by a positive number, the segregation of the group does not change. Namely, if $c_j^g{}' = \lambda c_j^g$ and $t_j{}' = \lambda t_j$ for any $\lambda > 0$ and $j = 1,...,J$, then $\Theta^g\left(c^g{}', t'\right) = \Theta^g\left(c^g, t\right)$.
b) *Scale Invariance* refers to the fact that the group's segregation does not change if the number of individuals in the group doubles, for example, and the total number of individuals triples. Namely, if $c_j^g{}' = \lambda c_j^g$ and $t_j{}' = \beta t_j$ for $j = 1,...,J$ (where $\lambda > 0$, $\beta > 0$, and $\lambda c_j^g \le \beta t_j$), then $\Theta^g\left(c^g{}', t'\right) = \Theta^g\left(c^g, t\right)$.
c) *Symmetry*, which means that if the organizational units are permuted, the segregation of the group remains unaltered. Namely, if $c_j^g{}' = c_{\Pi(j)}^g$ and $t_j{}' = t_{\Pi(j)}$,

where $(\Pi(1),...,\Pi(J))$ is a permutation of units $(1,...,J)$, then $\Theta^g\left(c^g{}',t'\right)=\Theta^g\left(c^g,t\right)$.

d) *Insensitivity to Proportional Subdivisions* of organizational units, i.e., the segregation level of the group does not change if a unit is split into several units of equal size with identical number of individuals of the group. Namely, assuming for the sake of simplicity that we split the last unit in K>0 units, if $c_j^g{}'=c_j^g$ and $t_j{}'=t_j$ for any $j=1,...,J-1$, and $c_{J+i}^g{}'=\frac{c_J^g}{K}$ and $t_{j+i}{}'=\frac{t_j}{K}$ for $i=0,...,K-1$, then $\Theta^g\left(c^g{}',t'\right)=\Theta^g\left(c^g,t\right)$.

e) *Sensitivity to Disequalizing Movements (type I)*: Disequalizing movements of the group between equally-sized organizational units, the size of which does not change after that movement (i.e., if a unit with a lower number of individuals of the target group than another loses some of those individuals in favor of the latter, other things being equal) increase the group's segregation.[10] Namely, if $c_i^g{}'=c_i^g-d$ and $c_h^g{}'=c_h^g+d$, where *i* and *h* are two units such that $c_i^g<c_h^g$ and $t_i=t_h$, whereas $c_j^g{}'=c_j^g$ for $j\neq i,h$, then $\Theta^g\left(c^g{}',t\right)>\Theta^g\left(c^g,t\right)$.

Note that alternative definitions of *sensitivity to disequalizing movements* may be articulated depending on how strictly we conceive of the circumstances under which we expect segregation to increase. This is why we put forth two new properties here:

f) *Sensitivity to Disequalizing Movements (type II)*: Disequalizing movements of the group between one organizational unit and another unit with a higher representation of the group (i.e., if the group's representation diminishes in the former unit and rises in the latter), while the size of these units do not change, produce an increase in the group's segregation. Namely, if $c_i^g{}'=c_i^g-d$ and $c_h^g{}'=c_h^g+d$, where *i* and *h* are two units such that $\frac{c_i^g}{t_i}<\frac{c_h^g}{t_h}$, whereas $c_j^g{}'=c_j^g$ for $j\neq i,h$, then $\Theta^g\left(c^g{}',t\right)>\Theta^g\left(c^g,t\right)$.

g) *Sensitivity to Disequalizing Movements (type III)*: Disequalizing movements of the group between one organizational unit and another unit with a higher

[10] In Alonso-Villar and Del Río (2010) this property appears as "movement between groups."

representation of the group (i.e., if the group's representation diminishes in the former unit and rises in the latter), whereas the sizes of these units change accordingly, result in an increase in the group's segregation. Namely, if $c_i^g{}' = c_i^g - d$, $c_h^g{}' = c_h^g + d$, $t_i{}' = t_i - d$, and $t_h{}' = t_h + d$, i and h being two units such that $\frac{c_i^g}{t_i} < \frac{c_h^g}{t_h}$, whereas $c_j^g{}' = c_j^g$ and $t_j{}' = t_j$ for $j \neq i,h$, then $\Theta^g\left(c^g{}', t'\right) > \Theta^g\left(c^g, t\right)$.

The question we now pose is whether properties (f) and (g) are too restrictive or, instead, commonly fulfilled. Propositions 1 and 2 reveal that these new properties, which allow us to compare more scenarios than does property (e), are not difficult to satisfy. In fact, as Corollaries 1 and 2 show, many local indices (unstandardized and standardized) meet them.

Proposition 1. *If a local segregation index $\Theta^g\left(c^g, t\right)$ satisfies insensitivity to proportional divisions and sensitivity to disequalizing movements type I, then it also fulfills sensitivity to disequalizing movements type II.*

Proof. See Appendix A.

Proposition 2. *Any local segregation index $\Theta^g\left(c^g, t\right)$ consistent with the dominance criterion given by the local segregation curves satisfies sensitivity to disequalizing movements type III.*

Proof. See Appendix A.

Before continuing any further, let us summarize the properties that extant local segregation measures fulfill using Corollary 1.

Corollary 1. *The indices G^g and Φ_α^g satisfy size and scale invariance, symmetry, insensitivity to proportional divisions, sensitivity to disequalizing movements type I, sensitivity to disequalizing movements type II, and sensitivity to disequalizing movements type III. Index D^g fulfills size and scale invariance, symmetry, and insensitivity to proportional divisions.*

Proof. See Appendix A.

Finally, Corollary 2 shows the properties fulfilled by our standardized indices.

Corollary 2. *The indices* $\tilde{G}^g$ *and* $\tilde{\Phi}^g_\alpha$ *satisfy size invariance, symmetry, insensitivity to proportional divisions, sensitivity to disequalizing movements type I, sensitivity to disequalizing movements type II, and sensitivity to disequalizing movements type III. The index* $\tilde{D}^g$ *fulfills size invariance, symmetry, and insensitivity to proportional divisions.*

Proof. See Appendix A.

The next theorem demonstrates the relationship that exists between the dominance criterion associated with the local segregation curves and the standardized indices.

Theorem. *If the local segregation curve of a group dominates that of another group whereas the opposite holds for the curves of maximum segregation, then segregation will be lower in the first case than in the second for any standardized local segregation index* $\tilde{\Theta}^g\left(c^g,t\right)=\dfrac{\Theta^g\left(c^g,t\right)}{\Theta^{g*}}$, *where* $\Theta^g\left(c^g,t\right)$ *satisfies scale invariance, symmetry, insensitivity to proportional divisions, and sensitivity to disequalizing movements type I.*[11]

Proof. See Appendix A.

Note that the properties that we require Θ^g meet are the properties that render these indices consistent with the dominance criterion established by Alonso-Villar and Del Río (2010) in their Theorem 1. In other words, for any local segregation index Θ^g that satisfies *scale invariance*, *symmetry*, *insensitivity to proportional divisions*, and *sensitivity to disequalizing movements type I*, Θ^g is lower in a given scenario than in another if, and only if, the local segregation curve of the former case dominates the latter.

Accordingly, it follows that if the local segregation curve of a group is above another (i.e., the former dominates the latter) and the ranking is the reverse for the groups' curves of maximum segregation, we need not calculate any $\tilde{\Theta}^g$ index (included in the set of indices established in the theorem) because all of them would lead to the same conclusion: segregation is lower for the first group.

Proposition 3. *The local segregation curve of a group associated with that group's maximum segregation dominates that of another group if, and only if, in the former case the group accounts for a larger share of the population than it does in the latter.*

[11] If there is dominance in one case and the curves are equal in the other case, the theorem still holds.

Proof. See Appendix A.

It follows from the foregoing proposition that to determine whether the curve of maximum segregation for a group dominates that of another group we need only know these groups' demographic shares.

2.3 The Segregation Level of a Group and Overall Segregation

In their 2002 paper, Reardon and Firebaugh derived several standardized overall measures using the notion of disproportionality (i.e., the overrepresentation and underrepresentation of groups in units), and assessed them against James and Taeuber's (1985) criteria.

As Table 2 shows, these overall indices, *G*, *D*,[12] *H*,[13] and *C*,[14] can be decomposed, respectively, in terms of local indices $\tilde{G}^g$, $\tilde{D}^g$, $\tilde{\Phi}_1^g$, and $\tilde{\Phi}_2^g$, in such a way that overall segregation is the weighted average of the local segregation of the groups involved. In the case of indices *G* and *D*, the weight of a group, $\tilde{w}^g$, represents the percentage of individuals, relative to total population, who must change units in group *g* for this group, starting from full segregation, to end up in an even distribution, divided by the sum of the corresponding proportions for all the groups.

---

[12] As these authors mentioned, the *D* index they developed is equivalent to that proposed earlier by Morgan (1975) and Sakoda (1981). To build overall index *D*, Sakoda (1981) drew inspiration from an expression like $\tilde{D}^g$, although the segregation of a group was not explored. Note that *D* can also be expressed as the I_p index proposed by Silber (1992) divided by its maximum, $D = \frac{I_p}{I_p^*}$, where $I_p^* = \sum_g \frac{C^g}{T} D^{g*}$ is the weighted average of the maximum segregation of the groups.

[13] *H* is the mutual information index, *M*, divided by its maximum ($H = \frac{M}{M^*}$, where $M^* = \sum_g \frac{C^g}{T} \Phi_1^{g*}$).

[14] *C* is the quotient between an unstandardized index based on the squared coefficient of variation, which is here labeled *UC*, and its maximum segregation. Namely, $C = \frac{UC}{UC^*}$, where $UC^* = 2\sum_g \frac{C^g}{T} \Phi_2^{g*}$.

Table 2. Decomposition of standardized overall segregation measures in terms of standardized local segregation measures

Standardized Overall Segregation Measures	Relationship between Standardized Overall and Local Segregation Measures	Weights
$G = \dfrac{1}{2\sum_g \frac{C^g}{T}\left(1-\frac{C^g}{T}\right)} \sum_g \frac{C^g}{T} \sum_{i,j} \frac{t_i}{T}\frac{t_j}{T}\left\| \frac{c_i^g / t_i}{C^g / T} - \frac{c_j^g / t_j}{C^g / T} \right\|$	$G = \sum_g \tilde{w}^g \tilde{G}^g$	$\tilde{w}^g = \dfrac{\tilde{C}^g / T}{\sum_g \tilde{C}^g / T}$; $\tilde{C}^g = C^g\left(1-\frac{C^g}{T}\right)$
$D = \dfrac{1}{2\sum_g \frac{C^g}{T}\left(1-\frac{C^g}{T}\right)} \sum_g \frac{C^g}{T} \left(\sum_j \frac{t_j}{T}\left\| \frac{c_j^g / t_j}{C^g / T} - 1 \right\| \right)$	$D = \sum_g \tilde{w}^g \tilde{D}^g$	$\tilde{w}^g = \dfrac{\tilde{C}^g / T}{\sum_g \tilde{C}^g / T}$; $\tilde{C}^g = C^g\left(1-\frac{C^g}{T}\right)$
$H = \dfrac{1}{\sum_g \frac{C^g}{T}\ln\left(\frac{T}{C^g}\right)} \sum_g \frac{C^g}{T} \left(\sum_j \frac{c_j^g}{C^g} \ln \frac{c_j^g / t_j}{C^g / T} \right)$	$H = \sum_g \widehat{w}^g \tilde{\Phi}_1^g$	$\widehat{w}^g = \dfrac{\widehat{C}^g / T}{\sum_g \widehat{C}^g / T}$; $\widehat{C}^g = C^g \ln\left(\frac{T}{C^g}\right)$
$C = \dfrac{1}{N-1} \sum_g \frac{C^g}{T} \left[\sum_j \frac{t_j}{T} \left(\frac{c_j^g / t_j}{C^g / T} - 1 \right)^2 \right]$	$C = \sum_g \hat{w}^g \tilde{\Phi}_2^g$	$\hat{w}^g = \dfrac{\hat{C}^g / T}{\sum_g \hat{C}^g / T}$; $\hat{C}^g = \frac{C^g}{2}\left[\left(\frac{C^g}{T}\right)^{-1} - 1\right]$

The analysis undertaken thus far allows us to enlarge our knowledge of the measurement of overall segregation. Reardon and Firebaugh (2002) proved that the information theory index, *H*, is the only one of the four overall indices mentioned above that verifies the principle of transfers in a multigroup context (i.e., the only one that decreases when an individual in a group moves to a unit where the group has a lower representation). This is why these authors recommend the use of *H* to measure multigroup overall segregation.[15] However, they also question "whether the violation of the principle of transfers seriously undermines the non-H indices, or instead is of little practical consequence in most research applications" (p. 58). In light of the local segregation approach shown here, that *H* is alone, among these standardized overall indices, in verifying the principle of transfers does not seem too problematic. As we have shown, both *G* and *C* can be generated via standardized local segregation indices satisfying sensitivity to disequalizing movements type III (which is the principle of transfers applied to the segregation of each group). This suggests that, unlike *D*, in the case of *G* and *C*, the violation of that property does not undermine the essence of the principle of transfers. The idea is that, when using *G* and *C*, we cannot ensure that the reduction in overall segregation arising from an equalizing movement of individuals in a group (from one unit to another) does more than offset the possible rise in segregation derived from the impact of the changes in the size of those units on other groups (especially if those groups are highly overrepresented in the unit of origin and underrepresented in the unit of destination). To assume these changes have to reduce overall segregation seems more a value judgment than a requirement that we cannot waive.[16]

3. An Illustration: Occupational Segregation of White Women in U.S. Metropolitan Areas

Using the above tools, we examine the occupational segregation of white women in the largest metropolitan areas in the U.S. We choose this group because it has a large presence in all large metropolitan areas while its demographic weight differs notably across them. Our

[15] Mora and Ruiz-Castillo (2011) instead defend the use of *M* vs. *H* based on decomposability properties.
[16] This rationale can be extended to the corresponding unstandardized overall measures (*UG*, *M*, and *UC*).

analysis focuses only on those cities that allow us to highlight the main similarities and differences between unstandardized and standardized local segregation measures.

We use the 2012-16 American Community Survey (ACS) provided by the IPUMS-USA (Ruggles et al., 2017). We select the 51 metropolitan areas (MAs) with more than 1 million inhabitants (based on the 2010 census). White women are identified on the basis of the information reported by the interviewees about their gender and race/ethnicity, considering only those women who are white and non-Hispanic. The percentage of white female workers ranges between 14.6% in Miami (FL) and 42.3% in Pittsburgh (PA).

Our occupational classification distinguishes among 458 categories, which allows us to measure segregation in a highly precise way. For each MA we calculate 12 local segregation indices (6 unstandardized and 6 standardized): D^g ($\tilde{D}^g$), G^g ($\tilde{G}^g$), and Φ^g_α ($\tilde{\Phi}^g_\alpha$) for $\alpha = 0.1$, 0.5, 1, and 2. For the sake of simplicity, the presentation focuses on indices D^g and $\tilde{D}^g$, referring to the others only when necessary.

Figure 2 shows the segregation of white women in each MA according to D^g and $\tilde{D}^g$.[17] The dotted lines represent the mean values of the indices. Washington D.C. is among the MAs in which white women have the lowest overrepresentation and underrepresentation in occupations, whether we use standardized and unstandardized measures. According to D^g (=0.256), the percentage of white women in Washington D.C. who must switch occupations in order for the group to be evenly distributed is slightly above 25%.[18] On the other hand, $\tilde{D}^g = 0.33$, i.e., the number of white women in this MA who must change occupation represents 33% of all white women who must move in case of maximum segregation. This suggests that the segregation of white women in Washington D.C. is far from reaching its maximum level.[19]

[17] Table A1 in Appendix B provides the corresponding values, together with the share of white women in each MA. Figure A1 shows the other indices.

[18] These findings are analogous to those reported by Alonso-Villar and Del Río (2017b) for an earlier period.

[19] In Washington, D^{g*}=0.77, i.e., if white women were completely segregated, 3 out of 4 would have to change occupations to achieve an even distribution.

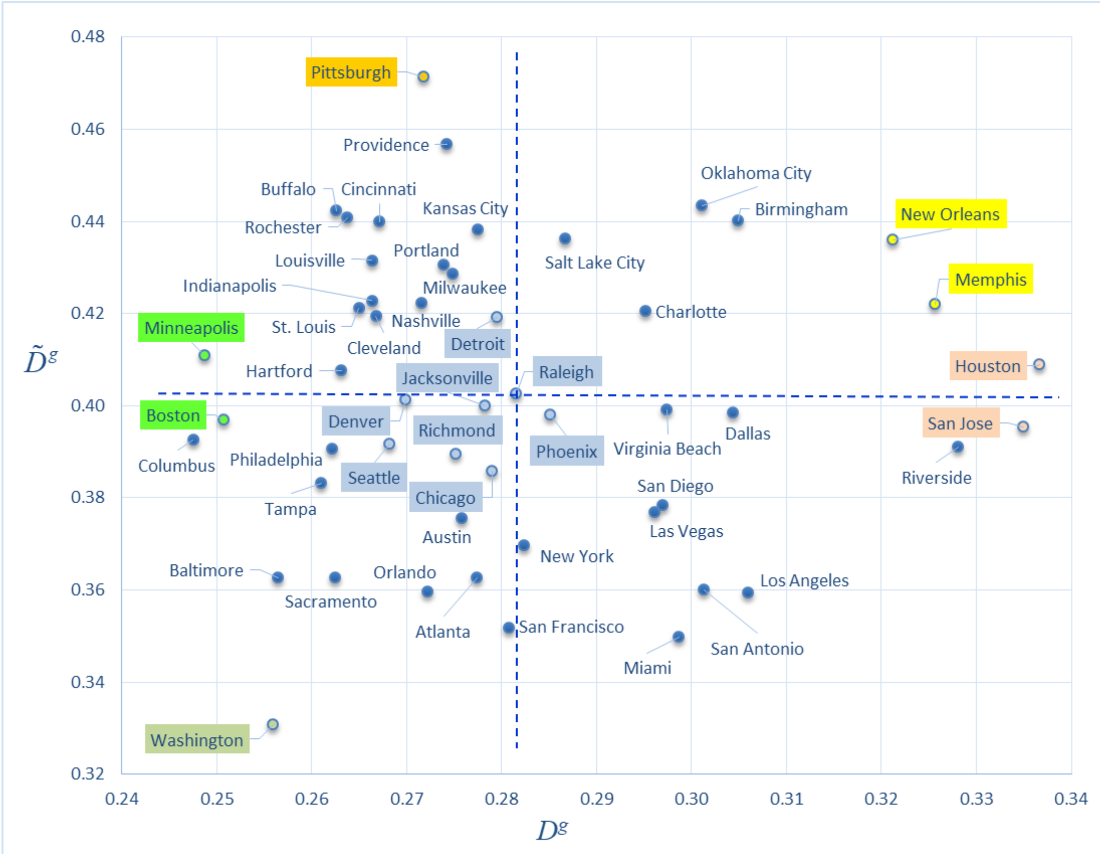

Figure 2. Values of the indices D^g and $\tilde{D}^g$

The remaining indices used in this illustration lead to the same conclusion: Washington D.C. has a low level of segregation. Moreover, Washington has a lower level of segregation than other MAs for the wider range of indices consistent with the dominance criterion provided by the theorem presented in Section 2. Thus, for example, Figure 3 shows that Washington's local segregation curve dominates that of New Orleans, while the opposite obtains for the curves of maximum segregation, thereby ensuring a lower level of segregation for white women in Washington for all the indices consistent with the dominance criterion (standardized or not).[20]

[20] Other large MAs having a similar position in the ranking with indices D^g and $\tilde{D}^g$ include Chicago, Seattle, Denver, Phoenix, and Detroit (see Figure 2). According to most of the (standardized and unstandardized) indices, all these cities have intermediate levels of segregation.

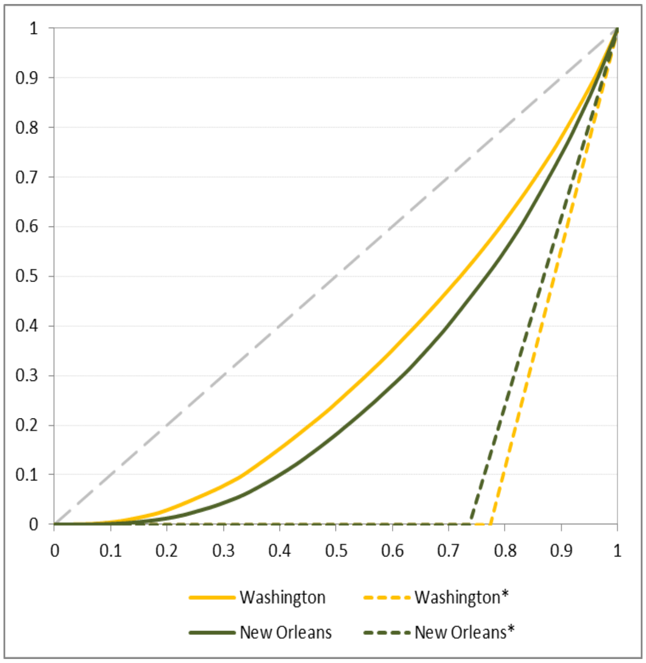

Figure 3. Local segregation curves (actual and maximum), Washington and New Orleans

Boston and Minneapolis exhibit a different pattern (Figure 2). They share with Washington D.C. a low segregation value according to D^g (=0.25). However, this figure represents around 40% of the maximum value of the index, which means these cities have an intermediate rather than a low position in the ranking based on standardized measure $\tilde{D}^g$. How do we interpret this? On the one hand, D^g shows that the three MAs have something in common: 1 out of 4 white women working there must change occupation for this group to have in each occupation the same weight it has in the corresponding MA. On the other hand, $\tilde{D}^g$ allows us to take a step further by accounting for another dimension, the demographic size of the group under consideration; this reveals that segregation is a more acute phenomenon in Boston and Minneapolis than it is in Washington D.C. This is so because the 25% of white women requiring occupation changes to achieve no segregation represents a higher proportion of total workers (or jobs) in the labor markets of the former cities.[21] Thus, the number of white women who must move to another occupation in Washington D.C., in order for the group to have zero segregation, represents around 6% of the city workers, whereas in Minneapolis and Boston this figure is higher, at nearly 10%.

[21] White women account for a larger share of the population in Boston and Minneapolis than in Washington D.C., while Hispanic and, especially, African American population have a remarkable presence.

By contrast, although San Jose and Houston have the highest levels of segregation according to D^g, they have an intermediate value with $\tilde{D}^g$.[22] The extraordinarily low proportion of white women workers in these cities (15.3% and 17.7%, respectively) allows them to have a standardized value similar to the one they have in Boston and Minneapolis (although they have a very different unstandardized segregation level).

New Orleans and Memphis represent cases that stand in opposition to Washington because they have high levels of segregation regardless of the approach followed. Moreover, this is so although the three cities have a similar share of white women workers: 22.6% in Washington, 22.8% in Memphis, and 26.3% in New Orleans. This suggests that demographic size does not, in isolation, determine the segregation level of a group.

Pittsburgh stands out as a paradigmatic case. The relatively low value of index D^g (=0.27) in this area represents almost half of the maximum segregation attainable by the group.[23] Pittsburgh is therefore the MA with the highest standardized segregation of the country according to index $\tilde{D}^g$ (=0.47). The 27% of white women who must change occupations to achieve an even distribution represents 11.5% of all the workers in the city.

To verify the robustness of this result, we estimate the local segregation curves of Pittsburgh and those of other MAs to explore whether there is a relationship of dominance between them. As an example, Figure 4 compares the curves of local segregation and maximum local segregation for Pittsburgh with those for New Orleans, thereby illustrating that there is no dominance relationship in the terms of the mentioned theorem.

[22] Something similar happens when comparing indices G^g and $\tilde{G}^g$ (see Figure A1 in the Appendix B).

[23] Remember that white women here account for 42.3% of the workers (the highest share in our selected MAs).

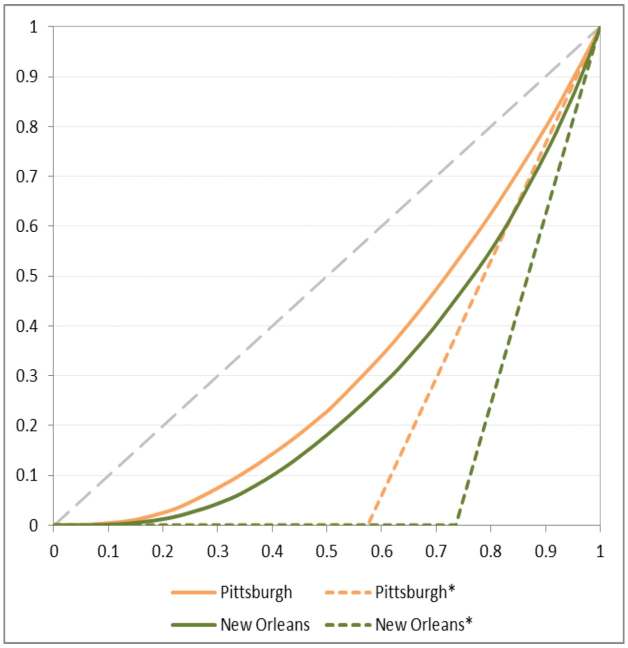

Figure 4. Local segregation curves (actual and maximum), Pittsburgh and New Orleans

The segregation curve of white women in Pittsburgh dominates that of white women in New Orleans, which implies that not only index D^g but all the indices consistent with that dominance criterion (as is the case of G^g and Φ_α^g) will conclude that white women have a lower level of segregation in Pittsburgh. However, the curve of maximum segregation for Pittsburgh also dominates that of New Orleans (white women have a lower weight in the latter), which thereby demonstrates the impossibility of dominance in terms of standardized segregation between these two cities. In other words, we cannot state that Pittsburgh has a higher level of standardized segregation than New Orleans based on the wide set of indicators consistent with the dominance criterion established in Section 2. Neither can we assert the opposite. In fact, 4 out of the 6 standardized indices that we have calculated ($\tilde{D}^g$, $\tilde{G}^g$, $\tilde{\Phi}_1^g$, and $\tilde{\Phi}_2^g$) indicate that Pittsburgh has a higher level of standardized segregation than New Orleans because, according to them, the local segregation curve of the former is “closer” to that of maximum segregation. However, two indices whose value judgements are more extreme ($\tilde{\Phi}_{0.1}^g$ and $\tilde{\Phi}_{0.5}^g$) attribute a greater standardized segregation to New Orleans.[24]

[24] This is especially true in the case of $\tilde{\Phi}_{0.1}^g$ (see Figure A1 in the Appendix B) because $\Phi_{0.1}^g$ focuses on the intensity of underrepresentation at an extraordinarily high level. That white women in New Orleans are virtually

4. Final comments

What can we conclude in light of the results presented in the previous section? Are white women in Pittsburgh highly concentrated in some occupations (as most standardized indices suggest) or is the segregation of this group below average and, especially, smaller than in New Orleans (as the unstandardized indices display)? The answers to these questions depend on what we mean by segregation. If we think of a group's segregation as the extent to which its occupational distribution departs from an even distribution—where the group accounts for the same share of workers in each occupation as it does in the MA—we would say that Pittsburgh exhibits an intermediate-low level, whereas New Orleans is among the MAs with the highest values. If, instead, we are interested in determining how close the group is to its maximum concentration in occupations, which depends on the group's relative size, Pittsburgh is the MA with the highest level of segregation, as several standardized indices display.

As this research has shown, the standardized local segregation indices developed here have several desirable properties and are consistent with existing standardized overall segregation indices, given that the latter can be written as the weighted average of the standardized local segregation of the groups involved. Although all of these standardized (overall and local) segregation indices are, like most unstandardized indices, margin-dependent, they result from a normalization that ensures the index is equal to 1 when there is maximum segregation. This allows overcoming the limitation of most unstandardized indices, the maximum values of which depend on the relative size of the groups under study. The cost of using standardized measures is that they focus on the "proximity" to a maximum segregation level, which depends on demographic composition, whereas segregation from an evenness perspective is usually viewed as separation from an egalitarian distribution, the segregation level of which is always equal to 0.

For this reason, the standardized local segregation indices developed here are not proposed as an alternative to existing local segregation indices, but as a complementary tool to explore

absent from occupations that account for 6% of total employment implies that the value of $\Phi^{g}_{0.1}$ in this MA is almost double that in Pittsburgh (where occupations with no white women represent less than 2% of total employment).

segregation from a different angle. In our opinion, the standardized local (respectively, overall) indices can be especially useful in empirical studies that involve segregation comparisons among groups (respectively, societies) of highly distinct relative sizes (respectively, composition).

References:

Alonso-Villar, O. and Del Río, C. (2010): "Local versus Overall Segregation Measures," *Mathematical Social Sciences* 60, 30-38.

Alonso-Villar, O. and Del Río, C. (2017a): "Occupational Segregation and Well-being," *Review of Income and Wealth* 63(2), 269-287.

Alonso-Villar, O. and Del Río, C. (2017b): "Mapping the Occupational Segregation of White Women in the US: Differences across Metropolitan Areas," *Papers in Regional Science* 96(3), 603-625.

Boisso, D, Hayes, K., Hirschberg, J. and Silber, J. (1994): "Occupational Segregation in the Multidimensional Case. Decomposition and Tests of Significance," *Journal of Econometrics* 61, 161-171.

Del Río C. and Alonso-Villar, O. (2015): "The Evolution of Occupational Segregation in the U.S., 1940-2010: The Gains and Losses of Gender-Race/Ethnic Groups," *Demography* 52 (3), 967-988.

Frankel, D.M. and Volij, O. (2011): "Measuring School Segregation," *Journal of Economic Theory* 146(1), 1-38.

James, D.R. and Taeuber, K.E. (1985): "Measures of segregation," *Sociological Methodology* 15, 1-32.

Moir, H. and Selby Smith, J. (1979): "Industrial Segregation in the Australian Labour Market," *Journal of Industrial Relations* 21, 281-291.

Mora, R. and Ruiz-Castillo, J. (2011): "Entropy-based Segregation Indices," *Sociological Methodology* 63(2), 269-287.

Morgan, B. (1975): "The Segregation of Socioeconomic Groups in Urban Areas: A Comparative Analysis," *Urban Studies* 12, 47-60.

Reardon, S.F. and Firebaugh, G. (2002): "Measures of Multigroup Segregation," *Sociological Methodology* 32, 33-76.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2017): Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis, MN: University of Minnesota.

Sakoda, J. (1981): "A Generalized Index of Dissimilarity," *Demography* 18, 245-250.

Silber, J. (1992): "Occupational Segregation Indices in the Multidimensional Case: A Note," *The Economic Record* 68, 276-277.

Theil, H. and Finizza, A. (1971): "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts," *Journal of Mathematical Sociology* 1, 187-194.

Appendix A

Maximum values of the indices. Making use of the graphical interpretation of G^g, it is clear that G^{g^*} is twice the difference between the area of the triangle of base 1 and that of the triangle of base $\frac{C^g}{T}$ (see Figure 1 and note that $2\left(\frac{1}{2}-\frac{C^g}{2T}\right)=1-\frac{C^g}{T}$). Analogously, taking into account that the 45º line represents points at which the ordinate is equal to the abscissa, the highest vertical distance of the curve is equal to the length of the horizontal part of the curve. In other words, $D^{g^*}=1-\frac{C^g}{T}$.

To obtain $\Phi_\alpha^{g^*}$ $(\alpha \neq 0,1)$, note that if the group is fully segregated

$$\Phi_\alpha^g=\frac{1}{\alpha(\alpha-1)}\left(1-\frac{C^g}{T}\right)(-1)+\frac{1}{\alpha(\alpha-1)}\frac{C^g}{T}\left[\left(\frac{1}{C^g/T}\right)^{\alpha}-1\right]=\frac{1}{\alpha(\alpha-1)}\left[\left(\frac{C^g}{T}\right)^{1-\alpha}-1\right].$$

Likewise, it is straightforward to see that $\Phi_1^{g^*}=\ln\left(\frac{T}{C^g}\right)$ since $\lim\limits_{c_j^g\to 0}\frac{c_j^g}{C^g}\ln\left(\frac{c_j^g/C^g}{t_j/T}\right)=0$.

Proof of Proposition 1. Assume that i and h are two units such that $\frac{c_i^g}{t_i}<\frac{c_h^g}{t_h}$. Taking into account that the local segregation index $\Theta^g\left(c^g,t\right)$ satisfies *insensitivity to proportional divisions*, the segregation of group g remains the same if units i and h are split into t_i and t_h subunits (of size 1 each), where the former subunits each account for $\frac{c_i^g}{t_i}$ "individuals" of group g and the latter for $\frac{c_h^g}{t_h}$.

If $\frac{d}{t_i t_h}$ "people" of group g leave one of the subunits of unit i to move to one of the subunits of unit h, the segregation of group g will increase, given that the two subunits have the same size and the index satisfies the property of *disequalizing movements type I*. Reiterating this for all other subunits of h, we will have a sequence of t_h *disequalizing movements type I* between units of the same size, which leads to a higher segregation for the group (a total of $\frac{d}{t_i}$ individuals of group g are moving from a subunit of unit i to unit h). If we repeat this process for any other subunit of unit i, eventually, $t_i \frac{d}{t_i} = d$ individuals will have switched from i to h.

In other words, a transfer of d individuals of group g from i to h, which does not alter the size of these units,[25] can be expressed as a sequence of *disequalizing movements type I* between units of the same size, which signifies a rise in the level of segregation of the group. Once more employing the *insensitivity to proportional divisions*, the segregation of the group is the same in the case of either having these small subunits or aggregating them to give rise to units i and h, which completes the proof.

Proof of Proposition 2. Assume that i and h are two units such that $\frac{c_i^g}{t_i} < \frac{c_h^g}{t_h}$ and that d individuals ($d < c_i^g$) are transferred from i to h without replacement, i.e., $c_i^g{}' = c_i^g - d$, $c_h^g{}' = c_h^g + d$, $t_i{}' = t_i - d$, and $t_h{}' = t_h + d$ (no changes in the other units, i.e., $c_j^g{}' = c_j^g$ and $t_j{}' = t_j$ for $j \neq i, h$). To facilitate the graphical analysis, let us assume, without loss of generality, that i is the unit in which the group has the lowest representation and h is the next unit in the ranking (see Figure A1).

First, we prove that, at point $\frac{t_i - d}{T}$, the post-transfer curve is below the other (the situation depicted in the chart), making use of simple trigonometric analysis.[26] To this end, we need

[25] This implies that an equal number of individuals from other groups has moved in the opposite direction.

[26] If $d = c_i^g$, it is trivial to prove that the curve after the transfer, which is equal to zero up to point $\frac{t_i - d}{T}$, is below the other.

only prove that $\tan(\alpha) > \tan(\beta)$. Note that $\tan(\alpha) = \dfrac{\frac{c_i^g}{C^g}}{\frac{t_i}{T}}$ and $\tan(\beta) = \dfrac{\frac{c_i^g - d}{C^g}}{\frac{t_i - d}{T}}$. It is straightforward to see that $\tan(\alpha) > \tan(\beta) \Leftrightarrow t_i > c_i^g$. Given that in unit i the group's representation is below that in h, then $\frac{c_i^g}{t_i} < 1$ and, consequently, $\tan(\alpha) > \tan(\beta)$.

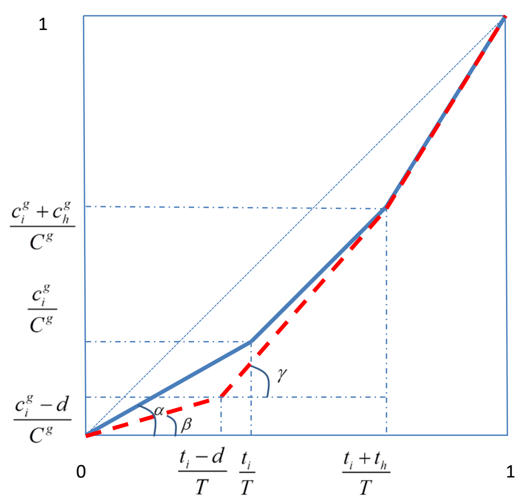

Figure A1. The segregation curve before (solid line) and after transfers (dash line)

Second, we must show that, at point $\frac{t_i}{T}$, the curve after the transfer is below (or equal to) the other. If we denote by x the difference between the curve after the transfer at point $\frac{t_i}{T}$ and $\frac{c_i^g - d}{C^g}$, then $\tan(\gamma) = \dfrac{x}{\frac{d}{T}} = \dfrac{\frac{c_h^g + d}{C^g}}{\frac{t_h + d}{T}}$. Therefore, $x = \dfrac{d}{d + t_h} \dfrac{c_h^g + d}{C^g}$. To finish the proof we need only demonstrate that $\dfrac{c_i^g - d}{C^g} + x \le \dfrac{c_i^g}{C^g}$. It is easy to see that this inequality holds because the number of individuals of the group in unit h is always lower than or equal to the total number of individuals in that unit ($c_h^g \le t_h$).

Proof of Corollary 1. As shown by Alonso-Villar and Del Río (2010), G^g and Φ_α^g satisfy *scale* and *size invariance*, *symmetry*, *insensitivity to proportional divisions*, and *sensitivity to disequalizing movements type I*. That these indices satisfy *sensitivity to disequalizing movements type II* follows from Proposition 1. As these indices are consistent with the dominance criterion given by the local segregation curve (Alonso-Villar and Del Río, 2010), it follows from Proposition 2 that they satisfy *sensitivity to disequalizing movements type III*. The properties of index D^g follow immediately from its definition.

Proof of Corollary 2. This follows straightforwardly from the fact that the unstandardized versions of these indices satisfy the corresponding properties and the standardized indices are obtained through the former by dividing them by a constant.

Proof of Theorem. If the local segregation curve in case A dominates that in case B, then any local segregation index $\Theta^g\left(c^g,t\right)$ that satisfies *scale invariance*, *symmetry*, *insensitivity to proportional divisions*, and *sensitivity to disequalizing movements type I* will have a lower value in case A than in B (Alonso-Villar and Del Río, 2010; Theorem 1). For the same reason, the value of the index in case of maximum segregation, Θ^{g^*}, is higher in scenario B than in A given that the curve of the former dominates that of the latter. Therefore, $\tilde{\Theta}^g\left(c^g,t\right)=\dfrac{\Theta^g\left(c^g,t\right)}{\Theta^{g^*}}$ is higher in A than in B, which completes the proof.

Proof of Proposition 3. As explained above, for a group with a population share of $\dfrac{C^g}{T}$, the curve of maximum segregation is equal to 0 up to an abscise value of $1-\dfrac{C^g}{T}$ and after that point the curve is a straight line that links points $\left(1-\dfrac{C^g}{T},0\right)$ and $\left(1,1\right)$. Therefore, if one group has a larger share of the population than another group, the curve of maximum segregation will be equal to 0 up to a point that is lower than that of the other group and after that point the curve will be above the other (by construction of the curve, see Figure 1). This means that the curve of the larger group dominates that of the smaller.

We prove the other implication by proof by contradiction. Assume that the local segregation curve of a group dominates that of another group, but that the former group has a lower share of the population than the latter. If the group in the first case has a lower share, this means that the curve of maximum segregation is equal to 0 up to an abscissa greater than that of the other group and that its curve is thereby dominated by the other (i.e., is below than or equal to the other curve), which contradicts the assumption.

Appendix B

Table A1. Population share of white women and indices D^g and $\tilde{D}^g$ in each MA

	Segregation indices			Population share of white women
Metropolitan Areas ranked by D^g	D^g	$\tilde{D}^g$	D^{g*}	
Columbus, OH	0.2475	0.3926	0.6304	37.0
Minneapolis-St. Paul-Bloomington, MN-WI	0.2488	0.4109	0.6055	39.4
Boston-Cambridge-Newton, MA-NH	0.2508	0.3968	0.6321	36.8
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.2559	0.3308	0.7736	22.6
Baltimore-Columbia-Towson, MD	0.2565	0.3626	0.7074	29.3
Tampa-St. Petersburg-Clearwater, FL	0.2610	0.3832	0.6811	31.9
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.2622	0.3907	0.6711	32.9
Sacramento--Roseville--Arden-Arcade, CA	0.2625	0.3625	0.7242	27.6
Buffalo-Cheektowaga-Niagara Falls, NY	0.2626	0.4423	0.5936	40.6
Hartford-West Hartford-East Hartford, CT	0.2631	0.4076	0.6455	35.4
Rochester, NY	0.2638	0.4408	0.5984	40.2
St. Louis, MO-IL	0.2650	0.4211	0.6292	37.1
Louisville/Jefferson County, KY-IN	0.2663	0.4316	0.6171	38.3
Indianapolis-Carmel-Anderson, IN	0.2664	0.4228	0.6301	37.0
Cleveland-Elyria, OH	0.2668	0.4194	0.6360	36.4
Cincinnati, OH-KY-IN	0.2672	0.4400	0.6073	39.3
Seattle-Tacoma-Bellevue, WA	0.2682	0.3917	0.6847	31.5
Denver-Aurora-Lakewood, CO	0.2699	0.4012	0.6728	32.7
Nashville-Davidson--Murfreesboro--Franklin, TN	0.2716	0.4224	0.6428	35.7
Pittsburgh, PA	0.2718	0.4713	0.5767	42.3
Orlando-Kissimmee-Sanford, FL	0.2722	0.3595	0.7572	24.3
Portland-Vancouver-Hillsboro, OR-WA	0.2739	0.4305	0.6364	36.4
Providence-Warwick, RI-MA	0.2742	0.4568	0.6003	40.0
Milwaukee-Waukesha-West Allis, WI	0.2749	0.4287	0.6412	35.9
Richmond, VA	0.2752	0.3895	0.7065	29.4
Austin-Round Rock, TX	0.2759	0.3754	0.7349	26.5
Atlanta-Sandy Springs-Roswell, GA	0.2774	0.3626	0.7651	23.5
Kansas City, MO-KS	0.2775	0.4381	0.6335	36.6
Jacksonville, FL	0.2782	0.4000	0.6956	30.4
Chicago-Naperville-Elgin, IL-IN-WI	0.2790	0.3857	0.7233	27.7
Detroit-Warren-Dearborn, MI	0.2795	0.4191	0.6671	33.3
San Francisco-Oakland-Hayward, CA	0.2809	0.3516	0.7989	20.1
Raleigh, NC	0.2816	0.4025	0.6996	30.0
New York-Newark-Jersey City, NY-NJ-PA	0.2824	0.3695	0.7643	23.6
Phoenix-Mesa-Scottsdale, AZ	0.2851	0.3979	0.7165	28.3
Salt Lake City, UT	0.2867	0.4362	0.6573	34.3
Charlotte-Concord-Gastonia, NC-SC	0.2952	0.4204	0.7021	29.8
Las Vegas-Henderson-Paradise, NV	0.2962	0.3768	0.7861	21.4
San Diego-Carlsbad, CA	0.2970	0.3782	0.7853	21.5
Virginia Beach-Norfolk-Newport News, VA-NC	0.2974	0.3990	0.7453	25.5
Miami-Fort Lauderdale-West Palm Beach, FL	0.2987	0.3497	0.8540	14.6
Oklahoma City, OK	0.3011	0.4434	0.6791	32.1
San Antonio-New Braunfels, TX	0.3013	0.3600	0.8370	16.3
Dallas-Fort Worth-Arlington, TX	0.3044	0.3984	0.7640	23.6
Birmingham-Hoover, AL	0.3050	0.4401	0.6930	30.7
Los Angeles-Long Beach-Anaheim, CA	0.3059	0.3594	0.8513	14.9
New Orleans-Metairie, LA	0.3212	0.4360	0.7367	26.3
Memphis, TN-MS-AR	0.3257	0.4220	0.7718	22.8
Riverside-San Bernardino-Ontario, CA	0.3281	0.3911	0.8389	16.1
San Jose-Sunnyvale-Santa Clara, CA	0.3350	0.3954	0.8472	15.3
Houston-The Woodlands-Sugar Land, TX	0.3366	0.4090	0.8231	17.7

Figure A1. Standardized versus unstandardized local segregation indices in each MA