Universidade Vigo

Departamento de Economía Aplicada
https://webx06.webs.uvigo.es

Working Paper: 24/02. December 2024

From the Extent of Segregation to Its Consequences in Terms of Wellbeing: A Methodological Reflection With an Application to the Spanish Labor Market

Olga Alonso-Villar
Coral del Río

Coordinator: Olga Alonso-Villar
ovillar@uvigo.es

From the Extent of Segregation to Its Consequences in Terms of Wellbeing: A Methodological Reflection With an Application to the Spanish Labor Market*

Olga Alonso-Villar# and Coral del Río♦

ECOBAS, Universidade de Vigo

Abstract

We offer a reflection on the measurement of segregation, gathering methodological contributions from sociology and economics, and we use some of them to explore occupational segregation by gender and nativity in Spain. Our goal is to offer a guide to the tools that can be used in empirical analysis, connecting them with theoretical discussions. Our empirical analysis shows that the occupational segregation of immigrant women is a more intense phenomenon than that of native women or immigrant men, although it decreased significantly over the period 2006-2024. Unlike their male peers, occupational sorting strongly penalizes immigrant women after controlling for characteristics.

JEL Classification: D63, J15, J16, J31

Keywords: Segregation, gender, migration status, wage gaps, intersectionality

---

* This research was supported by grants PID2020-113440GB-100, and PID2022-137352NB-C42 funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe." This work will appear in *Hacienda Pública Española / Review of Public Economics*, special issue of the XXI *Spanish Meeting on Public Economics*.

Correspondence address: ECOBAS, Universidade de Vigo, Departamento de Economía Aplicada, Fac. de CC. Económicas e Empresariais, Campus Lagoas-Marcosende s/n, 36310 Vigo, Spain. E-mail: ovillar@uvigo.es. Orcid ID: 0000-0002-7355-0279.

♦ ECOBAS, Universidade de Vigo. E-mail: crio@uvigo.es. Orcid ID: 0000-0002-3173-3624.

1. Introduction

The measurement of segregation has a long tradition in sociology, with approaches linked to the measurement of income inequality in economics and others developed outside this field (James and Taeuber, 1985; Massey and Denton, 1988; Charles and Grusky, 1995; Reardon and Firebaugh, 2002). As in the income inequality literature, this scholarship has followed an axiomatic approach because there is general agreement on the desirability of defining the basic properties that any segregation measure must meet. It is within sociology that these properties have been discussed to the greatest extent, especially to quantify overall or aggregate segregation, i.e., the simultaneous discrepancies among all the mutually exclusive groups into which society had been partitioned across organizational units (whether these units are occupations, industries, workplaces, schools, neighborhoods of a city, etc.).

For a long time, this literature focused on the segregation between two groups (e.g., occupational segregation by gender, residential segregation between Black and White households, and school segregation between pupils living in poverty and those who are not) using binary segregation measures (Duncan and Duncan, 1955; Karmel and MacLachlan, 1988). However, as societies grew more diverse, it became necessary to develop measures with which to quantify overall segregation in a multigroup context (e.g., residential segregation among Whites, Blacks, and Hispanics and occupational segregation by both gender and race/ethnicity; Reardon and Firebaugh, 2002).

The contributions of economics to this field are various. They include, inter alia, the development of new binary measures (Silber, 1989; Hutchens, 1991, 2004), the extension of binary measures to a multigroup context (Boisso et al., 1994; Silber, 1992), decomposition methods (Deutsch et al., 2009; Frankel and Volij, 2011), and the development of an axiomatic framework to deal with the measurement of the segregation of a group in a multigroup context, called local segregation to distinguish it from overall segregation (Alonso-Villar and Del Río, 2010a; Del Rio and Alonso-Villar, 2022), which has opened the door to explore the consequences of segregation in terms of wellbeing (Alonso-Villar and Del Río, 2017a).[1]

Throughout these pages, we offer a reflection on the measurement of segregation, gathering methodological contributions from both sociology and economics, and we use some of them to explore occupational segregation by gender and nativity in Spain. Our aim is not to provide a

[1] Silber (2012) discusses how to apply the measurement of segregation to other domains (inequality in life chances, inequality in happiness, and inequality in health).

review of the extensive literature on segregation but to offer a guide to the tools that can be used in empirical analysis, connecting them with theoretical discussions.

First, we briefly present the measurement of income inequality and show why and how measuring segregation requires moving away from it. Second, focusing on segregation from an evenness perspective, we provide both indicators that allow us to quantify overall segregation between two groups (e.g., women and men), as well as those that measure overall segregation in a multigroup context (e.g., segregation by both gender and nativity) and those with which to measure the segregation of a group in a multigroup context (e.g., segregation of immigrant women). Third, we delve deeper into what the local segregation approach brings to the analysis. We show how the segregation of a group relates to overall segregation, and we address the wellbeing consequences of segregation for the incumbent groups (i.e., we assess the sorting of the groups across organizational units). Fourth, we discuss the methodological approaches that the literature has used to explain why segregation in some economies is greater than it is in others, allowing us to explain differences across time and space.

Finally, we use several of these measures to explore the extent of occupational segregation by gender and nativity in Spain and its consequences in monetary terms for the incumbent groups. Drawing on the Spanish Labor Force Survey (*Encuesta de Población Activa*, EPA henceforth), we document the evolution of the phenomenon for the period 2006-2024. Using the most recent data sets, the 2024 EPA and the 2022 Structure of Earnings Survey (*Encuesta de Estructura Salarial*, EES henceforth), we also explore the role that occupations play in explaining the gender-native wage gaps, both before and after controlling for characteristics.

2. From Measuring Income Inequality to Segregation

Many of the segregation measures that exist in the literature, together with the properties required to them, are borrowed from the literature on income inequality. In this section, we briefly discuss how scholarship has addressed the measurement of income inequality and then we deal with the measurement of segregation. This will allow us to clarify which views of inequality have dominated the literature on segregation.

There is general agreement on how to rank income distributions that have the same total income in terms of inequality. The basic properties that any inequality measure should satisfy are *symmetry* and the *Pigou-Dalton principle of transfers*.[2] The former means that inequality does

[2] Additional properties, of a more technical nature, are often invoked as well (e.g., *normalization*, *continuity*, *differentiability*, and the *population principle*).

not depend on who enjoys each income level, which brings anonymity to the measurement. The latter requires that inequality decrease when there is a progressive transfer without re-ranking between two individuals (i.e., when a richer person gives money to a poorer person in such a way that the initially poorer person does not end up with more income than the other). However, in empirical analysis, we usually need to compare income distributions with different means. This implies that some value judgments must be incorporated into the measurement because there is no consensus about the mean-invariance property the indices must meet.

Following a relative approach, which is the most common path in empirical research, we say that, when income increases, inequality does not change if the additional income is distributed across individuals according to the share of income that individuals had in the initial distribution (*scale invariance* axiom). And, therefore, relative inequality increases when individuals with higher incomes receive higher proportions of the extra income than they had initially. The absolute approach requires instead that inequality remain unaltered when the extra income is distributed in equal amounts among all individuals (*translation invariance* axiom). Consequently, absolute inequality increases when individuals with higher incomes receive higher amounts of the extra income. Kolm (1976) labeled relative and absolute inequality measures as rightist and leftist measures, respectively, and proposed an intermediate or centrist view between these two extremes.[3]

When using intermediate and absolute inequality measures, an additional axiom has been recently invoked, the *unit-consistency* axiom. It requires that the ranking between two distributions does not depend on the currency unit in which those incomes are expressed (Zheng, 2007), a criterion that all relative measures satisfy. This requirement allows us to incorporate centrist and leftist views into the inequality measurement without compromising its empirical usefulness.

The literature provides a wide set of relative inequality measures satisfying good properties. For example, the Lorenz curves and the indices consistent with the dominance criterion given by these curves, as is the popular Gini index, the Atkinson family of indices, or the generalized entropy family of indices (which includes the Theil indices).[4] Thus, when the Lorenz curve of an income distribution A is never below that of distribution B (i.e., A is equal to B at some

---

[3] For intermediate inequality measures, see Krtscha (1994), Seidl and Pfingsten (1997), Del Río and Ruiz-Castillo (2000), and Del Río and Alonso-Villar (2010). Intermediate inequality approaches have been used to measure tax progressivity (Pfingsten,1986; Besley and Preston, 1988; Moyes, 1992; Ledic´et al., 2023).

[4] The literature also provides absolute Lorenz-type curves (Moyes, 1987) and various intermediate Lorenz-type curves (Yoshida, 2005; Azpitarte and Alonso-Villar, 2014).

points and is above it at others), we know that all these indices would rank these two distributions as the Lorenz criterion does (Foster, 1985): A is more egalitarian than B.

To measure segregation, scholarship has followed this relative view of inequality and has adapted it to the needs of the new context, which has required incorporating additional properties. However, before discussing how the above properties have been moved to the new scenario, it is necessary to clarify what segregation means. Reskin (1999, p. 183) argues that "segregation is a social mechanism that preserves inequality among groups." Reardon and O'Sullivan (2004, p. 122) point out that "segregation can be thought of as the extent to which individuals of different groups occupy and experience different social environments." But how can we operationalize it? Segregation can be seen from different angles, but evenness is the most popular one (Reardon and Firebaugh, 2002).[5] According to this perspective, segregation exists when the mutually exclusive groups into which society has been partitioned are unevenly distributed across organizational units, i.e., when some groups are underrepresented in some units and overrepresented in others.

The measurement of segregation involves important changes with respect to that of income inequality.[6] First, the analysis is not undertaken at the individual level because the subject of interest is the group. Second, the egalitarian distribution is not one in which all individuals have the same income, but rather the distribution (among units) of a reference population with which the distributions of other groups are compared. In the case of occupational segregation by gender, for example, this reference can be the occupational sorting of the total population or the occupational sorting of men. Third, to build Lorenz-type curves with which to quantify segregation, the units must be ranked from those in which the target group is underrepresented to those in which it is overrepresented (rather than ranking individuals from low to high income). And what under and overrepresentation means varies depending on whether we are measuring overall segregation or local segregation, as we show later. Fourth, to compare different economies, we must specify how segregation measures should behave when an organizational unit is split in several units, which gives rise to the *organizational equivalence* axiom (in the case of overall segregation; James and Taeuber, 1985) and the *insensitivity to*

[5] In the case of residential segregation, Massey and Denton (1988) distinguish among evenness, clustering, concentration, centralization, and exposure. Reardon and O'Sullivan (2004) propose grouping them into spatial exposure (as opposed to isolation) and spatial evenness (as opposed to clustering). In the case of school segregation, Frankel and Volij (2011) differentiate between evenness and representativeness.

[6] For a discussion about basic properties in the case of (overall) segregation between two groups, see James and Tauber (1985) and Hutchens (1991, 2001, 2004). Reardon and Firebaugh (2002) articulate them to measure multigroup overall segregation, and Alonso-Villar and Del Río (2010a) and Del Río and Alonso-Villar (2022) develop the properties required to measure local segregation.

proportional divisions axiom (in the case of local segregation; Alonso-Villar and Del Río, 2010a). Fifth, the translation of the Pigou-Dalton transfers principle has led to two different versions in the case of overall segregation (*transfers* axiom and *exchanges* axiom) and three versions in the case of local segregation (*sensitivity to disequalizing movements type I, II,* and *III*).[7] This is so because when a member of a group moves from one unit to another, this has an effect on the sizes of the units and the representations of the other groups and, therefore, it is necessary to specify the circumstances under which we expect the index to increase. We will come back to this discussion in Section 4.

3. Overall Segregation: Binary and Multigroup Measures

3.1 Binary Measures

The dissimilarity index put forward by Janh et al. (1947) and popularized by Duncan and Duncan (1955), who assessed it in terms of basic properties, is the segregation measure most widely used in empirical work, despite the criticisms it has received. In the case of occupational segregation by gender, this index can be expressed as:

$$D = \frac{1}{2}\sum_j \left| \frac{c_j^f}{C^f} - \frac{c_j^m}{C^m} \right|, \tag{1}$$

where j stands for organizational units (e.g., occupations), c_j^f and c_j^m are the number of female and male workers, respectively, in unit j, and C^f and C^m are the total number of females and males. This index is connected to a Lorenz-type curve (Figure 1), called the segregation curve.

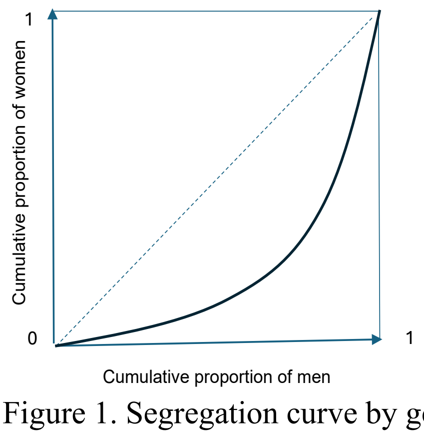


Figure 1. Segregation curve by gender

To build this curve (Duncan and Duncan, 1955), first, the units have to be ranked in ascendent order by the ratio $\frac{c_j^f}{c_j^m}$ and, then, the cumulative proportion of women in those units is plot on

[7] The corresponding definitions can be found in Del Río and Alonso-Villar (2022).

the vertical axis against the cumulative proportion of men in the same units on the horizontal axis. This means that we first look at what happens in the first unit, then we move to the first two units, then to the first three units, and so on. The dissimilarity index measures the highest vertical distance between the curve and the 45º line.[8]

The *D* index ranges between 0 (achieved when there is no segregation, i.e., when the proportion of men and women in each occupation equals their population shares in the total) and 1 (achieved when there is full segregation, i.e., when women work in occupations with no men and reciprocally). This makes this index very convenient to compare different scenarios. However, *D* is not consistent with the dominance criterion given by the segregation curves, so that when the curve of an economy is above another (i.e., when the former is closer to the 45º line than the latter), the index does not necessarily rank these two economies in the same way.[9] The reason is that this index does not satisfy the Pigou-Dalton *transfers principle* adapted to this context (i.e., the index may not decrease when a member of a group moves from a unit to another in which the group has a lower representation), which is one of its shortcomings (James and Taeuber, 1985).[10]

The dissimilarity index can be interpreted as the percentage of women (or men) who would have to move from one occupation to another to have the same occupational sorting as the other gender. This means that the reference against which the occupational sorting of women (men) is compared is that of men (women), which is not the occupational structure of the economy. To avoid this problem, inter alia, Karmel and MacLachlan (1988) proposed the I_p index:

$$I_p = \frac{C^m}{T}\frac{1}{2}\sum_j \left|\frac{c_j^m}{C^m} - \frac{t_j}{T}\right| + \frac{C^f}{T}\frac{1}{2}\sum_j \left|\frac{c_j^f}{C^f} - \frac{t_j}{T}\right|, \tag{2}$$

where T is the total number of workers in the economy (i.e., $T = C^f + C^m$) and t_j is the number of workers in occupation j ($t_j = c_j^f + c_j^m$).

This index can be seen as a modified version of the dissimilarity index, $I_p = 2(C^m / T)(C^f / T)D$, which satisfies the transfers principle. The I_p index can be interpreted as the percentage of workers who would have to swift among occupations to ensure that there

[8] In the inequality literature, the corresponding index is the Pietra index. It represents the income share that individuals with incomes above the mean must transfer to those below the mean to achieve the egalitarian distribution (Sarabia and Jordá, 2014). This index is also called the relative mean deviation (Atkinson, 1970).
[9] The dominance criterion is analogous to that established with the Lorenz curves (Hutchens, 1991).
[10] Other criticisms arise from the fact that the index is dependent on the population shares of the groups, what is called in the literature margin dependency (Charles and Grusky, 1995), and whether the benchmark should be evenness or randomness (Mazza and Punzo, 2015). However, these criticisms are of a different nature, there is no consensus on their convenience, and they not only affect this index.

is no gender segregation, without altering the occupational structure of the economy. This index has become quite popular because, together with the dissimilarity index, has been used to monitor occupational and industrial segregation by gender in the European Union (Bettio and Veraschchagina, 2009).

To measure overall segregation in the case of two groups, the literature also uses the Gini index (Jahn et al., 1947; Duncan and Duncan, 1955), which is equal to twice the area between the segregation curve and the 45º line, and the square root index (Hutchens, 2001, 2004), which is adapted from the generalized entropy family of inequality indices. These two indices are consistent with the dominance criterion provided by the segregation curves, as shown by Hutchens (1991, 2001) following what Foster (1985) did with Lorenz curves. *Scale invariance*, *symmetry*, the *transfers principle*, and *organizational equivalence* is what render these indices consistent with the segregation curves.[11]

3.2 Multigroup Measures

Some of the binary segregation measures mentioned above have been extended to measure segregation in a multigroup context and, therefore, can be easily obtained as particular cases of more general expressions. This is the case of the *generalized* I_p index (which we call GI_p here) proposed by Silber (1992):

$$GI_p = \frac{1}{2}\sum_g \frac{C^g}{T}\left(\sum_j \left|\frac{c_j^g}{C^g} - \frac{t_j}{T}\right|\right), \qquad (3)$$

where g denotes any of the demographic groups into which the population has been partitioned, c_j^g is the number of group g individuals in unit j, and C^g is the total size of group g.

Other (unstandardized) multigroup measures proposed in the literature are the Gini index proposed by Alonso-Villar and Del Río (2010a):[12]

$$G_u = \frac{1}{2}\sum_g \sum_{i,j} \frac{t_i}{T}\frac{t_j}{T}\left|\frac{c_i^g}{t_i} - \frac{c_j^g}{t_j}\right|, \qquad (4)$$

and the mutual information index proposed by Theil and Finizza (1971):

$$M = \sum_g \frac{C^g}{T}\left(\sum_j \frac{c_j^g}{C^g}\ln\frac{c_j^g / C^g}{t_j / T}\right). \qquad (5)$$

[11] Organizational equivalence, which Hutchens names "insensitivity to proportional divisions of units," means that segregation does not change when a unit is split in several units of equal size in which the representation of the groups does not change.

[12] Boisso et al. (1994) developed a different generalization of the binary Gini index.

Another (unstandardized) multigroup segregation measure satisfying basic properties, which as the M index is related to the generalized entropy family of inequality indices, is the C_u index based on the squared coefficient of variation (Alonso-Villar and Del Río, 2010a):

$$C_u = \sum_g \frac{C^g}{T}\left[\sum_j \frac{t_j}{T}\left(\frac{c_j^g / C^g}{t_j / T} - 1\right)^2\right]. \tag{6}$$

Multigroup segregation measures GIp, G_u, M, and C_u are unstandardized indices. In other words, their maximum value is not 1 when there is complete segregation (i.e., when each group works in occupations with no members of other groups). Although there is no consensus in the literature about whether standardized or unstandardized indices should be used, given that normalization usually comes at a cost in terms of decomposability properties (Mora and Ruiz-Castillo, 2011; Del Río and Alonso-Villar, 2022), standardized versions of the above indices also exist.[13]

The standardized version of M is the popular H index (Theil and Finizza, 1971), an index advocated by Reardon and Firebaugh (2002) in terms of terms of desirable properties, a matter that will be discussed below. The standardized versions of G_u and C_u were also proposed by Reardon and Firebaugh (2002), who called them the generalized Gini index (G) and the square coefficient of variation (C), respectively. The standardized version of the GIp index is the generalized dissimilarity index, proposed by Morgan (1971) and Sakoda (1981) and assessed by Reardon and Firebaugh (2002) in terms of properties. In Table 1, we use the names given by Reardon and Firebaugh (2002) to these indices, although in some cases they are the same as the binary versions.

In addition to the above aspatial measures, the literature also provides measures that account for distances among units, which seems especially convenient when dealing with some types of segregation, as is the case of residential segregation (Reardon and O'Sullivan, 2004).[14] There are also ordinal segregation measures that rank units according to a certain criterion, which allows the incorporation of another dimension in the analysis, given that not all units are equally good (Reardon, 2009).

[13] To obtain the standardized or normalized versions, the unstandardized indices given above must be divided by their maximum values, which in the case of GIp and G_u is $\sum_g \frac{C^g}{T}\left(1-\frac{C^g}{T}\right)$, for M is $\sum_g \frac{C^g}{T}\ln\left(\frac{T}{C^g}\right)$, and for C_u is the total number of groups in the economy minus 1.

[14] For a review of spatial measures, see Yao et al. (2018).

4. Local Segregation

Along with the extent of overall segregation (e.g., occupational segregation by gender and race/immigration), one may be interested in finding out the degree of unevenness of each target group (e.g., the occupational segregation of immigrant women). To deal with this, the literature has mainly undertaken pairwise comparisons between the target group and any other group using binary segregation measures (Reskin and Cassirer, 1996; Mintz and Krymkowski, 2011; Iceland et al., 2014), a process that becomes cumbersome when many groups are involved. Other scholars opt to compare the distribution of the target group across units with that of its complementary (Queneau, 2009; Marcinczak et al., 2016).

Alonso-Villar and Del Río (2010a) follow a different approach and develop a formal framework within which this measurement can be formally addressed. They establish the basic properties that these measures, called local segregation measures, must satisfy, and propose several measures meeting them. In what follows, we provide the (unstandardized) local measures associated with the above (unstandardized) overall measures:

$$D^g = \frac{1}{2}\sum_j \left| \frac{c_j^g}{C^g} - \frac{t_j}{T} \right|, \tag{7}$$

$$G^g = \frac{\sum_{i,j} \frac{t_i}{T}\frac{t_j}{T}\left| \frac{c_i^g}{t_i} - \frac{c_j^g}{t_j} \right|}{2\frac{C^g}{T}}, \tag{8}$$

$$\Phi_1^g = \sum_j \frac{c_j^g}{C^g} \ln\left(\frac{c_j^g / C^g}{t_j / T} \right), \text{ and} \tag{9}$$

$$\Phi_2^g = \frac{1}{2}\sum_j \frac{t_j}{T}\left[\left(\frac{c_j^g / C^g}{t_j / T} \right)^2 - 1 \right]. \tag{10}$$

D^g is a local index linked to overall index GI_p, given that the latter can be written as the weighted average of the segregation of each group using the former: $GI_p = \sum_g \frac{C^g}{T} D^g$. Likewise, the local index G^g is associated with the overall index G_u, $G_u = \sum_g \frac{C^g}{T} G^g$. The local index Φ_1^g is the one associated with the mutual information index, $M = \sum_g \frac{C^g}{T} \Phi_1^g$, and the local index Φ_2^g is the one corresponding to the overall index C_u, $C_u = \sum_g \frac{C^g}{T} \Phi_2^g$.

The D^g index was proposed by Moir and Shelby Smith (1979) in an empirical paper to analyze gender segregation in the Australian labor market. However, these authors did not explore the properties of this index to measure the segregation of a group in a multigroup context. Their goal was to explore gender segregation using a modified version of D that compared the distribution of women across units with the distribution of total workers across the same units, rather than the distribution of men (as D does). This index is known as Gorard index in the literature on school segregation.

D^g can be interpreted as the percentage of group g individuals who would have to move among units to reach zero segregation (i.e., to be distributed across units in the same way as the total population is). It also has a graphical interpretation: it measures the highest vertical distance between the local segregation curve proposed by Alonso-Villar and Del Río (2010a) and the 45º line. To construct this curve, the units must be ranked ascendingly by $\frac{c_j^g}{t_j}$ and then we look at the first unit, the first two units, the first three units, and so on. Next, we plot the cumulative proportion of the target group in the corresponding units against the cumulative proportion of total population in the same units (Figure 2).

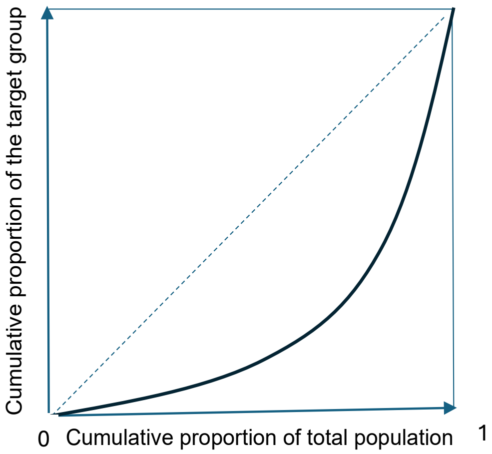


Figure 2. Local segregation curve

As happens with the original dissimilarity index, D^g does not meet the *transfers principle* (adapted to measure the segregation of a group), although it satisfies other basic properties (*scale invariance*, *symmetry*, and *insensitivity to proportional subdivisions of units*) (Alonso-Villar and Del Río, 2010a). Unlike D^g, the local indices G^g, Φ_1^g, and Φ_2^g do satisfy the local version of the *transfers principle* (either *disequalizing movements type I, II,* or *III*), along with the other properties mentioned above. These three indices are consistent with the dominance criterion given by the local segregation curves (Alonso-Villar and Del Río, 2010a). In other words, when a local segregation curve is above another (or equals it at some points), these indices rank these two distributions in the same way as the curves do: the segregation of the group is lower for the distribution whose curve is closer to the 45º line.

Standardized versions of the above local measures, whose values range between zero and one, can also be built. They are linked to standardized overall measures in a way analogous to the one mentioned for the unstandardized measures, although with different weights (Del Río and Alonso-Villar, 2022). To facilitate the reading of this paper, we do not bring here the corresponding expressions. However, it is worth mentioning that when measuring the segregation of a group using the above unstandardized local measures, we are measuring how far the local segregation curve is from the egalitarian distribution (given by the distribution of the total population among units). When using instead the standardized versions, we are quantifying how far the local curve is from the curve of maximum segregation (which depends on the group's size). Therefore, the use of standardized and unstandardized measures permits to look at the segregation of a group from a different angle. Given the link between local and overall measures, these ideas are also behind overall measures.

The distinction between the three axes we have mentioned so far—binary versus multigroup measures, local versus overall measures, and standardized versus unstandardized measures—allows us to classify many of the segregation measures that the literature provides and connect them (see Table 1, where arrows link local and overall indices, solid lines link binary and multigroup indices, and dashed lines connect standardized and unstandardized). This classification, adapted from Del Río and Alonso-Villar (2022), allows us to shed some light on academic debates about the convenience of using one index or another to measure segregation. As these authors discuss, in the case of school segregation between two groups, there has been a debate about whether to use index D^g (known as Gorard index) or the dissimilarity index D (Allen and Vignoles, 2007; Gorard and Taylor, 2002), without realizing that the former is an unstandardized local segregation measure while the latter is a standardized overall measure. In a binary context, the standardized version of D^g equals the dissimilarity index, which means that, ultimately, the matter is whether to use a standardized measure or not.

Table 1. Segregation in a three-axis setting: Multigroup versus binary, overall versus local, and standardized versus unstandardized indices

	UNSTANDARDIZED INDICES					STANDARDIZED INDICES				
BINARY			I_p	I_p index	Karmel & MacLachlan (1988)	D	Dissimilarity index	Jahn et al. (1947); Duncan & Duncan (1955)		
						G	Gini index	Jahn et al. (1947); Duncan & Duncan (1955); Silber (1989)		
						V / Eta^2	Variance ratio index/ the square of the correlation ratio index	Duncan & Duncan (1955)		
						O	Square root index	Hutchens (2001)		
MULTIGROUP	**Local indices** [Alonso-Villar & Del Río, 2010a]		**Overall indices**			**Overall indices**			**Local indices** [Del Río & Alonso-Villar, 2022]	
	D^g	Local dissimilarity index (Moir & Selby Smith, 1979)	GI_p	Generalized I_p index	Silber (1992)	D	Generalized dissimilarity index	Morgan (1975); Sakoda (1981); Reardon & Firebaugh (2002)	$\tilde{D}^g$	Standardized local dissimilarity index
			G_s	Multidimensional G-Segregation index	Boisso et al. (1994)					
	G^g	Local Gini index	G_u	Overall Gini index	Alonso-Villar & Del Río (2010a)	G	Generalized Gini index	Reardon & Firebaugh (2002)	$\tilde{G}^g$	Standardized local Gini index
	Φ_1^g	Local entropy index	M	Mutual information index	Theil & Finizza (1971); Frankel & Volij (2011)	H	Information theory index	Theil & Finizza (1971)	$\tilde{\Phi}_1^g$	Standardized local entropy index
	Φ_2^g	Local index based on the squared coefficient of variation	C_u	Overall index based on the squared coefficient of variation	Alonso-Villar & Del Río (2010a)	C	Squared coefficient of variation	Reardon & Firebaugh (2002)	$\tilde{\Phi}_2^g$	Standardized local index based on the squared coefficient of variation

Source: Adapted from Del Río and Alonso-Villar (2022).

4.1 Connecting the Segregation of a Group with Overall Segregation

We have mentioned that measuring the segregation of a group with local measures is consistent with the way overall segregation is quantified, given that the latter can be expressed as the weighted average of the segregation of the groups involved. These weights equal their demographic shares in the case of using unstandardized measures:

$$I = \sum_{g} \frac{C^g}{T} I^g , \tag{11}$$

where I stands for overall segregation and I^g is the segregation of group g using the corresponding local segregation index. What we have not discussed yet is what this decomposition implies. This decomposition tells us that the use of local measures can be especially helpful when working with small groups (e.g., racial minorities, sexual minorities, or recent immigrants) because in these cases their impact on overall segregation is small (given their low weights) and, therefore, what happens to them may be hidden when using overall measures (Del Río and Alonso-Villar, 2019a,b).

The use of this local segregation framework has also permitted us to delve deeper into the properties that overall measures should satisfy. First defining the properties of the local segregation measures and building from them the properties that overall segregation must meet seems a natural and simple way to approach this measurement. (This is not how scholarship has done it, since overall segregation was thought of before addressing the segregation of a group).[15] Reardon and Firebaugh (2002) show that, among the standardized indices they assess (*GIp*, *G*, *H*, and *C*), *H* is the only one that satisfies the *principle of transfers* (i.e., the index decreases whenever an individual in a group moves from one unit to another in which that group has a lower representation), although they wonder whether this invalidates the other indices. Del Río and Alonso-Villar (2022) address this issue and suggest that, in light of the local segregation approach, violation of *the principle of transfers* by an overall measure should not be a problem as long as the corresponding standardized local segregation measures comply with *sensitivity to disequalizing movements type III*. This means that *G* and *C* are sensible segregation measures as well.

[15] Unlike the segregation literature, in other economics fields the measurement of the "parts" was addressed before the measurement of the "total." This is the case of regional economics, in which the geographical concentration of a sector was tackled, often using inequality-based indices, before dealing with the geographical concentration of the whole economic activity (Brülhart and Traeger, 2005; Aiginger and Davis, 2004; Mulligan and Schmidt, 2005; Alonso-Villar and Del Río, 2013a).

When an overall index does not satisfy the *principle of transfers*, the reduction in overall segregation arising from an individual moving from a unit to another in which the group has a lower presence may not offset the possible rise in segregation associated with the impact that the change in the size of the incumbent units may have on the other groups. However, if the local index satisfies *sensitivity to disequalizing movements type III*, we know that it decreases whenever a member of the group moves to another unit in which the group has a lower representation (no matter how this affects the other groups). In other words, when using local measures, we look at the effect of *disequalizing* movements only for the target group. Requiring that equalizing movements in a group always reduce overall segregation may seem unnecessary. In this sense, *the principle of exchanges* (according to which overall segregation decreases when two individuals of two different groups exchange their positions moving from a unit in which the incumbent group has a higher representation to another in which it has a lower representation), which is another translation of the Pigou-Dalton principle to this context, may appear as a more reasonable requirement for the measurement of overall segregation.

4.2 Measuring the Consequences of Segregation in Terms of Wellbeing

The literature on segregation has paid little attention to the connection between segregation and welfare from a normative point of view. A few proposals deal with the measurement of unevenness while accounting for the status or "quality" of units (Reardon, 2009; Del Río and Alonso-Villar, 2012a), but they measure segregation, not the welfare associated with that situation.

By looking at each group separately, the local segregation framework opens new avenues of research, allowing for questions that have not been answered before. Thus, we may consider whether an uneven distribution of a group across units gives it advantages or disadvantages. This inquiry would be pointless in the field of income inequality because in that context inequality is a bad thing per se. However, when addressing segregation, let's say occupational segregation, unevenness benefits or harms a group as long as it is concentrated in well-paid or low-paid jobs, respectively. In what follows, we expose how this topic can be addressed focusing on occupational segregation, although it could be adapted to other types of segregation as long as the quality or status of units can be measured cardinally.

We denote by $t \equiv (t_1, t_2, ..., t_J)$ the distribution of total employment across J occupations and by $c^g \equiv (c_1^g, c_2^g, ..., c_J^g)$ the distribution of group g across these same units. To assess the occupational sorting of group g we need to incorporate another element in the analysis. We have to be able to distinguish between good and bad occupations, and the relative wage of each occupation seems a sensible way to do this. Let's $w \equiv (w_1,...,w_J)$ be the occupational wage distribution (i.e., the average wage of each occupation) and $\frac{w_j}{\bar{w}}$ be the relative wage of occupation j, where $\bar{w} = \sum_j \frac{t_j}{T} w_j$ is the average wage of the economy. Following Alonso-Villar and Del Río (2017a) and using a social welfare function *W*(.) to assess the state in which a society is, we can measure the well-being loss or gain that group *g* has associated with its occupational sorting, $\Psi(c^g;t;w)$, as the loss or gain *g* has from departing from evenness (in per capita terms). Namely,

$$\Psi(c^g;t;w) = \frac{1}{C^g}[W(c^g;t;w) - W(\frac{C^g}{T}t;t;w)], \tag{12}$$

where $\frac{C^g}{T}t \equiv \left(\frac{C^g}{T}t_1,...,\frac{C^g}{T}t_J\right)$ is the egalitarian distribution in this case, so that in each occupation the group accounts for $\frac{C^g}{T}$ percent of the employment there.

Alonso-Villar and Del Río (2017a) established the properties that an index aimed at assessing the sorting of a group across units in terms of wellbeing, rather than the extent of segregation, should meet and proposed a family of indices based on them:

$$\Psi_\varepsilon(c^g;t;w) = \begin{cases} \sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T}\right) \frac{\left(\frac{w_j}{\bar{w}}\right)^{1-\varepsilon} - 1}{1-\varepsilon} & \varepsilon \neq 1 \\ \sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T}\right) \ln \frac{w_j}{\bar{w}} & \varepsilon = 1 \end{cases} \tag{13}$$

where $\varepsilon > 0$ is a parameter of aversion to inequality within the group (i.e., to the fact that some members of the group are in highly paid occupations while others are in low-paid occupations). As these authors point out, underrepresentation (respectively, overrepresentation) in an occupation only penalizes the index when this occurs in highly (respectively, poorly) paid jobs.

In the limit case where $\varepsilon = 0$, $\Psi_0(c^g;t;w) = \sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T} \right) \frac{w_j}{\overline{w}}$, which is the Γ^g index defined by Del Río and Alonso-Villar (2015) to quantify the monetary loss or gain that group *g* has associated with its occupational sorting (in per capita terms), expressed as a proportion of the average wage of the economy ($\overline{w}$). In other words, Γ^g measures the loss or gain in wellbeing of the group without taking into account the inequality that exists within it due to some of its members working in well-paid occupations and others in poorly paid occupations.

Note that the wellbeing of each group into which the whole economy has been partitioned can be aggregated to determine the welfare loss that society experiences due to segregation (Del Río and Alonso-Villar, 2018). This has been done adding the losses of disadvantaged groups in a similar way to what is done in the literature on economic deprivation, since these losses can be thought of as deprivation gaps, and defining curves of social welfare losses associated with segregation (WLAS curves) and indices consistent with them.[16]

4.3 The Total Advantage or Disadvantage of a Group

Along with the advantage or disadvantage of a group due to its occupational sorting, whether in monetary terms or in terms of wellbeing, the group may receive wages above or below those of other groups working in the same occupations. In what follows, we explain how the total wage advantage or disadvantage of a group can be decomposed into these two components. Here we provide the decomposition when there is no inequality aversion, although similar expressions exist for $\varepsilon > 0$ (Alonso-Villar and Del Río, 2017a).

As shown by Del Río and Alonso-Villar (2015), the earning gap of a group, $EGap^g = \left(w^g - \overline{w} \right) \frac{1}{\overline{w}}$ (i.e., the difference between the average wage of the group and the average wage of the economy divided by the latter), can be decomposed as follows:

$$EGap^g = \underbrace{\left[\sum_j c_j^g \left(w_j^g - w_j \right) \right] \frac{1}{C^g \overline{w}}}_{\Delta^g} + \underbrace{\sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T} \right) \frac{w_j}{\overline{w}}}_{\Gamma^g}, \qquad (14)$$

where w_j^g denotes the (average) wage of group *g* within occupation *j*, Γ^g is the group's gain/loss associated with its occupational sorting, and Δ^g is the gain/loss arising from whether the group's

[16] These curves are similar to the TIP curves (Three Is of Poverty) used in the literature of income distribution.

wage within each occupation is higher or lower than that of other groups. This decomposition allows us to determine whether the fact that a group has a wage above or below average arises mainly from the occupations in which it tends to work, from earning more/less than other groups who work in the same occupations, or from both sources. We will use this decomposition in our empirical analysis, which for simplicity is called the within-between decomposition.

Each component can be further decomposed to determine the contribution of each occupation to the $EGap^g$ (Alonso-Villar and Del Río, 2023a). Thus,

$$\Gamma^g = \sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T} \right) \frac{w_j}{\bar{w}} = \sum_j \underbrace{\left(\frac{c_j^g}{C^g} - \frac{t_j}{T} \right) \frac{\left(w_j - \bar{w} \right)}{\bar{w}}}_{\Gamma_j^g} \text{ and} \tag{15}$$

$$\Delta^g = \sum_j c_j^g \left(w_j^g - w_j \right) \frac{1}{C^g \bar{w}} = \sum_j \underbrace{\left(\frac{c_j^g}{C^g} \right) \left(\frac{w_j^g - w_j}{\bar{w}} \right)}_{\Delta_j^g}. \tag{16}$$

Expression (15) allows us to single out the occupations that bring advantages to the group. Occupations that contribute positively to Γ^g, $\Gamma_j^g > 0$, are those well-paid occupations in which the group is overrepresented and those low-paid occupations in which the group is underrepresented. On the contrary, $\Gamma_j^g < 0$ when j is either a low-paid occupation in which the group is overrepresented or a highly paid occupation in which the group is underrepresented. Analogously, expression (16) allows us to identify the occupations in which the group receives wages below the occupational wage ($w_j^g - w_j < 0$) or above it. The magnitude of this effect also depends on the importance of that occupation for the group ($\frac{c_j^g}{C^g}$).

5. Explaining Segregation

So far, we have discussed how to measure the extent of segregation, or its consequences in terms of wellbeing. In this section, we examine the methods the literature has used to explain segregation. When comparing two countries, regions, or cities, or when comparing a country over time, we may be interested in determining whether segregation is greater in an economy than the other because some groups have more difficulties integrating there (i.e., "pure" segregation is larger) or because the size of the organizational units or groups differs between these two economies. To address this, some scholars opt to use margin-free segregation measures, which do not change when the units or groups change their relative sizes (Charles and Grusky, 1995). In other words, with these measures, the differences in segregation between two economies do not arise from differences in their marginal distributions (e.g., the distributions of genders and occupations).

However, there is no consensus in the literature about the convenience of using margin-free segregation measures, in part because we may want the index to increase when more individuals are in segregated units, which involves margin dependency, and also because margin independency may conflict with other properties (Elbers, 2023). This is why some scholars advocate the use of margin-dependent measures and propose decomposing the change in segregation between two economies into marginal changes (i.e., changes in the distributions of groups or in the distributions of units) and structural changes (i.e., changes in "pure" segregation). Thus, drawing on Theil (1972) and on the decomposition procedure provided by Karmel and MacLachlan (1988) to the I_p index, which involves gradual adjustments of the marginal distributions using iterative proportional fitting, Elbers (2023) propose a similar method to disentangle the change in segregation in these two components for the M index.[17] Others use simple decompositions of the difference in segregation between two scenarios using an intermediate step in which the index is applied to a counterfactual economy built by changing either the unit marginals or the group marginals (Blau et al., 2013; Alonso-Villar and Del Río, 2017b; Azpitarte et al., 2021). The problem with this method is that the results may vary depending on whether the unit or instead the group marginals are the first to change, a problem that can be addressed by averaging the values obtained by the two paths (using the Shapley decomposition; Chantreuil and Trannoy, 2013; Shorrocks, 2013).

Scholarship has also explored whether differences in characteristics may explain differences in segregation across locations, with the country, region, city, or establishment being the unit of analysis. To do this, these studies make use of geographical variability in the characteristics of groups and locations to explain segregation, or its consequences, using regression models (Abrahamson and Sigelman, 1987; Massey and Denton, 1987; Lorence, 1992; Huffman et al.,

[17] The marginal change is further decomposed in changes associated with the units' marginals and those involving the groups' marginals, as Deutsch et al. (2009) do for the generalized I_p index (GI_p), whereas the structural change can be decomposed to determine the contribution of each unit.

2010; Alonso-Villar et al., 2012; Alonso-Villar and Del Río, 2017c; Borrowman and Klasen, 2020).

Another way to address how differences in group characteristics can explain intergroup differences or changes in segregation, across locations or over time, is to construct a counterfactual distribution that removes the composition effect (Alonso-Villar et al., 2013; Gradín, 2013, 2019; Gradín et al., 2015; Palencia-Esteban, 2022). This involves controlling for groups' attributes that could affect their positions in the labor market. For example, in the case of occupational segregation by gender and immigration, these attributes could include educational achievements, age, region of residence, etc. We discuss this approach in more detail below because it is the one we follow in our empirical analysis.

To construct this counterfactual, we must partition each group into several subgroups or cells, defined by the combination of the characteristics for which we want to control, and choose the group (e.g., native men) whose characteristics are taken as a reference for the other groups. Next, we keep the distribution of each subgroup (e.g., immigrant women with tertiary education, in a certain age range, and living in a certain region) across units (e.g., occupations) as we observe in the actual distribution, but change the weight of that subgroup to make it equal to the weight that the corresponding subgroup has in the reference group. To do this, we can follow a simple nonparametric method, which only requires reweighting the subgroups according to their weights in the reference group (Alonso-Villar and Del Río, 2023b), or we can instead use a parametric method whose reweighting scheme involves logit estimates (based on the popular propensity score procedure proposed by DiNardo et al.,1996, as adapted by Gradín, 2013).[18] The advantage of the parametric method is that it allows for an easy decomposition, by factors, of the change between unconditional values (e.g., segregation in the actual economy) and conditional values (e.g., segregation in the counterfactual). To avoid the problem of path dependency that the original method has, which makes the decomposition depend on the order in which the different covariates are included in the analysis, Gradín (2013) proposes to use the Shappley value. The same technique can be used to determine conditional wages and the role that occupations play in explaining intergroup wage disparities after controlling for characteristics.

[18] When working with small groups, it is easier to replicate the characteristics of the reference group using the nonparametric method, as illustrated in the case of conditional poverty by sexual orientation (Alonso-Villar and Del Río, 2023c).

6. An Illustration: Occupational Segregation and Wages in Spain

Despite social advances, gender inequalities persist in the labor markets of Western countries, including Spain (Blau and Kahn, 2017; Anghel et al., 2019; Adserà and Querin, 2023). However, women are not a homogeneous group, nor are men. Individuals' labor opportunities are also shaped by their migration status, race, or ethnicity (Algan et al., 2010; De la Rica et al., 2014; Del Río and Alonso-Villar, 2015; O'Higgins, 2015; Cantalini et al., 2023; Alonso-Villar and Del Río, 2023b; Palencia and Del Río, 2024).

In this section, we use several tools presented above to explore occupational segregation by gender and nativity, along with its effects on wages, in Spain, a country with significant migration flows since the end-1990s. We use an intersectionality framework that distinguishes among four groups: immigrant women, native women, immigrant men, and native men.[19] Our analysis draws on the 2006-2024 EPAs (second quarter) and the 2022 EES (the quadrennial survey). We define immigrants as those who do not have Spanish citizenship. It is important to note that the EPA provides a smaller sample of employed individuals than the EES but has better coverage of occupations related to domestic service and agriculture, which is relevant when analyzing segregation by gender and nativity.[20]

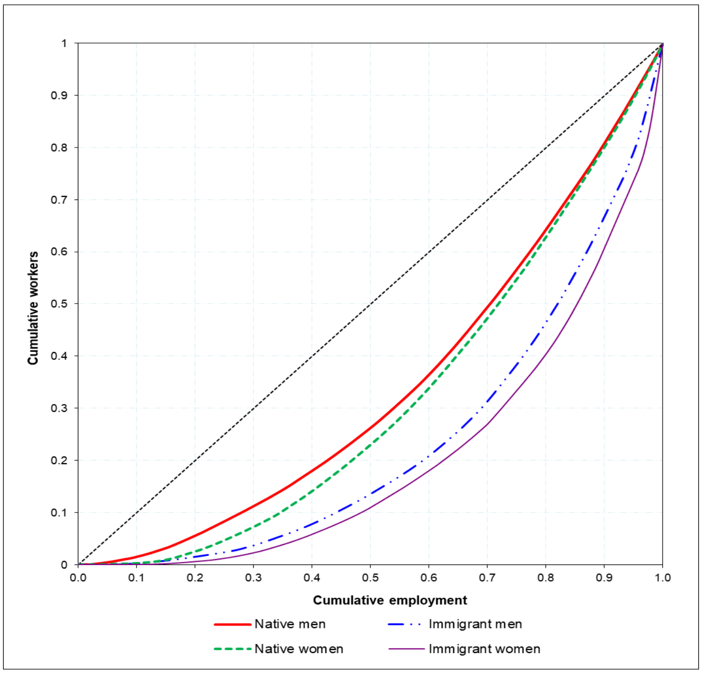
Our first goal is to document the extent of segregation for the period 2006-2024, thus enlarging the period covered in previous studies (Alonso-Villar and Del Río, 2017d, 2024). Drawing on the 2024 EPA and the 2022 EES, our second goal is to explore the role that occupations play in explaining each group's position on the wage ladder before and after controlling for group characteristics. To perform the conditional analysis, we build a counterfactual economy using both parametric (DiNardo et al., 1996; Gradín, 2013) and nonparametric methods (Alonso-Villar and Del Río, 2023b) mentioned earlier. To build these counterfactuals, we take native men as the reference group and adjust the characteristics of the other three groups to make them equal to those of native men. Unlike Palencia and Del Río (2024), we do not explore why the occupational sorting of immigrants differs between countries, but rather why the occupational sorting of immigrants (and their average wages) differs from those of natives, which implies that the reference group is Spanish men and not immigrants from a country of reference (in their case, immigrant women and men from the UK).

[19] Intersectionality means that the interaction of the various identities that individuals possess provides them with unique experiences, as some identities can bring advantages and others disadvantages.
[20] Alonso-Villar and Del Río (2017d) provide more details on these two data sets, including the definition of immigrant.

By exploring the conditional wages of the four groups simultaneously, our analysis departs from what is usually done in the literature, which tends to analyze gender and nativity wage gaps separately or explore the nativity wage gap within a gender group (Antón et al., 2010; Nicodemo and Ramos, 2012; García-Pérez et al., 2012; Anghel et al., 2019; Pinto et al., 2023). We also depart from most studies on wage gaps by exploring the role that occupations play without including them as covariates, allowing us to use a broader list of occupations. A limitation of our analysis is that, unlike methods based on wage equations, our approach requires including only basic covariates to have enough observations of immigrant women and men with a given combination of characteristics.

The Extent of Segregation

Figure 3 displays the local segregation curves for each group in 2024 using the national classification of occupations at the three-digit level (169 occupations).[21]



Source: Authors' calculations.

Figure 3. Local segregation curves for gender-nativity groups, 2024 EPA

Since the curve for immigrant women is below those of the other groups, occupational segregation is greater for them using a broad set of segregation indicators (like G^g, Φ_1^g, and Φ_2^g, inter alia). We also see that the curve for immigrant women is closer to zero for a wider range of abscissa values. Immigrant women have no presence in occupations that account for

[21] The graph at the two-digit level, with 62 occupations, is similar. One could also combine occupations and sectors to take into account that women and men do not distribute equally among sectors (Alonso-Villar and Del Río, 2010b).

about 10% of employment (and is extremely small up to 20%), while this proportion is 5% in the case of immigrant men.

Hereafter, we use the national classification of occupations at the two-digit level given that, to later assess the occupational distributions of the groups, we need information on wages and the EES does not provide it at the three-digit level.[22] Figure 4 displays the evolution of the D^g and Φ_1^g indices for each group for 2006-2024, while Figure 5 illustrates the evolution of segregation by gender, by nativity, and by gender-nativity using the corresponding overall segregation indices (I_p and M, respectively).[23][24]

Figure 4 shows that immigrant women experienced an important drop in segregation throughout the period, although the process stopped in 2021. In 2024, around 40% of immigrant women would have to change occupations to be evenly distributed across them (D^g=0.4), while in 2006, it was 56%. This drop arises from their lower presence in occupations in which they were highly overrepresented (such as catering service workers and domestic employees) and a greater presence in many occupations in which they were underrepresented (including different types of technicians and professionals, support professionals, and office employees who do not deal with the public). This change in their occupational sorting could be associated with changes in the composition of the group. Over the period, immigrant women experienced a sharp increase in education, an increase in age, and changes in country of origin (Table A1 in the Appendix).

Segregation also decreased for native women and immigrant men over the period, although with much lower intensity and only lasting a few years (until around 2012). In 2024, 26% of native women would have to change their occupation to be evenly distributed among them, while in 2006 it was 32%. The percentages for immigrant men are 37% and 40%, respectively. However, the degree of unevenness for native men has remained quite stable throughout the period (around 23% of them would have to change occupation). Females and males tend to converge over the period in terms of segregation, especially among the immigrant population.

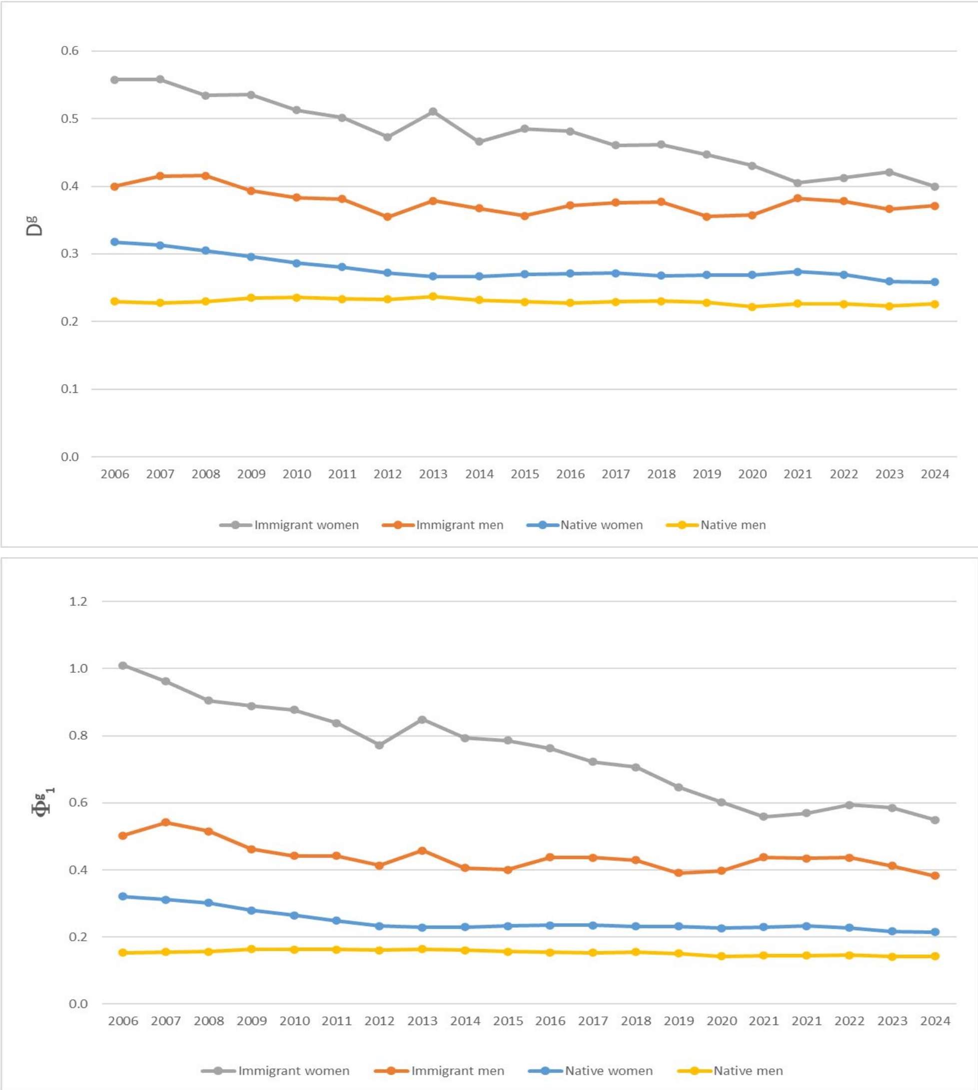
Alonso-Villar and Del Río (2013b) showed that the increase in the immigrant population that took place in Spain between 1996 and 2006 was accompanied by an increase in the occupational segregation of immigrants (women and men considered together), especially in the period 1998-

[22] The EPA does not provide information on individual wages.

[23] Note that there is a break in the series in 2011 since the national classification of occupations is CNO-94 until 2010 (66 occupations) and CNO-11 (62 occupations) since 2011. It is worth keeping in mind that segregation estimates are sensitive to the classification used.

[24] The values corresponding to Figures 4-7 are given in the Appendix (Tables A2-A4).

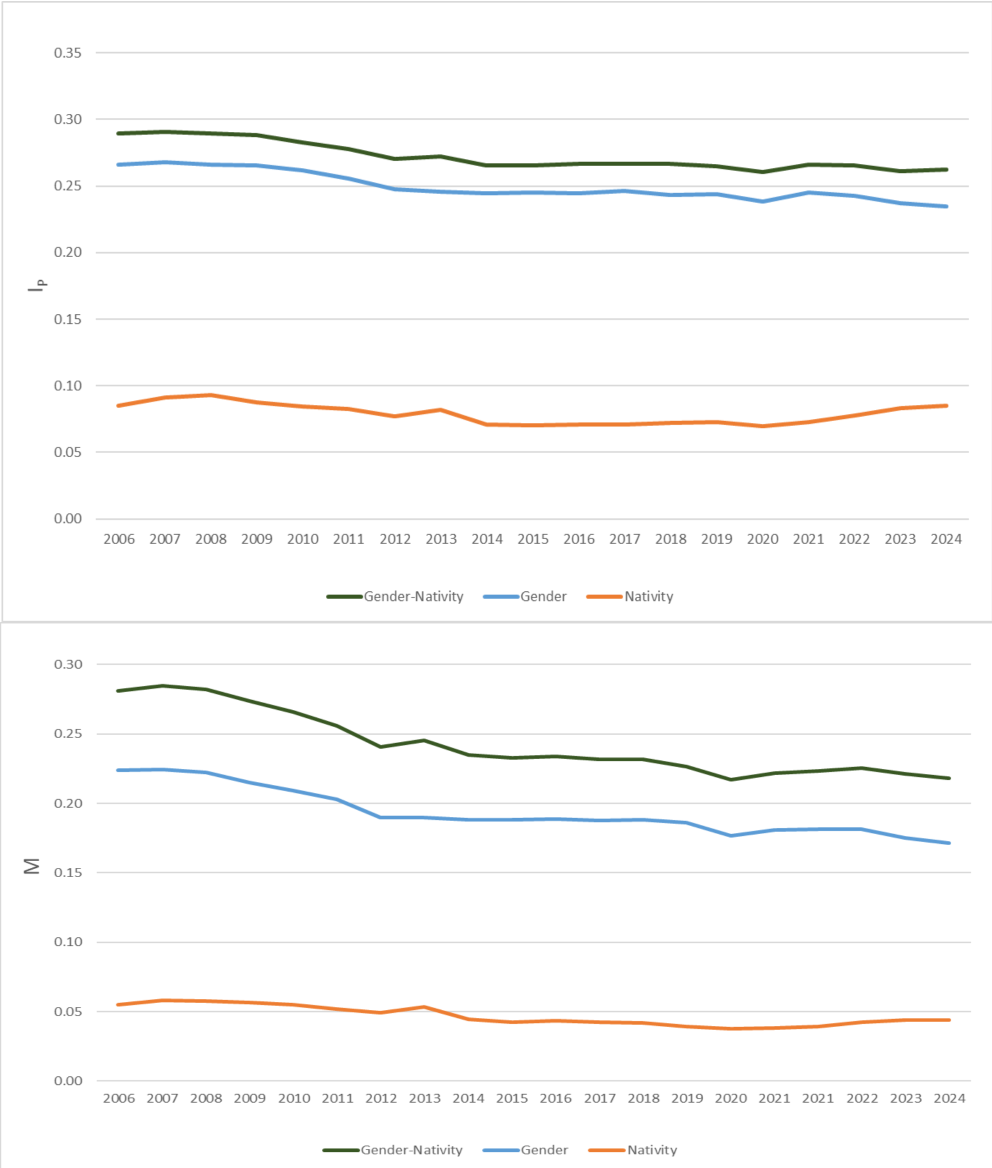
2004. This increase was clearly linked to the model of economic growth that prevailed in Spain during that decade. Figure 4 suggests that this trend changed in subsequent years since there is a certain convergence among the four groups, mainly driven by immigrant women catching up with immigrant men.



Source: Alonso-Villar and Del Río (2024) and authors' calculations.

Figure 4. Evolution of local segregation (D^g and Φ_1^g), 2006-2024 EPAs.

Figure 5 reveals that most of the gender-nativity segregation that we observe in the Spanish labor market is driven by gender; segregation by nativity is of lesser magnitude. Moreover, the evolution of the gender-nativity segregation seems to parallel that of gender segregation, both of which experienced a fall in the first years of the great recession (2008-2012). Gender segregation also appears to have experienced a small decline after the pandemic (2022-2024), a pattern that is not reflected in terms of gender-nativity segregation, perhaps due to the small increase in nativity segregation in recent years.



Source: Alonso-Villar and Del Río (2024) and authors' calculations.

Figure 5. Evolution of overall segregation (I_p and M) by gender, nativity, and gender-nativity, 2006-2024 EPAs.

Note that even though segregation by nativity does not change much over the period, the segregation of immigrant women did experience an important drop, as already mentioned. The fact that this decline is not fully reflected in overall segregation may be because a group's contribution to overall segregation depends on its size, as discussed above. The use of local segregation measures is especially helpful when working with small groups.

The Consequences of Segregation for the Incumbent Groups

As explained in Section 4.3, a group's occupational sorting may bring it a monetary advantage or disadvantage, depending on whether the group tends to be concentrated in high- or low-paying occupations. The group may also have a wage advantage or disadvantage within occupations compared to other groups. Figure 6 provides the hourly wage advantage or disadvantage (*EGap*) of each group with respect to the national average wage, as well its two components, Γ and Δ (see expression (14)). For simplicity, the superscript that refers to the group is removed. Employment data comes from the 2024 EPA, while wages for each group in each occupation come from the 2022 EES.[25] Results are shown before and after controlling for characteristics. The values obtained in the counterfactual economy that we construct using the parametric method are denoted by the subscript *p*, and those based on the nonparametric are denoted by *np*.

Source: Author's calculations using employment from the 2024 EPA and wages from the 2022 EES.
Note: *p* (parametric) and *np* (nonparametric) counterfactuals.

Figure 6. Within-between decomposition of the hourly wage advantage or disadvantage of each group (relative to the national average) in the actual economy and in the counterfactuals, 2024 EPA.

The chart reveals that the unadjusted hourly wages of immigrant women and men are well below the national average, especially those of the former (25% versus 16%), with occupational sorting explaining most of their wage disadvantage.[26] Figures A1-A4 in the Appendix display the contribution of each occupation to Γ and Δ using the decompositions provided in

---

[25] Immigrant wages may be overestimated. The EES only includes workers in enterprises in industry, construction, or services who have been affiliated with social security during the month of October of the corresponding year.

[26] Drawing on the 2019 European Labor Force Survey and the 2014 Structure of Earnings Survey, Palencia and Del Río (2024) document that in Germany, Finland, Italy, and Slovenia, the occupational sorting of immigrant women penalizes them significantly as well, while in Portugal the penalty is small, and France, the Netherlands, Norway, and Sweden show intermediate levels. The concentration of immigrant men in low-paying occupations is much smaller. In fact, in the UK, the Czech Republic, and especially Portugal, immigrant men are not concentrated in low-paying jobs. Unlike Palencia and Del Río (2024), our data sets allow us to determine not only the value of Γ for each group but also the value of Δ, and therefore the *EGap*.

expressions (15) and (16). Figure A1 shows that there are occupations that penalize immigrant women due to the group's overrepresentation in low-paying occupations (e.g., catering, cleaning work, personal care, and especially domestic work) or their underrepresentation in highly paid ones (sciences and engineering professions and health professions). In all of them, immigrant women are also underpaid compared to other groups, especially in the last two. The problem of overrepresentation in bad occupations and underrepresentation in good ones is not as intense for immigrant men, although they are overrepresented in agriculture and fishing, construction, and catering and underrepresented in health professions (Figure A3). While they are also underpaid in these occupations, their underpayment is not as intense as previously shown for immigrant women, except in the case of agriculture and fishing. Native women are also highly concentrated in cleaning, retail sales, or care work (Figure A2), but this group's position on the wage ladder does not seem to arise from a disadvantaged occupational sorting as a whole (Figure 6). Notwithstanding, native women do not have the occupational advantage that native men have. And native women also have wage disadvantages within occupations, which are especially significant in some of them (Figure A2).

As already mentioned, to calculate the conditional *EGap* of each group, together with the conditional Γ and Δ values, we build a counterfactual economy taking native men as the reference group with respect to which the characteristics of the other groups are adjusted. This involves reweighting the cells in each group to replicate those of native men. The counterfactual economy shows how the occupational sorting of the groups, and their wages, would be if the four groups had been equal in terms of basic characteristics. The covariates included in the conditional analysis are: educational attainment, age, and location.[27] To have enough observations of immigrants in each category (cell), we distinguish only among three educational levels: lower secondary education at most, upper secondary education, and tertiary education. We have three age groups: up to 30 years old, between 30 and up to 55, and 55 or older. Since immigrants are not evenly distributed across regions, and taking into account that there are wage disparities among regions, we also distinguish between individuals living in regions with wages above and below the national average. This means that, in the conditional analysis, we partition each group into 18 (3x3x2) subgroups or cells.

[27] There are important differences among immigrants by region of origin, which affects them in terms of occupations and wages (Del Río and Alonso-Villar, 2012b; García et al., 2012). However, that analysis is beyond the scope of this illustration.

Figure 6 shows that the situation improves significantly for immigrant men after controlling for characteristics, their conditional wages are 7-8% below average, mainly because they have a more favorable distribution across occupations. Unlike them, the conditional hourly wage of immigrant women, which is 21% below the average, barely improves compared to their unconditional value (25%). Our analysis shows that occupational sorting strongly penalizes immigrant women even after controlling for characteristics. Unlike what happens with immigrants, when comparing the actual economy and the counterfactual (i.e., before and after removing the composition effect), the wages of native women decrease, since in the counterfactual the occupational sorting of native women penalizes them, although only slightly.

Figure 7 provides the same information as Figure 6 but using only the 2022 EES for both employment and wages. In this case, for each individual, we have not only their education, age, and location, but also their wage, allowing us to build a more accurate counterfactual for the sectors included in this data set.

Source: Author’s calculations using employment and wages from the 2022 EES.
Note: *p* (parametric) and *np* (nonparametric) counterfactuals.

Figure 7. Within-between decomposition of the hourly wage advantage or disadvantage of each group (relative to the national average) in the actual economy and in the counterfactuals, 2022 EES.

When comparing Figures 6 and 7, we observe that intergroup disparities are more evident with the EPA, perhaps because the EES does not include all sectors, which has an important effect on the size of some occupations in which immigrants tend to be concentrated. This is the case of domestic service, for example, which may explain why the wage disadvantage of immigrant women, and the role that occupations play, is smaller with the EES. In any case, the patterns detected earlier for the EPA remain when using the EES.

Finally, the intersectionality approach followed in this research allows us to show that the gender wage gap in Spain is higher among the native population than it is among immigrants (10.3% of the average wage versus 8.6% with the 2024 EPA in the actual economy), and the wage nativity gap is larger for men than it is for women (24.3% of the average wage versus 22.6%).[28]

7. Final Comments

After offering some reflections about how the literature has approached the measurement of segregation, this paper has explored occupational segregation by gender and nativity in Spain in an intersectionality framework. The analysis has revealed that the occupational segregation of immigrant women is a more intense phenomenon than that of native women or immigrant men, although it decreased significantly over the period 2006-2024 (the process came to a halt in 2021). Part of the segregation of immigrant women and men arises from a composition effect, mainly from their lower educational attainments. However, occupational disparities seem to go further, especially for immigrant women. Their occupational sorting strongly penalizes them even after controlling for education (and other basic characteristics). In fact, immigrant women are highly concentrated in low-paying occupations in the counterfactual economy that we build in which the four groups are alike in terms of characteristics. Unlike them, the distribution of immigrant men across occupations is not especially harmful after accounting for characteristics.

By using an intersectionality approach, we can compare the wages of the four groups simultaneously. This has allowed us to move beyond comparisons between immigrants and natives of the same gender, as usually done in the literature. We show that, after removing the composition effect, the (conditional) hourly wages of immigrant women are well below those of the other three groups, mainly due to the occupational disadvantage of the former just mentioned. The (conditional) hourly wages of native women are also lower than those of native men, although occupations play a much smaller role in explaining their wage gap (with respect to the national average wage). As expected, in the counterfactual economy the gender wage gap among natives is larger than it is in the actual economy, due to the higher educational attainments of native women.[29] Native women's (conditional) hourly wages are closer to those of their immigrant male peers than to those of their native male peers.

[28] The gender gap with the 2022 EES is 9.4% for natives and 6.2% for immigrants. The nativity gap is 22.8% for men and 19.7% for women.

[29] The covariates included in the counterfactual analysis were intended to control for differences between natives and immigrants and to have enough observations in the corresponding cells for both immigrant women and men.

This paper contributes to the literature on intersectionality documenting that the gender wage gap in Spain is (slightly) greater among natives than it is among immigrants while the nativity gap is (slightly) greater among men than it is among women. The fact that the gender gap is smaller for minorities and/or that the minority-majority gap is larger for men is also detected in racial and ethnic studies for the US and some European countries (Algan et al., 2010; Alonso-Villar and Del Río, 2023b; Cantalini et al., 2023; Sprengholz and Hmjediers, 2024). However, in other countries, the gender wage gap for ethnic minorities or immigrants is larger than for the majority group (Algan et al., 2010; O'Higgins, 2015; Drolet and Amini, 2023), which evidences distinctive patterns among countries and also among minorities.

References

Abrahamson M. and Sigelman, L. (1987): "Occupational Sex Segregation in Metropolitan Areas," *American Sociological Review* 52, 588-597.

Adserà, A. and Querin, F. (2023): "The Gender Wage Gap and Parenthood: Occupational Characteristics Across European Countries," *European Journal of Population* 39: 34.

Aiginger, K. and Davies, S. (2004): "Industrial Specialization and Geographic Concentration: Two Sides of the Same Coin? Not for the European Union," *Journal of Applied Economics*, 7 (2), 231-248.

Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010): "The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom," *The Economic Journal* 120: F4-F30.

Allen, R. and Vignoles, A. (2007): "What Should an Index of School Segregation Measure?," *Oxford Review of Education* 33(5), 643-668.

Alonso-Villar, O. and Del Río, C. (2010a): "Local Versus Overall Segregation Measures," *Mathematical Social Sciences* 60(1), 30-38.

Alonso-Villar, O. and Del Río, C. (2010b): "Segregation of Female and Male Workers in Spain: Occupations and Industries," *Hacienda Pública Española* 194-(3/2010): 91-121.

Alonso-Villar, O. and Del Río, C. (2013a): "Concentration of Economic Activity: An Analytical Framework," *Regional Studies* 47(5), 756-772.

Alonso-Villar, O. and Del Río, C. (2013b): "Occupational Segregation in a Country of Recent Mass Immigration: Evidence from Spain," *Annals of Regional Science* 50, 109-134.

Alonso-Villar, O. and Del Río, C. (2017a), "Local Segregation and Wellbeing", *The Review of Income and Wealth*, vol. 63 (2), pp.269–287.

Alonso-Villar, O. and Del Río, C. (2017b), "The Occupational Segregation of African American Women: its Evolution from 1940 to 2010," *Feminist Economics* 23 (1), 108-134.

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

A more in-depth analysis of disparities between native women and men would require, in particular, a more detailed list of educational levels.

Alonso-Villar, O. and Del Río, C. (2017d): "Segregación ocupacional por razón de género y estatus migratorio en España y sus consecuencias en términos de bienestar," *Ekonomiaz* 91(1), 124-163.

Alonso-Villar, O. and Del Río, C. (2023a): "Disentangling Occupational Sorting from Within-Occupation Disparities: Earnings Differences among 12 Gender-Race/Ethnicity Groups in the U.S.," *Population Research and Policy Review* 42(3): 45.

Alonso-Villar, O. and Del Río, C. (2023b): "Privilege and Hindrance on the USA Earnings Distribution by Gender and Race/Ethnicity: An Intersectional Framework With 12 Groups," *International Journal of Manpower* 44(4), 635-652.

Alonso-Villar, O. and Del Río, C. (2023c): "Poverty among Same-Sex Couple Families in the United States: Is There a Premium for Married Couples?," *The Journal of Economic Inequality* 22, 495-517.

Alonso-Villar, O. and Del Río, C. (2024): "Dime o teu sexo (e nacionalidade) e direiche onde traballas," in I. Vázquez (coord.) *Non nacemos para coidar*. Vigo: Servizo de Publicacións, Universidade de Vigo, forthcoming.

Alonso-Villar, O., Del Río, C., and Gradín, C. (2012): "The Extent of Occupational Segregation in the United States: Differences by Race, Ethnicity, and Gender," *Industrial Relations* 51(2), 179-212.

Alonso-Villar, O., Gradín, C. and Del Río, C. (2013): "Occupational Segregation of Hispanics in U.S. Metropolitan Areas," *Applied Economics* 45 (30), 4298-4307.

Anghel, B., Conde-Ruiz, I., and Marra de Artíñano, I. (2019): "Brechas Salariales de Género en España," *Hacienda Pública Española* 229-(2/2019), 87-119.

Antón, J.I., Muñoz de Bustillo, R., and Carrera, M. (2010): "From Guests to Hosts: Immigrant-Native Wage Differentials in Spain," *International Journal of Manpower* 31(6), 645-659.

Atkinson, A. (1970): "On the Measurement of Inequality," *Journal of Economic Theory* 2, 244-263.

Azpitarte, F. and Alonso-Villar, O. (2014): "On the Measurement of Intermediate Inequality: A Dominance Criterion for a Ray-Invariant Notion," *Research on Economic Inequality* 22, 401-420.

Azpitarte, F., Alonso-Villar, O., and Hugo-Rojas, F. (2021): "Socioeconomic Groups Moving Apart: An Analysis of Recent Trends in Residential Segregation in Australia's Main Capital Cities," *Population, Space and Place* 27(3): e2399.

Besley, T. and Preston, I. (1988): "Invariance and the Axiomatics of Income Tax Progression: A Comment," *Bulletin of Economic Rese*arch 40, 159-163.

Bettio, F. and Verashchagina, A. (2009): "Gender Segregation in the Labour Market: Root Causes, Implications and Policy Responses in the EU." Luxembourg: Publications Office of the European Union.

Blau, F., Brummund, P., and Liu, B (2013): "Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System," *Demography* 50, 471-492.

Blau, F., and Kahn, L. (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature* 55(3), 789-865.

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(1), 161-171.

Borrowman, M. and Klasen, S. (2020): “Drivers of Gendered Sectoral and Occupational Segregation in Developing Countries,” *Feminist Economics* 26(2), 62-94.

Brülhart, M. and Traeger, R. (2005): “An Account of Geographic Concentration Patterns in Europe,” *Regional Science and Urban Economics* 35 (6), 597-624.

Cantalini, S., Guetto, R., and Panichella, N. (2023): “Ethnic wage penalty and human capital transferability: a comparative study of recent migrants in 11 European countries,” *International Migration Review* 57(1), 328-356.

Chantreuil, F. and Trannoy, A. (2013): “Inequality Decomposition Values: The Trade-Off between Marginality and Consistency,” *Journal of Economic Inequality* 11(1), 83-98.

Charles, M. and Grusky, D. (1995): “Models for Describing the Underlying Structure of Sex Segregation,” *American Journal of Sociology* 100(4), 931-971.

De la Rica, S., Glitz, A., and Ortega, F. (2014): “Immigration in Spain: What Have We Learned from Recent Evidence?,” *Cuadernos Ecońomicos de ICE* 87, 9-28.

Del Río, C. and Alonso-Villar, O. (2010): “New Unit-Consistent Intermediate Inequality Indices,” *Economic Theory* 42, 505-521.

Del Río, C. and Alonso-Villar, O. (2012a): “Occupational Segregation Measures: A Role for Status,” *Research on Economic Inequality* 20, 37-62.

Del Río, C. and Alonso-Villar, O. (2012b): “Occupational Segregation of Immigrant Women in Spain,” *Feminist Economics* 18(2), 91-123.

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.

Del Río C. and Alonso-Villar, O. (2018): “Segregation and Social Welfare: A Methodological Proposal with an Application to the U.S.,” *Social Indicators Research* 137, 257-280.

Del Río, C. and Alonso-Villar, O. (2019a): “Occupational Achievements of Same-Sex Couples in the U.S. by Gender and Race,” *Industrial Relations* 58 (4), 704-731.

Del Río, C. and Alonso-Villar, O. (2019b): “Occupational Segregation by Sexual Orientation in the U.S.: Exploring its Economic Effects on Same-Sex Couples,” *Review of Economics of the Household* 17 (2), 439-467.

Del Río, C. and Alonso-Villar, O. (2022): “On Measuring Segregation in a Multigroup Context: Standardized Versus Unstandardized Indices,” *Social Indicators Research* 163, 633-659.

Del Río C. and Ruiz-Castillo J. (2000): “Intermediate Inequality and Welfare,” *Social Choice and Welfare* 17, 223-239.

Deutsch, J., Flückiger, Y., and Silber, J. (2009): “Analyzing Changes in Occupational Segregation: The Case of Switzerland (1970-2000),” *Research on Economic Inequality* 17, 171-202.

DiNardo, J., Fortin, N., and Lemieux. T. (1996): “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach,” *Econometrica* 64(5), 1001-1044.

Drolet, M. and Amini, M.M. (2023): "Intersectional Perspective on the Canadian Gender Wage Gap," in *Studies on Gender and Intersecting Identities*, Statistics Canada, catalogue no. 45200002.

Duncan, O. and Duncan, B. (1955): "A Methodological Analysis of Segregation Indexes," *American Sociological Review* 20(2), 210-217.

Elbers, B. (2023): "A Method for Studying Differences in Segregation Across Time and Space," *Sociological Methods and Research* 52(1), 5-42.

Foster, J. E. (1985): "Inequality Measurement," in H. Peyton Young (ed.), *Fair allocation. Proceedings of Symposia in Applied Mathematics* 33, 31–68. Providence RI: American Mathematical Society.

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

García-Pérez, I., Muñoz-Bullón, F., and Prieto-Rodríguez, M. (2012): "The Wage Gap between Foreign and Spanish Nationals in Spain: an Analysis Using Matched Employer-Employee Data," *International Migration* 52(6), 165-179.

Gorard, S. and Taylor, C. (2002): "What Is Segregation? A Comparison of Measures in Terms of 'Strong' and 'Weak' Compositional Invariance," *Sociology* 36(4), 875-895.

Gradín, C. (2013): "Conditional Occupational Segregation of Minorities in the US," Journal of Economic Inequality 11(4), 473-493.

Gradín, C. (2019): "Segregation of Women into Low-Paying Occupations in the United States," *Applied Economics* 52(17), 1905–1920.

Gradín, C., Del Río, C. and Alonso-Villar, O. (2015), "Occupational Segregation by Race and Ethnicity in the US: Differences across States," *Regional Studies* 49 (10), 1621-1638.

Huffman, M. and Cohen, P. (2010): "Engendering Change: Organizational Dynamics and Workplace Gender Desegregation, 1975-2005," *Administrative Science Quarterly* 55, 255-277.

Hutchens, R. (1991): "Segregation Curves, Lorenz Curves, and Inequality in the Distribution of People across Occupations," *Mathematical Social Sciences* 21, 31-51.

Hutchens, R. (2001): "Numerical Measures of Segregation: Desirable Properties and their Implications," *Mathematical Social Sciences* 42, 13-29.

Hutchens, R. (2004): "One Measure of Segregation," *International Economic Review* 45 (2), 555-578.

Iceland, J., Weinberg, D., and Hughes, L. (2014): "The Residential Segregation of Detailed Hispanic and Asian Groups in the United States: 1980-2010," *Demographic Research* 20, 593-624.

Jahn, J., Schmid, C., and Schrag, C. (1947): "The Measurement of Ecological Segregation," *American Sociological Review* 12(3), 293-303.

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

Karmel, T. and MacLachlan, M. (1988): "Occupational Sex Segregation—Increasing or Decreasing?," *The Economic Record* 64(3), 187-195.

Kolm, S. (1976): "Unequal Inequalities I." *Journal of Economic Theory* 12, 416-442.

Krtscha M. (1994): “A New Compromise Measure of Inequality,” in: W. Eichorn (ed) *Models and Measurement of Welfare and Inequality*. Heidelberg: Springer.

Ledic´, M., Rubil, I., and Urban, I. (2023): “Tax Progressivity and Social Welfare with a Continuum of Inequality Views,” *International Tax and Public Finance* 30, 1266-1296.

Lorence, J. (1992): “Service Sector Growth and Metropolitan Occupational Sex Segregation,” *Work and Occupations* 19, 128-156.

Marcińczak, S., Musterd, S., van Ham, M., and Tammaru, T. (2016): “Inequality and Rising Levels of Socio-Economic Segregation: Lessons from a Pan-European Comparative Study,” in T. Tammaru, S. Marcińczak, M. Ham, and S. Musterd (eds.) *Socioeconomic Segregation in European Capital Cities: East Meets West*, 358-382. London: Routledge.

Massey, D. and Denton, N. (1988): “The Dimensions of Residential Segregation,” *Social Forces* 67(2), 281-315.

Mazza, A. and Punzo, A. (2015): “On the Upward Bias of the Dissimilarity Index and Its Corrections,” *Sociological Methods and Research* 44(1), 80-107.

Mintz, B. and Krymkowski, D. (2011): “The Intersection of Race/Ethnicity and Gender in Occupational Segregation,” *International Journal of Sociology* 40(4), 31-58.

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.

Moyes, P. (1987): “A New Concept of Lorenz Domination,” *Economics Letters* 23, 203-207.

Moyes, P. (1992): “The Through-Time Redistributive Effect of Income Taxation,” *Mathematical Social Sciences* 24(1), 59-71.

Mulligan G. and Schmidt, C. (2005): “A Note on Localization and Specialization,” *Growth and Change* 36(4), 565-576.

Nicodemo, C. and Ramos, R. (2012): “Wage Differentials Between Native and Immigrant Women in Spain,” *International Journal of Manpower* 33(1), 118-136.

O’Higgins, N. (2015): “Ethnicity and gender in the labour market in Central and South-Eastern Europe” Cambridge Journal of Economics 39(2), 631-654.

Palencia-Esteban, A. (2022): “Occupational Segregation of Female and Male Immigrants in Europe: Accounting for Cross-Country Differences,” *International Labour Review* 161(3), 341-373.

Palencia-Esteban, A. and Del Río, C. (2024): “Winners and Losers from Occupational Segregation Across Europe: The Role of Gender and Migration Status,” *Migration Studies* 12(1), 21-41.

Pfingsten, A. (1986): *The Measurement of Tax Progression*. Berlin: Springer.

Pinto, F., Martínez, R., Delgado Rodríguez, M.J., and Murillo, E. (2023): “The Migrant Pay Gap in Spain: Where Do the Differences Come From?,” *The Economic and Labour Relations Review* 34, 468-490.

Queneau, H. (2009): “Trends in Occupational Segregation by Race and Ethnicity in the USA: Evidence from Detailed Data,” *Applied Economics Letters* 16 (13), 1347-1350.

Reardon, S. (2009): “Measures of Ordinal Segregation,” *Research on Economic Inequality* 17, 129-155.

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

Reardon, S. and O’Sullivan, D. (2004): “Measures of Spatial Segregation,” *Sociological Methodology* 34, 121-162.

Reskin, B. (1999): “Occupational Segregation by Race and Ethnicity Among Women Workers,” in I. Browne (ed.), *Latinas and African American Women at Work: Race, Gender, and Economic Inequality*, 183-204. New York: Russell Sage.

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

Sarabia, J.M. and Jordá, V. (2014): “Explicit Expressions of the Pietra Index for the Generalized Function for the Size Distribution of Income,” *Physica A* 416, 582-595.

Seidl C. and Pfingsten A. (1997): “Ray Invariant Inequality Measures,” in: S. Zandvakili and D. Slotje (eds) *Research on Taxation and Inequality*. Greenwich: JAI Press.

Shorrocks, A. (2013): “Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value,” *Journal of Economic Inequality* 11(1), 99-126.

Silber, J. (1989): “On the Measurement of Employment Segregation,” *Economics Letters* 30(3), 237-243.

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

Silber, J. (2012): “Measuring Segregation: Basic Concepts and Extensions to Other Domains,” *Research on Economic Inequality* 20, 1-35.

Sprengholz, M. and Hamjediers, M. (2024): “Intersections and commonalities: using matching to decompose wage gaps by gender and nativity in Germany” Work and Occupations 5(2), 249-286.

Theil, H. (1972): *Statistical Decomposition Analysis*. Amsterdam, the Netherlands: North Holland.

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(2), 187-194.

Yao, J., Wong, D., Bailey, N., and Minton, J. (2018): “Spatial Segregation Measures: A Methodological Review,” *Tijdschrift voor Economische en Sociale Geografie* 110(3), 235-250.

Yoshida, T. (2005): “Social Welfare Rankings of Income Distributions. A New Parametric Concept of Intermediate Inequality,” *Social Choice and Welfare* 24, 557-574.

Zheng, B. (2007): “Unit Consistent Decomposable Inequality Measures,” *Economica* 74 (293), 97-111.

Appendix

Table A1. Basic characteristics of the groups in 2006 and 2024.

	2006 EPA					2024 EPA				
	Native men	**Immigrant men**	**Native women**	**Immigrant women**	**Total**	**Native men**	**Immigrant men**	**Native women**	**Immigran t women**	**Total**
Sample size	37,497	2,029	25,959	1,638	67,123	24,401	2,217	22,685	2,032	51,335
Population share (%)	52.4	6.8	35.6	5.2	100	45.5	8.0	39.7	6.8	100
Educational attainment (%)										
Lower secondary education at most	50.0	46.4	36.9	36.6	44.4	31.9	49.3	22.1	36.9	29.7
Upper secondary education	20.7	33.0	22.7	39.4	23.2	23.6	27.1	22.4	30.7	23.9
Tertiary education	29.4	20.7	40.4	24.0	32.4	44.6	23.6	55.6	32.5	46.4
Age (%)										
Up to 30 years old	22.4	30.8	25.2	36.1	24.7	14.0	19.6	14.1	17.6	14.7
Between 30 and up to 55	63.9	65.4	64.9	60.3	64.2	63.1	66.5	63.5	71.9	64.1
Over 55	13.6	3.7	9.8	3.6	11.1	22.9	13.8	22.5	10.5	21.1
Location (%)										
Regions with wages below the national average	63.7	53.7	59.7	52.2	61.0	62.8	53.5	60.8	51.3	60.5
Regions with wages above the national average	36.3	46.3	40.3	47.8	39.0	37.2	46.6	39.2	48.7	39.5
Citizenship (%)										
Bulgaria & Romania		15.5		18.6	17.6		14.4		15.8	15.0
Rest of Europe		15.6		15.7	16.1		18.5		22.2	20.2
Africa		23.3		7.4	17.0		20.6		8.1	14.9
Central and South America		42.3		56.0	46.5		37.1		46.1	41.2
Asia		3.2		2.0	2.4		8.9		7.1	8.1
Rest of the world and stateless		0.2		0.3	0.3		0.5		0.7	0.6

Source: Author's calculations based on the EPAs.

Table A2. Evolution of overall and local segregation (I_p, GI_p and D^g), 2006-2024 EPAs.

	Overall Segregation: I_p and GI_p indices			**Local Segregation: D^g index**			
	Gender	**Nativity**	**Gender-Nativity**	**Native men**	**Immigrant men**	**Native women**	**Immigrant women**
2006	0.266	0.085	0.289	0.230	0.400	0.318	0.558
2007	0.268	0.091	0.290	0.228	0.415	0.313	0.558
2008	0.266	0.093	0.290	0.229	0.416	0.305	0.534
2009	0.265	0.088	0.288	0.235	0.393	0.296	0.535
2010	0.262	0.085	0.282	0.236	0.383	0.287	0.513
2011	0.256	0.083	0.278	0.233	0.381	0.281	0.502
2012	0.248	0.077	0.270	0.233	0.355	0.272	0.473
2013	0.245	0.082	0.272	0.237	0.379	0.267	0.511
2014	0.245	0.071	0.266	0.232	0.367	0.267	0.466
2015	0.245	0.070	0.266	0.229	0.356	0.270	0.485
2016	0.244	0.071	0.266	0.228	0.372	0.271	0.481
2017	0.246	0.071	0.267	0.229	0.376	0.271	0.461
2018	0.244	0.072	0.267	0.230	0.377	0.268	0.462
2019	0.244	0.073	0.265	0.228	0.355	0.269	0.447
2020	0.238	0.070	0.261	0.222	0.357	0.269	0.431
2021	0.245	0.073	0.266	0.227	0.382	0.274	0.405
2022	0.243	0.078	0.265	0.226	0.378	0.269	0.413
2023	0.237	0.083	0.261	0.223	0.366	0.260	0.421
2024	0.234	0.085	0.262	0.226	0.371	0.258	0.400

Source: Alonso-Villar and Del Río (2024) and author's calculations based on the EPAs.

Table A3. Evolution of overall and local segregation (M and Φ_1^g), 2006-2024 EPAs.

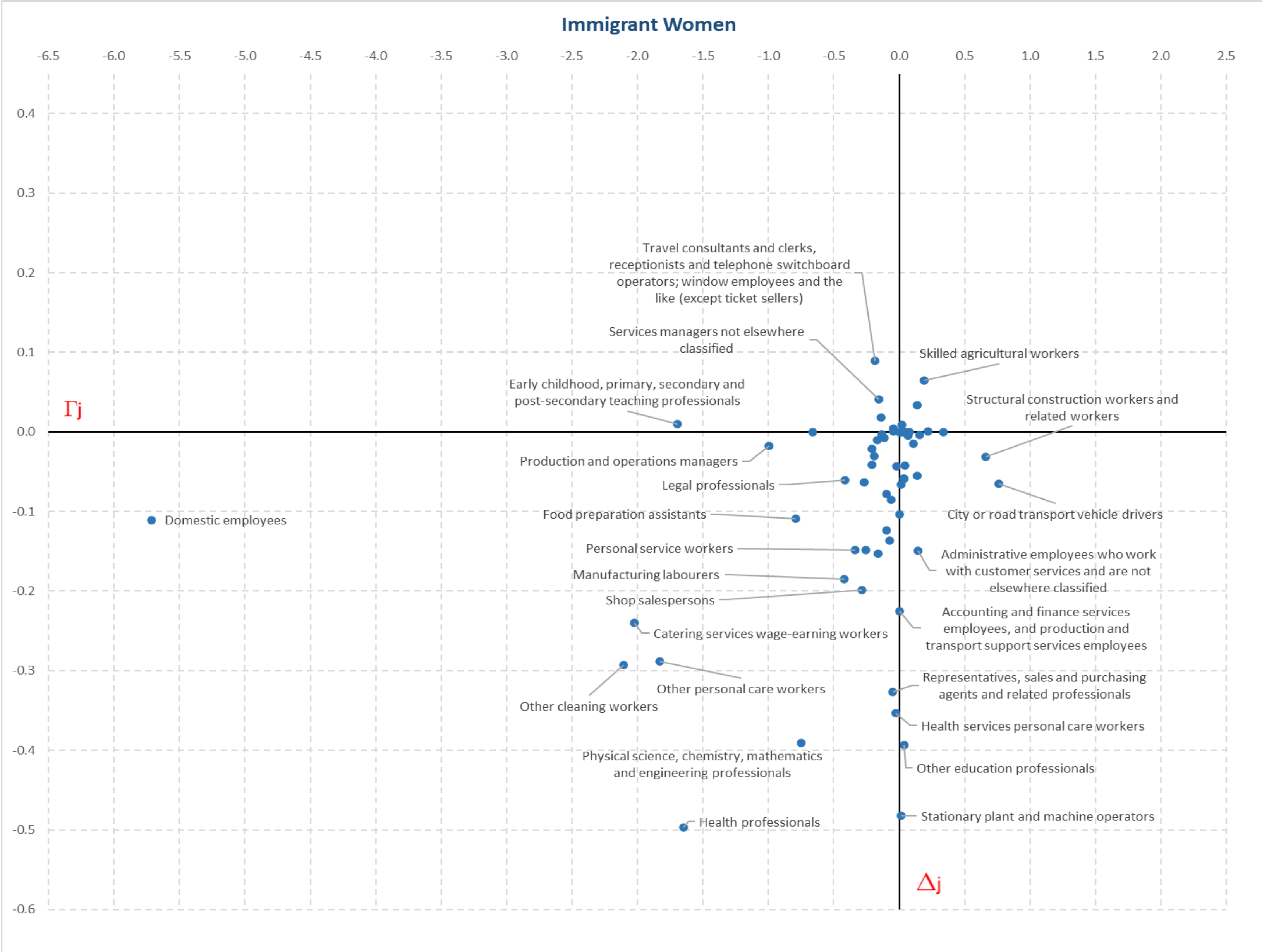
	Overall Segregation: M index			**Local Segregation: Φ_1^g index**			
	Gender	**Nativity**	**Gender-Nativity**	**Native men**	**Immigrant men**	**Native women**	**Immigrant women**
2006	0.224	0.055	0.281	0.153	0.502	0.321	1.010
2007	0.224	0.058	0.285	0.155	0.541	0.311	0.962
2008	0.222	0.058	0.282	0.157	0.515	0.302	0.904
2009	0.215	0.057	0.274	0.164	0.463	0.280	0.889
2010	0.209	0.055	0.266	0.163	0.442	0.265	0.877
2011	0.203	0.052	0.256	0.162	0.441	0.249	0.838
2012	0.190	0.049	0.241	0.160	0.413	0.233	0.772
2013	0.190	0.053	0.245	0.164	0.458	0.228	0.848
2014	0.188	0.045	0.235	0.161	0.405	0.230	0.793
2015	0.188	0.043	0.233	0.156	0.400	0.233	0.786
2016	0.189	0.043	0.234	0.154	0.438	0.234	0.762
2017	0.188	0.042	0.232	0.153	0.436	0.235	0.722
2018	0.188	0.042	0.232	0.155	0.429	0.232	0.706
2019	0.186	0.039	0.227	0.151	0.391	0.232	0.647
2020	0.177	0.038	0.217	0.143	0.398	0.226	0.602
2021	0.182	0.039	0.223	0.145	0.434	0.233	0.569
2022	0.181	0.042	0.226	0.145	0.437	0.228	0.593
2023	0.175	0.044	0.221	0.142	0.412	0.217	0.586
2024	0.171	0.044	0.218	0.143	0.382	0.214	0.549

Source: Author's calculations based on the EPAs.

Table A4. The hourly wage advantage or disadvantage (*EGap*) and within-between decomposition (Γ and Δ) in the actual economy and counterfactuals, 2024 EPA and 2022 EES.

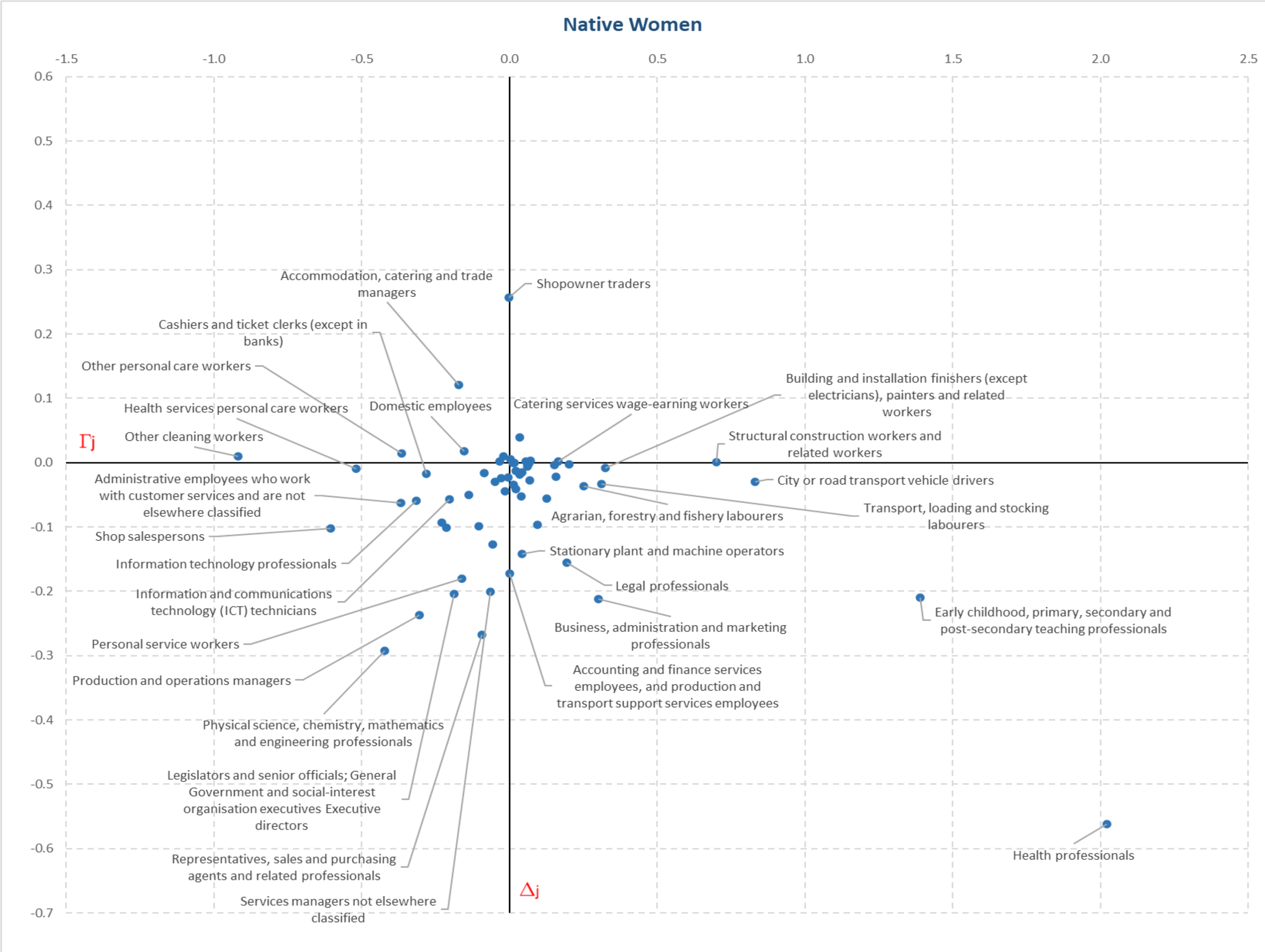
	Employment from the 2024 EPA			
Wages from the 2022 EES	**Native men**	**Immigrant men**	**Native women**	**Immigrant women**
EGap	8.263	-16.066	-2.076	-24.656
Γ	3.532	-12.572	1.686	-19.009
Δ	4.731	-3.494	-3.762	-5.646
$EGap_p$	9.431	-7.273	-5.799	-20.926
Γ_p	4.760	-3.523	-2.242	-14.771
Δ_p	4.672	-3.750	-3.557	-6.155
$EGap_{np}$	9.437	-7.593	-5.770	-20.759
Γ_{np}	4.770	-4.018	-2.185	-14.599
Δ_{np}	4.667	-3.575	-3.586	-6.160
	Employment from the 2022 EES			
Wages from the 2022 EES	**Native men**	**Immigrant men**	**Native women**	**Immigrant women**
EGap	6.459	-16.344	-2.901	-22.552
Γ	1.868	-13.547	1.096	-15.777
Δ	4.591	-2.797	-3.997	-6.775
$EGap_p$	8.132	-1.635	-7.282	-14.190
Γ_p	3.549	-1.923	-2.800	-8.685
Δ_p	4.583	0.289	-4.483	-5.505
$EGap_{np}$	8.143	-3.360	-7.033	-14.737
Γ_{np}	3.534	-2.442	-2.678	-9.143
Δ_{np}	4.609	-0.919	-4.355	-5.595

Source: Author's calculations based on the 2024 EPA and 2022 EES.



Source: Author’s calculations using employment from the 2024 EPA and wages from the 2022 EES.

Figure A1. Contribution of each occupation to the within and between components of the actual earning gap, immigrant women

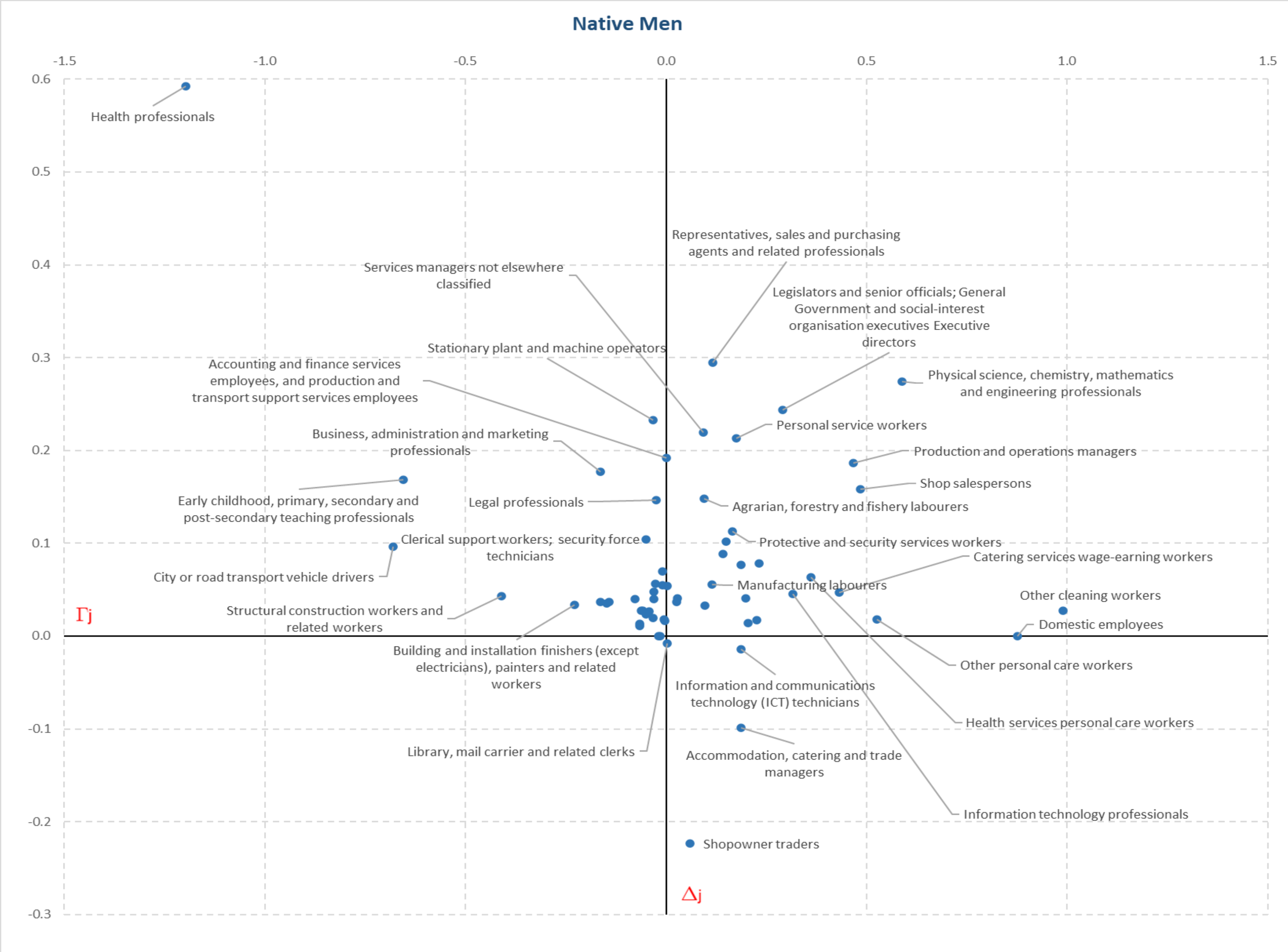


Source: Author's calculations using employment from the 2024 EPA and wages from the 2022 EES.

Figure A2. Contribution of each occupation to the within and between components of the actual earning gap, native women

Source: Author's calculations using employment from the 2024 EPA and wages from the 2022 EES.

Figure A3. Contribution of each occupation to the within and between components of the actual earning gap, immigrant men



Source: Author’s calculations using employment from the 2024 EPA and wages from the 2022 EES.

Figure A4. Contribution of each occupation to the within and between components of the actual earning gap, native men