# Measuring gender segregation: Evolution and decomposition analysis in the Brazilian labor market

Diego Camargo Botassio[1]
Rodolfo Hoffmann[2]

**ABSTRACT**

This paper analyzes the evolution and decomposition of gender segregation in the Brazilian labor market. First, we define the segregation measures. Further, we examine some characteristics of the women's participation in the labor market, the evolution of gender segregation from 1992 to 2015, and its additively decomposition from 2002 to 2015, all analyses based on the Brazilian National Household Sample Survey microdata. The empirical evidence of the study suggests that gender segregation has decreased considering both levels of analysis: the 53 sectors of activity (varying from 22% to 8%, depending on the index) and the 7 groups of activities (from 23.5% to 7.2%). The additively decomposition showed that, regardless of the measure used, the between-groups term explains at least partially this decreasing trend.
**KEYWORDS**: Gender segregation; Additively decomposition; Labor economics; Brazil; Labor market.

**RESUMO**

Este artigo analisa a evolução e a decomposição da segregação por gênero no mercado de trabalho brasileiro. Primeiro são apresentadas as medidas de segregação. Posteriormente, são analisadas algumas características da participação das mulheres no mercado de trabalho, a evolução da segregação por gênero de 1992 a 2015 e sua decomposição aditiva de 2002 a 2015, todas análises utilizando os microdados da Pesquisa Nacional por Amostra de Domicílios. As evidências empíricas sugerem que a segregação por gênero decresceu consideravelmente em ambos níveis de análise: para os 53 setores de atividade (variando de 22% a 8%, dependendo do índice) e para os 7 grupos de atividade (de 23,5% a 7,2%). A decomposição aditiva mostra que, independentemente da medida utilizada, a parcela entre-grupos explica em grande parte a tendência de queda da segregação.
**PALAVRAS-CHAVE**: Segregação por gênero; Decomposição aditiva; Economia do Trabalho; Brasil; Mercado de trabalho.

**Área 2: Desenvolvimento Econômico**

**JEL**: C60; J01; J16; J82

[1] Professor at Department of Economics, State University of Maringá (UEM). Email: dcbotassio2@uem.br.
[2] Professor at 'Luiz de Quiroz' College of Agriculture, University of São Paulo (ESALQ/USP). Email: hoffmannr@usp.br.

# 1 Introduction



Gender segregation in the labor market is an interesting phenomenon that cannot be neglected and, in the long-run, its decrease may reflect gender equality policies. Since the first studies on this issue (Gross, 1968; Oppenheimer, 1970), this branch of the welfare economics has been expanding continuously. Segregation is often related to discrimination and segmentation, and, specifically gender segregation in the labor market, and skin color segregation.

Gender segregation in the labor market indicates that men tend to be occupied in some sectors of activity, while the women concentrate on others. On the other hand, discrimination and segmentation are also related to individual income. For the segregation measurement, the income is not taken into account, because the relevant information regards the sectors of activities in which the individuals are allocated. For the issue relating segregation to income, specifically the wage gap in Brazil, see an interesting analysis of how the wage gap can be accounted for the degree of gender segregation by Foguel (2016).

In addition to the study of segregation by gender in the labor market, after the work by Albelda (1986), the analysis of occupational segregation by skin color became common. Some authors (e.g. King, 2009; Salardi, 2016) analyze these two situations to Brazil. Ribeiro and Araújo (2016), for example, prefer to analyze only the distinction by color and extend the discussion to the level of schooling. Although the development of segregation measures allows both analyzes, the sociological phenomena in the two situations are substantially different.

The sociological interpretation of segregation by color, in general, is related to the geographical differences between the categories, as already indicated in the works on spatial distribution between whites and blacks in American cities (Jahn et al., 1947; Duncan and Duncan, 1955b; Taeuber and Taeuber, 1965; Schelling, 1971). On the other hand, Anker (1997) highlights that occupational segregation by gender is related to the demand and supply aspects of work by men and women. In order to avoid comparisons in these two different sociological situations, we aim to measure only gender segregation by sectors of activity in Brazil.

The studies regarding the evolution of occupational gender segregation in Brazil analyze the data from the 1980s to the mid-2000s. The works that we highlight use the segregation indexes (dissimilarity or Gini) for two or more periods. Other works, and some of those listed ahead, also use a measure attributed to Karmel and MacLachlan (1988), but we do not use it or report its results to Brazil due to a critique associated with this index and explained in detail in Section 2. For 80s the works by Oliveira (2001) and Ometto et al. (1997) stand out. Both use the Brazilian National Household Sample Survey (PNAD, the abbreviation for *Pesquisa Nacional por Amostra de Domicílios* in Portuguese) data and perform their analyzes based on the index of dissimilarity. Oliveira (2001) uses data from 1981 to 1999, exclusive 1991 and 1994, considering occupations in three different levels of aggregation. She concludes that segregation has decreased 9% for the less aggregated analysis, while it has increased for more aggregated one, both considering the 80s and 90s data. Ometto et al. (1997) use data from 1981 to 1990, exclusive 1982 and 1986, for the States of São Paulo and Pernambuco. The work leads to the conclusions that segregation declined slightly for State of São Paulo, but rose up substantially in Pernambuco.

Also for the 90s and onwards, we highlight the works by King (2009) and Vaz and Hoffmann (2011). The first uses the 1989 and 2001 PNAD data and reach similar conclusions to Oliveira (2001). Vaz and Hoffmann (2011) analyze occupational segregation by sex in the Brazilian public sector from 1995 to 2008, using the Annual Report on Social Information (RAIS - *Relação Anual de Informações Sociais*) microdata. The indexes used by the authors are the Gini index for segregation and the index of dissimilarity. When analyzing the public sector, where the vast majority of occupations are formal employment, the RAIS database is an excellent source of information because of its census characteristic. The authors' conclusions indicate that segregation has generally decreased in the three spheres of government, especially at the counties and federal levels. On the studies about occupational segregation by gender in Brazil, this is the only one that does not use PNAD data. In this paper, we prefer to use the PNAD data to also incorporate the non-formal sector information.

Salardi (2016) uses the PNAD data from 1987 to 2006, except 1991, 1994 and 2000. In 2002, PNAD began to adopt a new classification for occupations. She is the first author that harmonizes the data for comparison in segregation analysis to Brazil. The harmonization used was also proposed by her (Salardi, 2013), and it results in 83 occupational codes. In this sense, our paper is similar to hers, but we harmonize the data for the groups of activity. She concludes that gender segregation decreased 6.6% in the period (index of dissimilarity), and notices an initial increase at the beginning of the 1990s, followed by a significant decline from 1995 to 2006 (the Gini index results are similar). We did not find papers that covered the 2010s analyzing this issue. For this reason, we analyze the gender segregation from 1992 to 2015, because this is the longest period available for which data can be harmonized for comparison.

A very interesting way to approach the segregation analysis is its additively decomposition. In this way, we can decompose the segregation into two components: one related to the segregation between groups of activity, and another one regarding the segregation within these groups. Very few studies in the literature use this approach to gender segregation in the labor market (e.g. Mora and Ruiz-Castillo, 2003), perhaps by the fact that the earlier studies are about racial segregation in the schools (Theil and Finizza, 1971) or because the properties of these measures have been exhaustively analyzed just recently (Hutchens, 2004; Frankel and Volij, 2011). We did not find papers that make the additively decomposition analysis of gender segregation in Brazil.

Taking into account the remarks made so far, this study aims to measure the gender segregation in the Brazilian labor market from 1992 to 2015 considering the sectors of activity. To do so we use two complementary approaches. The first is the temporal evolution of the segregation measures, and the second is the additively decomposition analysis.

This paper is structured as follows. Section 2 presents the indexes used in this paper. They are the index of dissimilarity, the Gini index, and the class of additively decomposable segregation measures. In Section 3 we describe the dataset and the sample delimitation. Section 4 presents some descriptive statistics regarding the labor market structure by gender. In subsections 4.1 and 4.2, we present the results about evolution and decomposition, respectively. Finally, in Section 5, we summarize our main results.

## 2 Methodology: Segregation measurement

Before describing the segregation measures used in this paper, we formally define the notation for segregation measurement signaling its extreme cases. Consider an economy divided into $J$ sectors and two types of people (women and men) represented by $X$ and $Y$. Let $j = 1, 2, \ldots, J$ be the $j$th sector, $X_j$ the number of women allocated in $j$, and $x_j = X_j/X$ its share in the total of women, with $X = \sum_{j=1}^{J} X_j$. Similar definitions follow for $Y_j$, $y_j$, and $Y$ representing the men's category. Let $\mathbf{D}$ be the $2 \times J$ matrix distribution, whose elements are the quantities of women and men in the $J$ sectors. A measure of segregation for an economy divided into $J$ sectors is a mapping from any $\mathbf{D}$ to a non-negative real number. We highlight two extreme situations. If the women's share in the $j$ sector is equal to the respective men's share, for all sectors, this is the no-segregation case. In general, it corresponds to the value zero of any measure. On the other hand, if all sectors are occupied exclusively by one category, this is the completely segregated case.

Undoubtedly, the most widely used measure of segregation in the literature is the index of dissimilarity $[D(\mathbf{D})]$. It was proposed by Jahn et al. (1947) and gained noticeable popularity after the work by Duncan and Duncan (1955a). In mathematical terms, it is

$$D(\mathbf{D}) = 0.5 \sum_{j=1}^{J} |x_j - y_j|. \tag{1}$$

According to Eq. (1), $D(\mathbf{D}) = 1$ if and only if all sectors are occupied exclusively by one gender (which correspond to the completely segregated case). If $x_j = y_j$ for all sectors, then $D(\mathbf{D}) = 0$ (the no-segregation case). So, for any distribution $\mathbf{D}$, we must have $D(\mathbf{D}) \in [0, 1]$.

Frequently, applied works on segregation use an index assigned to Karmel and MacLachlan (1988). It is defined as $I_P(\mathbf{D}) = 2XYD(\mathbf{D})/T^2$, where $T = X + Y$. This index can be found in at least two works previously to Karmel and MacLachlan (1988). They are Jahn et al. (1947, p. 294) and Walker et al. (1967, p. 6). Also, this measure does not satisfy two important desirable properties for segregation measures.

A desirable property that a measure should satisfy is the movements between strata (Hutchens, 2004). For measures that satisfy this property, regressive movements (also referred to as disequalizing movement) imply the increase of the measure. In short, a less desirable situation implies a higher value of segregation. For inequality, this condition is known as the Pigou-Dalton principle of transfers. It is well known that the dissimilarity index does not satisfy this property (Duncan and Duncan, 1955a; James and Taeuber, 1985), and consequently, the index $I_P(\mathbf{D})$ also does not.

The second property refers to the increase of one of the categories (Botassio, 2017). Defining $P = X/T$, we have $I_P(\mathbf{D}) = 2P(1 - P)D(\mathbf{D})$. That is, segregation depends on the composition (shares) of the categories in the population. When measuring gender segregation in the labor market, this fact must be taken into account. In general, the participation of women in the labor market is less than 1/2, and in several countries of the world, it is approaching one half. Thus, this index could indicate that segregation increased simply because of this phenomenon [$P(1 - P)$ is an increasing function for $P < 1/2$]. For an empirical example, this article shows that, from 1992 to 2001, segregation decreased 2.1% according to the index of dissimilarity, but increased 1.6% using the $I_P(\mathbf{D})$, because the share of women in the population increased 3.1 percentage points (*p.p.*). The women's increasing participation in the labor market is an important phenomenon, but it is better to differentiate it from segregation. It is not appropriate to mix the increasing of this participation and changes in the degree of segregation in the same index. Thus, in order to do not affect the results with the increasing participation of women in the labor market (also observed for Brazil), we do not use this measure. For examples of its application to Brazil, see Salas and Leite (2007) and Salardi (2016).

Other common measure in the literature is the Gini index for segregation proposed by Duncan and Duncan (1955a). For its definition, it is useful the concept of the segregation curve. Let us sort the $J$ sectors by their ratios $X_j/Y_j$, that is $X_{(j)}/Y_{(j)} \leq X_{(j+1)}/Y_{(j+1)}$, for $j = 1, 2, \ldots, J - 1$, where $(j)$ represents the $j$th sorted position. For effect of notation, denote $x_{(0)} = y_{(0)} = 0$. Let $\Psi_j = \sum_{k=0}^{j} x_{(k)}$ and $\rho_j = \sum_{k=0}^{j} y_{(k)}$ be the cumulative proportions of women and men, respectively. The segregation curve is defined by the line segments joining all points $\{(\rho_j, \Psi_j)\}_{j=0}^{J}$. The segment bounded by points $(0, 0)$ and $(1, 1)$ is the no-segregation line, and it is a special case of the segregation curve when $x_j = y_j$ for all $j$. So, the Gini index for segregation is defined by twice the area between the no-segregation line and the segregation curve. In mathematical terms, following Duncan and Duncan (1955a), this index is[3]

$$G(\mathbf{D}) = \sum_{j=1}^{J} (\Psi_j \rho_{j-1} - \Psi_{j-1} \rho_j). \tag{2}$$

It is not difficult to demonstrate that $G(\mathbf{D}) \in [0, 1]$, with zero and one corresponding to the extreme cases of no-segregation and completely segregation, similar to the index of dissimilarity. In the literature, the number of different expressions for the Gini index for segregation and the index of dissimilarity is striking. Botassio (2017) demonstrates the equivalence of a significant number of these equations. For example, the Duncan and Duncan (1955a) expressions for both indexes are equivalents to those found in Hoffmann (1998) and James and Taeuber (1985).

The last index is the class of additively decomposable segregation measures (hereinafter general measure of segregation) proposed by Hutchens (2004). Its mathematical expression is

$$I_\varepsilon(\mathbf{D}) = \frac{1}{\varepsilon(1 - \varepsilon)} \left( 1 - \sum_{j=1}^{J} y_j^{\varepsilon} x_j^{1-\varepsilon} \right). \tag{3}$$

[3] We note that the correct expression is Eq. (2), since that found in Duncan and Duncan (1955a, p. 211) must be multiplied by minus one. This fact is demonstrated by Botassio (2017).

Note that this measure depends on a parameter ($\varepsilon$), which can be any real number, usually from $-1$ to $1$. This parameter, as demonstrated by Botassio and Hoffmann (2019), is a segregation aversion parameter regarding the types' distribution in the sectors according to the ratios $X_j/Y_j$. We always have $I_\varepsilon(\mathbf{D}) = 0$ for the no-segregation case, upper limit $1/[\varepsilon(1-\varepsilon)]$ for $0 < \varepsilon < 1$, and this measure is unbounded for $\varepsilon \leq 0$ or $\varepsilon \geq 1$ (Hutchens, 2004).

Some of the usual values for $\varepsilon$ are $-1$ (transformation of the coefficient of variation for segregation - Hutchens, 1991, p. 48), $\varepsilon = 0$ (Theil's $T$ index for segregation - Hutchens, 1991, p. 48), $\varepsilon = 0.5$ (Square Root index multiplied by 4 - Hutchens, 2001, 2004), and $\varepsilon = 1$ (Theil's $L$ index for segregation). These are some of the values for $\varepsilon$ used in this paper.

The main advantage of the general measure of segregation, compared to the Gini index and the index of dissimilarity, is that it is additively decomposable. This means that segregation can be decomposed in two terms, one referring to the segregation between-groups of sectors and the other is a weighted sum of segregation within-groups. Let $g$ denote a group of sectors ($g = 1, 2, \ldots, G$). Allocate the $J$ sectors in the $G$ groups, so that $I_g(\mathbf{D}_g)$ represents the measure of segregation within the $g$th group with $n_g$ sectors, and $\sum_{g=1}^{G} n_g = J$. Since $x_g$ and $y_g$ represent the shares of each category in the $g$th group, the additively decomposition of Eq. (3) is

$$I_\varepsilon(\mathbf{D}) = \sum_{g=1}^{G} y_g^\varepsilon x_g^{1-\varepsilon} I_g(\mathbf{D}_g) + \frac{1}{\varepsilon(1-\varepsilon)}\left(1 - \sum_{g=1}^{G} y_g^\varepsilon x_g^{1-\varepsilon}\right). \tag{4}$$

In Eq. (4), the second term on the right-hand side is the between-groups term, and the first one is the within-groups component, which corresponds to a weighted sum of the within segregation measures. In this equation, $y_g^\varepsilon x_g^{1-\varepsilon}$ is the weight attached to the $g$th group. For those interested in a more detailed analysis of the special cases of the general measure by $\varepsilon$ values and the properties of the additively decomposition, see Botassio and Hoffmann (2018). These authors make an important remark regarding the parameter's choice in the decomposition. For the Theil's indexes, i.e., considering $\varepsilon = 0$ or 1, the between-groups and the within-groups terms are independent. In this sense, these values for $\varepsilon$ are preferably compared to the others in the decomposition analysis. Finally, we point out that the Gini index and the general measure satisfy the movement between strata property.

We use the PNAD's datasets to analyze the trend and decomposition of gender segregation in Brazil. The next section describes these data and the sample's delimitation.

## 3 Data

We use the PNAD microdata (individual level of information) from 1992 to 2015, except the years 1994, 2000 and 2010 in which the survey was not conducted. The PNAD is carried out by the Brazilian Institute of Geography and Statistics (IBGE - *Instituto Brasileiro de Geografia e Estatística*) and collects information about the individuals, such as schooling, employment, etc., and conditions of the household. Its scope is national since 2004 and allows population analysis using the sample expansion factor. The datasets are public and available at IBGE (2017).

Until 2003, the PNAD did not cover the rural area of the former North Region, which corresponds to the current one, except for the State of Tocantins. For this reason, we exclude the observations for this area from 2004 onwards to make the results comparable. Thus, in this article, when analyzing the time series results, "Brazil" means the Country without the rural area of the former North Region. Before 2002, PNAD's activities and occupations were classified according to the National Classification of Economic Activities (CNAE - *Classificação Nacional de Atividades Econômicas*) and the Brazilian Classification of Occupations (CBO - *Classificação Brasileira de Ocupações*), respectively. From 2002 onwards, IBGE adopted the CNAE-Domiciliar to classify activities and the CBO-Domiciliar to occupations. For the correct analysis of the results, it is indispensable to harmonize the data from one period to be comparable to the other.

We harmonized the variable "main groups of activity of the main job" (variable V4709 until 2001, and V4809 from 2002 onward) following the procedures established by IBGE.[4] This variable is used to analyze the segregation evolution, from 1992 to 2015. By harmonizing the data using the IBGE recommendation, this is the longest period of analysis available.

To analyze the decomposition, besides the group of activity variable, we need a more disaggregated one. Unfortunately, the harmonization of the less aggregated variable (sectors of activity - V9907) is not possible, because the classification of the CNAE and the CNAE-Domiciliar differ substantially for this level of aggregation. For example, one activity in the CNAE could correspond to two or more activities in the new classification. So, in order to make the decomposition, we consider the variables group of activities (V4809) and its activities (V9907), from 2002 to 2015.[5] Using the PNAD data, this is the longest period for a comparable decomposition analysis.

Regarding the sample delimitation, we consider occupied people with positive income and declaration of schooling. Groups of activity declared as other or undefined activities were dropped. As the group of activities variable considers only people that are 10-year-old or older, we dropped observations less than 10-year-old or no declared. We also exclude people with no declaration of work hours per week or with less than 15 or more than 98 hours (in all jobs), and observations in which position in the occupation was registered as unpaid. Usually, these filters are used in labor economics studies when estimating wage equations. After these delimitations, we get 2,709,162 observations from 1992 to 2015, varying from 106,944 (49,142,420) in 1992 to 128,445 (74,445,191) in 2015. The numbers in parenthesis are the population size represented by the sample. For all estimations in this article, we consider the individual sample weights.

The next section presents the results of evolution and decomposition analysis.

## 4 Results

For the segregation evolution from 1992 to 2015 (indexes' trend presented in subsection 4.1), we consider the seven main groups of activities: 1) Agriculture; 2) Industry; 3) Construction; 4) Trade and repair; 5) Accommodation and transport; 6) Public administration; and 7) Other services. But first, we analyze some preliminary statistics regarding the women participation in the Brazilian labor market. These data for 1992, 2004 and 2015 are in Table 1.

For all years, only in the Other services group the number of women is greater than the number of men. In this group, it reaches almost 4.2 women for each man in 1992, decreasing to 3.6 in 2015, values that are, respectively, the highest and lowest in the period.

In the Industry, Accommodation and transportation, Trade and repair, and Public administration groups, the $X_j/Y_j$ ratio is less than 1 for all years and with increasing trend. This means that, in each of these groups, we observe that the difference between the employed people by gender ($Y_j - X_j$) is decreasing. In these sectors, women's highest participation is observed in the Public Administration, followed by, since 2001, the Trade and repair, Industry, Accommodation and transport, Agriculture, and Construction sectors. Prior to 2001, the participation in the Trade and Repair group was lower than that of the Industry, with the same ordering for the other sectors.

The lowest female participation in the groups for all years is observed for the Construction (e.g. 2.1% in 1992 and 3.1% in 2015), followed by Agriculture (12.3% and 11.9%). In these sectors, the ratio between men and women increased from 47 to 31 men for each woman employed in Construction, and remained constant in approximately 7 men for each woman in Agriculture.

According to the last row of Table 1, the women's participation in the Brazilian labor market increased 18.5% (6.4 *p.p.*), rising from 34.5% (in 1992) to 40.9% (2015). The highest observed value for the historical

[4]We thank Daniela Verzola Vaz (Federal University of São Paulo) and Vandeli dos Santos Guerra (IBGE) for the explanations on the variables harmonization.

[5]After the sample delimitation, we noticed that for 2011 there was an observation whose group of the activity was transportation (V4809=7) and financial intermediation activity (V9907=65000). Since this activity corresponds to the group of other activities (V4809=12), we decided to exclude it from the sample.

Table 1: Proportion of women in the groups ($P_j$) and the ratio $X_j/Y_j$. Brazil, 1992, 2004 and 2015

| Group of activities | 1992 $P_j$ | 1992 $X_j/Y_j$ | 2004 $P_j$ | 2004 $X_j/Y_j$ | 2015 $P_j$ | 2015 $X_j/Y_j$ |
|---|---|---|---|---|---|---|
| Agriculture | 12.3 | 0.140 | 11.1 | 0.125 | 11.9 | 0.135 |
| Industry | 32.05 | 0.547 | 33.9 | 0.515 | 35.2 | 0.544 |
| Construction | 2.1 | 0.021 | 2.2 | 0.022 | 3.1 | 0.032 |
| Trade and repair | 27.2 | 0.374 | 35.5 | 0.550 | 39.5 | 0.652 |
| Accommodation and transport | 22.9 | 0.296 | 26.6 | 0.362 | 31.6 | 0.462 |
| Public administration | 33.0 | 0.492 | 37.0 | 0.588 | 39.4 | 0.649 |
| Other services | 80.7 | 4.179 | 79.4 | 3.858 | 78.1 | 3.561 |
| Brazil | 34.5 | 0.528 | 39.1 | 0.642 | 40.9 | 0.694 |

Source: Estimated by the authors using the PNAD microdata.

Notes: $P_j$ is defined as $X_j/(X_j + Y_j) \times 100\%$. Neither values, $P$ and $X/Y$ for Brazil, are obtained adding up the values of the columns.

series was 41.1% in 2014. The decrease of 0.2 *p.p.* in 2015 compared to 2014 broke off an interrupted increasing trend which has lasted since 1998. In 1997, the women's participation was 35.7%, and had decrease 0.25 *p.p.* compared to 1996. These years (1997 and 2015) are the only 2 out of 21 that showed a decreased compared to the previous year. The steepest increasing trend between two years was observed in 2004 (1.9% or 0.7 *p.p.*).

Table 2 shows the shares of each group in the total of the two genres in for 1992, 2004, and 2015. These are the basic quantities to compute the segregation measures. Sorting the ratios $x_j/y_j$ by non-decreasing order (or equivalently by $X_j/Y_j$ or $P_j$), for 2015, we have the ordering: Construction, Agriculture, Accommodation and transport, Industry, Public administration, Trade and repair, and Other services. This analysis plays a central role in the choice of the general measure's parameter $\varepsilon$, as we will see in the following sections.

Table 2: Share of each group in the total of women ($x_j$) and men ($y_j$) - in %. Brazil, 1992, 2004 and 2015

| Group of activities | 1992 $x_j$ | 1992 $y_j$ | 2004 $x_j$ | 2004 $y_j$ | 2015 $x_j$ | 2015 $y_j$ |
|---|---|---|---|---|---|---|
| Agriculture | 6.9 | 25.8 | 4.0 | 20.5 | 2.8 | 14.1 |
| Industry | 17.9 | 20.1 | 15.5 | 19.3 | 12.8 | 16.3 |
| Construction | 0.5 | 11.9 | 0.4 | 12.5 | 0.8 | 17.6 |
| Trade and repair | 13.5 | 19.1 | 18.3 | 21.3 | 20.5 | 21.8 |
| Accommodation and transport | 6.0 | 10.6 | 6.6 | 11.7 | 9.5 | 14.3 |
| Public administration | 5.9 | 6.3 | 6.0 | 6.5 | 6.2 | 6.6 |
| Other services | 49.3 | 6.2 | 49.3 | 8.2 | 47.4 | 9.2 |
| Total | 100 | 100 | 100 | 100 | 100 | 100 |

Source: Estimated by the authors using the PNAD microdata.

Throughout the period of analysis, approximately half of the women were employed in the Other services group, which includes, for example, Education, Health and Social Services, and Domestic Services activities. For men, the lowest and highest shares in this group are 6.2% and 9.2%, for 1992 and 2015, respectively. If we use the general measure of segregation with lower values for $\varepsilon$ (0 or $-1$, for instance), the measure is more sensitive for the Other services groups, because we observe a high value for $X_j/Y_j$. This means that the Other services group has greater participation in the general measure if compared to others whose ratios $X_j/Y_j$ are smaller.

We emphasize the women's shares in Construction and Agriculture. In 2015, for instance, 1 out of 125 women was employed in the Construction, while that value was 1 out of 40 in the Agriculture. Since the ratios $X_j/Y_j$ are lower to these groups compared to the others, they explain a more significant part of the segregation using the general measure with higher values for $\varepsilon$ (equal or greater than 1). We analyze this issue in more details after presenting the results of the segregation evolution from 1992 to 2015, in the following subsection.

## 4.1 Evolution: 1992-2015

For the evolution analysis of gender segregation in the labor market in Brazil, we consider the 7 main groups of activities defined in the beginning of section 4. This corresponds to consider $J = 7$ in Eqs. (1), (2), and (3). Figure 1 presents these results for the index of dissimilarity, the Gini index, and the general measure considering the Theil's $T$ index ($\varepsilon = 0$), $\varepsilon = 0.5$, and the Theil's $L$ ($\varepsilon = 1$).

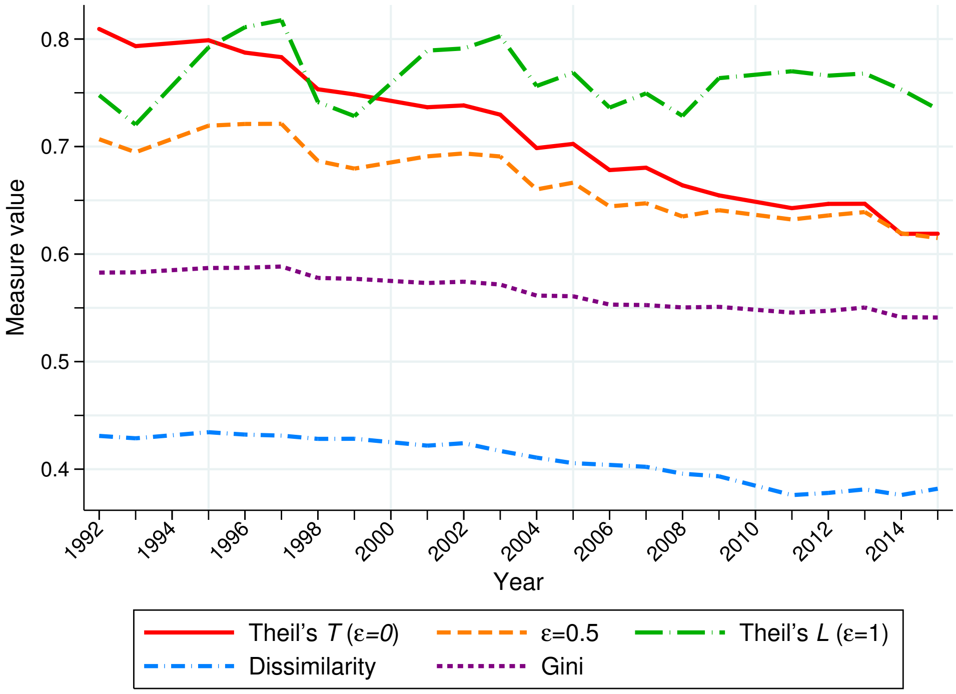


Figure 1: Gender segregation evolution considering 7 groups of activities. Brazil, 1992-2015

Except for the Theil's $L$ index, the other measures showed a downward trend on segregation. At 0.1% of statistical significance, we can affirm that segregation decreased from 1992 to 2015 for these indexes. The Theil's $T$ decreased 23.5%, followed by the index of dissimilarity (11.4%), the general measure with $\varepsilon = 0.5$ (13%), and the Gini index (7.2%). Table 3 present the trend regression analysis of these results. These analyses show that 97.7% of the variability of Theil's $T$ index has been accounted for by the temporal trend, followed by the Gini index (92.1%), the index of dissimilarity (91.8%), the general measure with $\varepsilon = 0.5$ (86.2%), and the Theil's $L$ index (3.6%, but not statistically significant).

For the Gini index, the results confirm the decreasing trend in segregation which was previously pointed out by Salardi (2016) with data up to 2006. Regarding the index of dissimilarity, its value in 1992 was 0.431. It is well known that the index of dissimilarity value is interpreted as the proportion of people in a category that must move from other sectors to achieve the no-segregation level (Duncan and Duncan, 1955a). Thus, in 1992, 43.1% of women should move from groups of activity predominantly female to achieve the no-segregation level. This value was 38.2% in 2015.

In Figure 1 and Table 3, considering the general measure for $\varepsilon = 0$ or 0.5, we conclude that segregation has decrease from 1992 to 2015. However, for the Theil's $L$, there is no trend on segregation (the null hypothesis is no trend and $p$ value=0.431). Further, considering a high value for $\varepsilon$ (1.5 or 2, for instance),

Table 3: Gender segregation evolution considering 7 groups of activities (regressions). Brazil, 1992-2015

| | Theil's $T$ | $\varepsilon = 0.5$ | Theil's $L$ | Dissimilarity | Gini |
|---|---|---|---|---|---|
| Trend | $-0.870^{***}$ | $-0.453^{***}$ | $-0.076$ | $-0.285^{***}$ | $-0.224^{***}$ |
| | (0.030) | (0.043) | (0.094) | (0.023) | (0.016) |
| $R^2$ | 0.9768 | 0.8621 | 0.0361 | 0.9178 | 0.9206 |

Source: Estimated by the authors using the PNAD microdata.
Notes: Robust standard errors in parentheses; All constants omitted; *** represent statistical significance at 0.1%; Trend correspond to year divided by 100; Number of observations: 21.

we can verify that there is an increasing trend on segregation. This phenomenon is associated with the interpretation of $\varepsilon$ and was studied by Botassio and Hoffmann (2019). The authors conclude that $\varepsilon$ is an aversion parameter to segregation considering the distribution of the ratios $x_j/y_j$. Considering higher values for $\varepsilon$, the general measure is more sensitive to changes in sectors with lower values for $x_j/y_j$ and vice versa. Thus, according to Table 1, using $\varepsilon$ less than 0, the general measure is more sensitive to the Other services sector, while for $\varepsilon$ greater than 1, the general measure is especially sensitive to the Agriculture and Construction sectors. To capture this variation in sensitivity, without giving too much weight to extreme cases, it seems reasonable to use $\varepsilon$ values from 0 to 1.

As the last result of this subsection, we present the correlation analysis between the Theil's $T$, the general measure for $\varepsilon = 0.5$, the Theil's $L$, and the dissimilarity and Gini indexes in Table 4. All indexes showed a positive correlation with the others. Excluding the Theil's $L$ index, all pairwise comparison presented strong and significant correlation at 0.1%, with values higher than 0.938.

Table 4: Correlation between the segregation measures. Brazil, 1992-2015

| | Theil's $T$ | $\varepsilon = 0.5$ | Theil's $L$ | Dissimilarity | Gini |
|---|---|---|---|---|---|
| Theil's $T$ | 1.000 | | | | |
| $\varepsilon = 0.5$ | $0.968^{***}$ | 1.000 | | | |
| Theil's $L$ | 0.299 | $0.521^{*}$ | 1.000 | | |
| Dissimilarity | $0.961^{***}$ | $0.938^{***}$ | 0.265 | 1.000 | |
| Gini | $0.985^{***}$ | $0.985^{***}$ | $0.389^{+}$ | $0.963^{***}$ | 1.000 |

Notes: +, *, and *** represent statistical significance at 10%, 5% and 0.1%. Number of observations: 21.

## 4.2 Decomposition: 2002-2015

For the decomposition, we consider 53 activities divided into 6 groups of activities. These groups are: 1) Agriculture; 2) Industry and Construction; 3) Trade and repair; 4) Accommodation and transport; 5) Public administration; and 6) Other services. The difference between this classification and that used in the previous section is the aggregation of the Industry and Construction groups.

According to the CNAE classification, the Construction group consists of the activities Construction (V9907=45999) and Construction Equipment Rentals (V9907=45005). In order to measure segregation within the Construction group, there are two essential problems. First, the proportions of people in these two activities are extremely disproportionate. See the 2015 data, for example. After the sample delimitation, there were 13,999 observations (13,574 men) for the Construction activity, whereas there were only 11 observations (9 men) for the Equipment Rental activity. Second, the difference in the proportions of women and men is extremely high. This implies that considering the category $X$ as women, segregation within this group would be extremely high using the general measure with $\varepsilon \geq 1$, that could even tend to infinity if there are no female observations in some activity. For 2002-2005, 2007-2009, and 2014, there was no female observation for the Equipment Rental activity. Thus, considering the Theil's $L$ index ($\varepsilon = 1$), for these years, the general measures tend to infinity. In addition, for 2006 in this same activity, there was also no observation for men. Thus, to circumvent this issue, the Construction and Equipment Rental activities were aggregated. As a result, the Construction group would be composed of only one activity. So, we aggregated the Industry and Construction groups.

Before we present the decomposition for segregation from 2002 to 2015, we illustrate the concept of segregation curve for 2002 and 2015, considering the 6 groups and 53 activities separately.

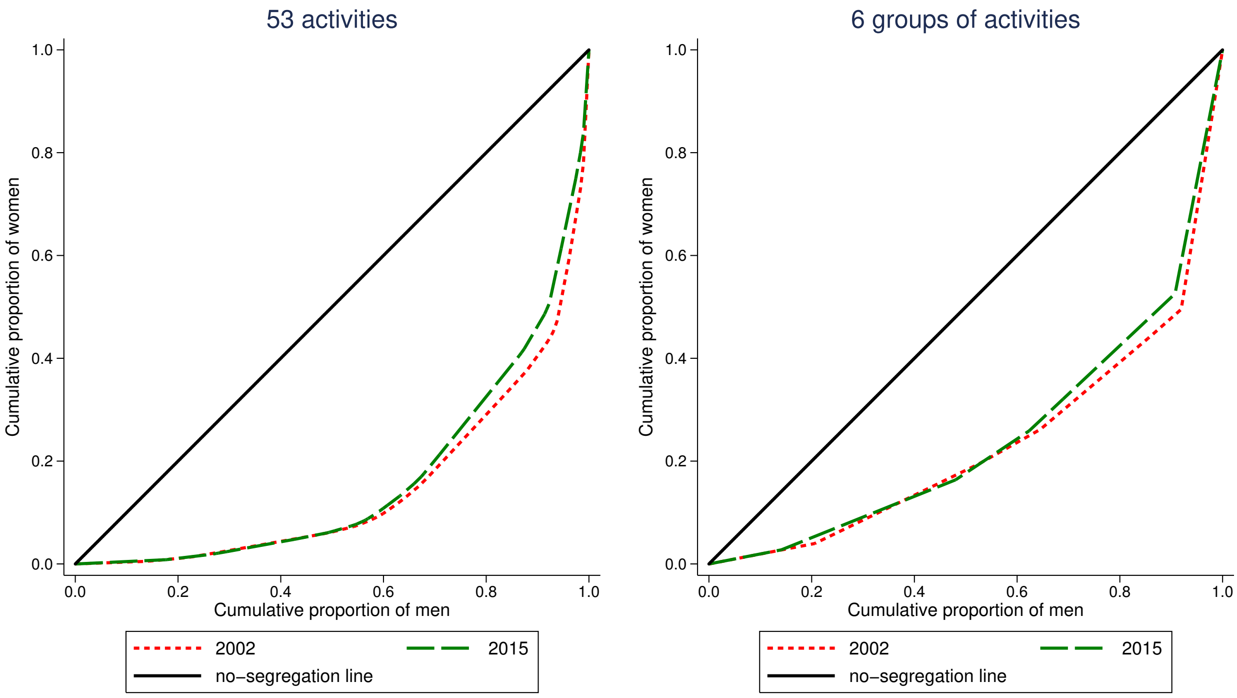


Figure 2: Segregation curves considering 53 activities and 6 groups. Brazil, 2002 and 2015

According to the criterion of first-order stochastic dominance, if we represent by $SC(\mathbf{D}_{i,J})$ the segregation curve for the $i$th year considering $J$ activities, the curve for year $i$ dominates the curve for year $k$ $[SC(\mathbf{D}_{k,J})]$ if it lies at no point below and at some point above the $i$'s curve. So, using this criterion to rank two different distributions with the same number of sectors, we can assert that distribution $\mathbf{D}_{i,J}$ is preferable to $\mathbf{D}_{k,J}$. In other words, the distribution associated with the dominant curve ($i$) is ranked as less segregated than the other one ($k$). The concept of stochastic dominance for segregation curves was adapted by Hutchens (1991) from the Lorenz curve in the context of inequality.

However, for both graphs in Figure 2, considering 53 activities or 6 groups of activities (which correspond to the between-groups in the decomposition analysis), neither of the segregation curves for 2002 or 2015 dominates another distribution. This occurs because the curves intersect (a subtle difference in the lower part for both graphs). In this case, the criterion of stochastic dominance for segregation curves provides an incomplete ranking of distributions. So, an alternative way is to use segregation measures, which was made in subsection 4.1 for the 7 groups of activities, from 1992 to 2015. Now we analyze the segregation decomposition.

For the decomposition analysis, we use the general measure of segregation, because it is additively decomposable. This requirement is not fulfilled by the index of dissimilarity and the Gini index for segregation. We consider 53 activities, and the 6 groups delimited at the beginning of this subsection (the same groups and activities used in the segregation curves). So, we consider $J = 53$ and $G = 6$ in (4). Figure 3 presents the segregation evolution results considering the overall segregation (53 activities) using the general measure for the Theil's $T$ index, $\varepsilon = 0.1$, 0.5, and 0.9, and the Theil's $L$ index. We use $\varepsilon = 0.1$ and

0.9 to comparison and analyze the behavior of the general measure for $\varepsilon$ close to 0 and 1 because for those former values the between and within groups components are independent.

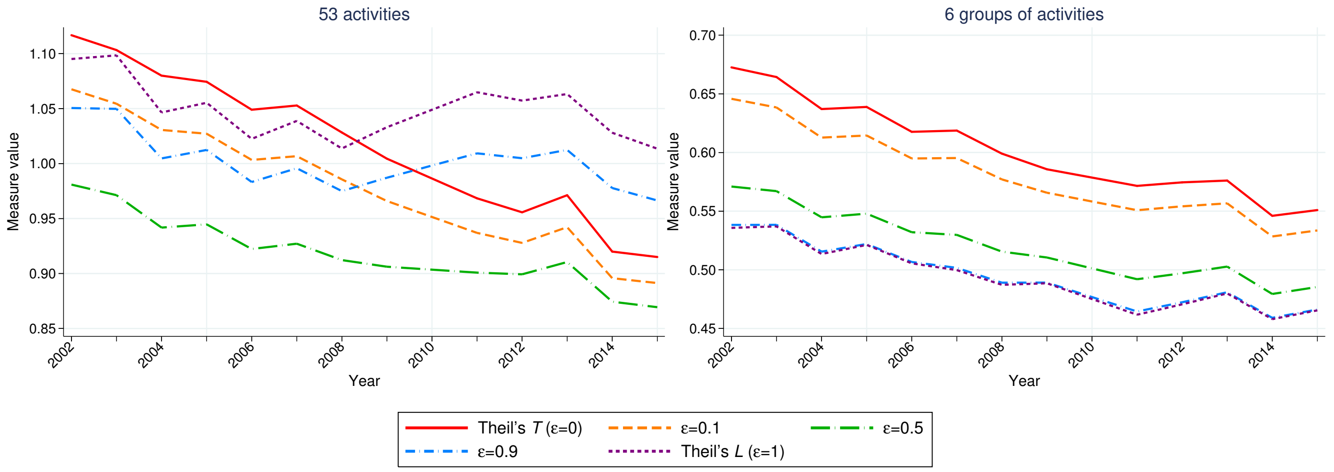


Figure 3: Gender segregation evolution considering the overall and between-groups segregation. Brazil, 2002-2015

We can assert that, using any value for $\varepsilon$, the overall segregation decreased, but at different levels of significance (at 0.1% for $\varepsilon = 0$, 0.1, and 0.5, at 5% using $\varepsilon = 0.9$, and at 10% for $\varepsilon = 1$). These regressions results are in the upper part of the Table 5. We observe an inverse relationship between the falling intensity of segregation and the $\varepsilon$ value. According to these measures, the overall segregation decreased 18.1% ($\varepsilon$=0), 16.5% (0.1), 11.4% (0.5), 8% (0.9), and 7.4% (1). As already mentioned, this variation in the strength of the trend is related to the effect of $\varepsilon$ on the sensitivity of the segregation measure to changes in different sectors, depending upon their $X_j/Y_j$ ratios.

Table 5: Gender segregation evolution considering the overall and between-groups segregation (regressions). Brazil, 2002-2015

| | $\varepsilon$ | | | | |
|---|---|---|---|---|---|
| | 0 | 0.1 | 0.5 | 0.9 | 1 |
| Overall (53 activities) | | | | | |
| Trend | $-1.550^{***}$ | $-1.324^{***}$ | $-0.722^{***}$ | $-0.379^{*}$ | $-0.313^{+}$ |
| | (0.068) | (0.065) | (0.084) | (0.137) | (0.160) |
| $R^2$ | 0.9790 | 0.9757 | 0.8849 | 0.3981 | 0.2413 |
| Between-groups (6 groups of activities) | | | | | |
| Trend | $-0.926^{***}$ | $-0.859^{***}$ | $-0.676^{***}$ | $-0.599^{***}$ | $-0.595^{***}$ |
| | (0.064) | (0.062) | (0.057) | (0.060) | (0.062) |
| $R^2$ | 0.9517 | 0.9497 | 0.9350 | 0.9067 | 0.8972 |

Source: Estimated by the authors using the PNAD microdata.
Notes: Robust standard errors in parentheses; All constants omitted; $^{+}$, $^{*}$, and $^{***}$ represent statistical significance at 10%, 5%, and 0.1%, respectively; Trend correspond to year divided by 100; Number of observations: 13.

For the between-group component, for all $\varepsilon$ values there is a decreasing trend on segregation at 0.1% of statistical significance (bottom part of Table 5). The percentage decreased varies from 18.1% for $\varepsilon = 0$, to 13.1% for $\varepsilon = 1$, from 2002 to 2015. It is also interesting to compare the results from 2002 to 2015 obtained considering 7 sectors (Figure 1) and those obtained for the segregation between 6 groups (Figure 3). Let us consider the evolution of Theil's $L$ index ($\varepsilon = 1$) The regression analysis using 7 groups (Figure 1), considering only the 13 observations from 2002 to 2015, does not indicate a decreasing trend on the overall segregation at 10% ($p$-value=0.111). This fact is associated with the aggregation of the Industry and the Construction groups. Consider the 2015 data for illustration. In this year, the ratio $x_j/y_j$ was 0.047 for Construction and 0.784 for Industry. With the aggregation of the groups, we eliminate a very low value for $x_j/y_j$ (the aggregated ratio is 0.401). In consequence, the general measure becomes less sensitive to higher $\varepsilon$ values. The comparison of these results alerts us to two important facts that we must be aware of: the parameter's choice and the delimitation of the sectors of activity. Differences in these points can modify substantially the results and, consequently, the conclusions obtained. To use only the dissimilarity and Gini measures avoids the problem, but limits the analysis, also because these indexes are not additively decomposable.

An important feature of the decomposition is the share analysis of each term (between and within) to the overall segregation. Figure 4 and Table 6 present the evolution of the between-groups share in the overall segregation. In mathematical terms, if we represent the between-groups term by $B_\varepsilon(\mathbf{D})$ (second term in the right-hand side of Eq. (4)), this share is $S_\varepsilon(\mathbf{D}) = B_\varepsilon(\mathbf{D})/I_\varepsilon(\mathbf{D})$, or in percentage,

$$S_\varepsilon(\mathbf{D}) = \frac{1 - \sum_{g=1}^{G} y_g^\varepsilon x_g^{1-\varepsilon}}{1 - \sum_{g=1}^{G} \sum_{j=1}^{n_g} y_{gj}^\varepsilon x_{gj}^{1-\varepsilon}} \times 100\% \tag{5}$$

where $x_{gj}$ is the share of the $j$th stratum, that belongs to the $g$th group, in the total of the $X$ category, with similar definition to $y_{gj}$. The cases for $\varepsilon = 0$ or 1 are obtained by L'Hôpital's rule.

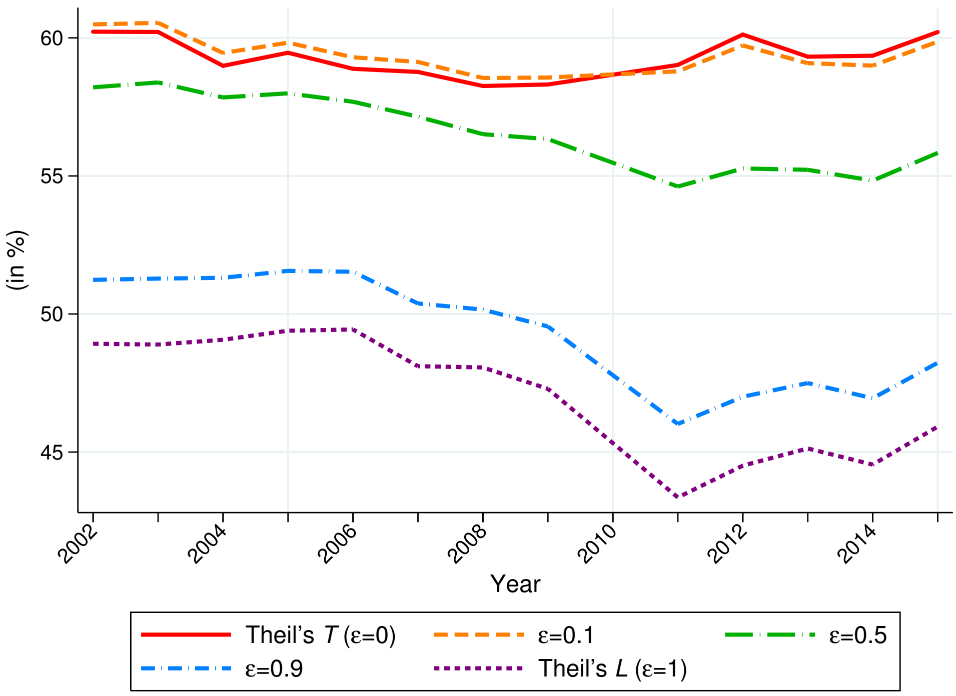


Figure 4: Share of the between-groups term in the overall segregation. Brazil, 2002-2015

We highlight that considering $\varepsilon = 0$ or 0.1, there are no statistical evidence that the share of the between-groups component in the overall segregation decreased or increased. However, as the overall segregation decreased using these values for $\varepsilon$, we can assert that the downward trend in the overall segregation is attributed to both components, the within-groups, and the between-groups.

The share of the between-groups component in the overall segregation decreased statistically at 0.1% using $\varepsilon = 0.5$ (4.2% or 2.4 *p.p.*), $\varepsilon = 0.9$ (6.2% or 3 *p.p.*), and $\varepsilon = 1$ (6.5% or 3 *p.p.*). It can be verified that for these values of $\varepsilon$, the downward trend of the overall segregation is accounted for by the between-group component decreasing trend. Use the general measure for $\varepsilon = 0.5$, for instance, which has the highest proportion of the share's variability explained by the regression (84.5%). In 2002, the between-

Table 6: Share of the between-groups term in the overall segregation (regressions). Brazil, 2002-2015

| | $\varepsilon$ | | | | |
|---|---|---|---|---|---|
| | 0 | 0.1 | 0.5 | 0.9 | 1 |
| Trend | 0.081 | $-6.860$ | $-28.826^{***}$ | $-40.856^{***}$ | $-42.343^{***}$ |
| | (4.960) | (4.276) | (4.193) | (6.722) | (7.420) |
| $R^2$ | 0.0000 | 0.2056 | 0.8451 | 0.7592 | 0.7254 |

Source: Estimated by the authors using the PNAD microdata.
Notes: Robust standard errors in parentheses; All constants omitted;
*** represent statistical significance at 0.1%, Trend correspond to year divided by 100; Number of observations: 13.

groups' share on the overall segregation was 58.2%, falling down to 55.8% in 2015. Note that for any measure, regardless of the $\varepsilon$'s choice, we can assign the decreasing trend in the overall segregation, in parts or completely, to the between-groups component.

## 5 Conclusions

This study is conducted to explore the evolution of gender segregation in the Brazilian labor market. The results show that considering 7 groups of activities, segregation decreased from 1992 to 2015 using several measures. They are: Theil's $T$ index (decreased 23.5%), followed by the general measure with $\varepsilon = 0.5$ (13%), the index of dissimilarity (11.4%), and the Gini index (7.2%).

In the decomposition analysis, considering 53 activities divided into 6 groups, both the overall and the between-groups segregation decreased statistically (data from 2002 to 2015). For overall segregation, this decreasing ranged from 18.1% for the Theil's $T$ index to 7.4% for Theil's $L$ index. For the between-groups component, this falling varied, respectively, 18.1% and 13.1%. In addition, the decreasing trend of the between-groups' share in the overall segregation explains partially the overall segregation decline considering Theil's $T$ index, while it fully explains this decline using the Theil's $L$ index.

These results show that gender segregation by activities, undoubtedly, has decreased in Brazil considering both approaches.

## Acknowledgments

Both authors thank the Brazilian National Council for Technological and Scientific Development (CNPq - *Conselho Nacional de Desenvolvimento Científico e Tecnológico*) for scholarship and research grant.

## References

Albelda, R.P.. Occupational segregation by race and gender, 1958-1981. Industrial and Labor Relations Review 1986;39(3):404–411. doi:10.2307/2524099.

Anker, R.. Theories of occupational segregation by sex: an overview. International Labour Review 1997;136(3):315–339. URL: http://heinonline.org/HOL/LandingPage?handle=hein.journals/intlr136&div=31&id=&page=.

Botassio, D.C.. Segregation and inequality: analogy in the measurement and analysis of gender segregation by activity sectors in Brazil [in Portuguese]. Master's thesis in Applied Economics; University of São Paulo; 2017. doi:10.11606/D.11.2017.tde-16082017-090239.

Botassio, D.C., Hoffmann, R.. A note on the class of additively decomposable segregation measures; 2018.

Botassio, D.C., Hoffmann, R.. Measuring gender segregation: Sensitivity to movements and to an increase in one category. Economic Theory Bulletin 2019;doi:10.1007/s40505-019-00162-8.

Duncan, O.D., Duncan, B.. A methodological analysis of segregation indexes. American Sociological Review 1955a;20(2):210–217. doi:10.2307/2088328.

Duncan, O.D., Duncan, B.. Residential distribution and occupational stratification. American Journal of Sociology 1955b;60(5):493–503. doi:10.1086/221609.

Foguel, M.N.. The effect of gender segregation at the occupation, industry, establishment, and job-cell levels on the male-female wage gap. 2016. URL: http://reap.org.br/wp-content/uploads/2016/03/091-The-Effects-of-Gender-Segregation_Miguel-Nathan.pdf.

Frankel, D.M., Volij, O.. Measuring school segregation. Journal of Economic Theory 2011;146(1):1–38. doi:10.1016/j.jet.2010.10.008.

Gross, E.. Plus ca change...? The sexual structure of occupations over time. Social Problems 1968;16(2):198–208. doi:10.2307/800005.

Hoffmann, R.. Distribuição de renda: medidas de desigualdade e pobreza. São Paulo: EDUSP, 1998.

Hutchens, R.. Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations. Mathematical Social Sciences 1991;21(1):31–51. doi:10.1016/0165-4896(91)90038-S.

Hutchens, R.. Numerical measures of segregation: desirable properties and their implications. Mathematical Social Sciences 2001;42(1):13–29. doi:10.1016/S0165-4896(00)00070-6.

Hutchens, R.. One measure of segregation. International Economic Review 2004;45(2):555–578. doi:10.1111/j.1468-2354.2004.00136.x.

IBGE, . Pesquisa Nacional por Amostra de Domicílios. 2017. URL: ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_anual.

Jahn, J., Schmid, C.F., Schrag, C.. The measurement of ecological segregation. American Sociological Review 1947;12(3):293–303. doi:10.2307/2086519.

James, D.R., Taeuber, K.E.. Measures of segregation. Sociological Methodology 1985;15:1–32. doi:10.2307/270845.

Karmel, T., MacLachlan, M.. Occupational sex segregation - increasing or decreasing? Economic Record 1988;64(3):187–195. doi:10.1111/j.1475-4932.1988.tb02057.x.

King, M.C.. Occupational segregation by race and sex in Brazil, 1989-2001. The Review of Black Political Economy 2009;36(2):113–125. doi:10.1007/s12114-009-9038-2.

Mora, R., Ruiz-Castillo, J.. Additively decomposable segregation indexes. The case of gender segregation by occupations and human capital levels in Spain. The Journal of Economic Inequality 2003;1(2):147–179. doi:10.1023/A:1026198429377.

Oliveira, A.M.H.. Occupational gender segregation and effect on wages in Brazil. In: XXIV General Population Conference. 2001. URL: http://archive.iussp.org/Brazil2001/s30/S38_03_Oliveira.pdf.

Ometto, A.M.H., Hoffmann, R., Alves, M.C.. A segregação por gênero no mercado de trabalho nos estados de São Paulo e Pernambuco. Economia Aplicada 1997;1(3):393–423.

Oppenheimer, V.K.. The female fabor force in the United States: demographic and economic factors governing its growth and changing composition. Berkeley: University of California Press, 1970.

Ribeiro, R., Araújo, G.S.. Segregação ocupacional no mercado de trabalho segundo cor e nível de escolaridade no Brasil contemporâneo. Nova Economia 2016;26(1):147–177. doi:10.1590/0103-6351/2652.

Salardi, P.. An analysis of pay and occupational differences by gender and race in Brazil - 1987 to 2006. Ph.D. thesis; 2013. URL: http://sro.sussex.ac.uk/45204/.

Salardi, P.. The evolution of gender and racial occupational segregation across formal and non-formal labor markets in Brazil, 1987 to 2006. Review of Income and Wealth 2016;62(S1):S68–S89. doi:10.1111/roiw.12159.

Salas, C., Leite, M.. Segregación sectorial por género: una comparación Brasil-México. Cadernos PROLAM/USP 2007;6(11):241–259. doi:10.11606/issn.1676-6288.prolam.2007.82307.

Schelling, T.C.. Dynamic models of segregation. The Journal of Mathematical Sociology 1971;1(2):143–186. doi:10.1080/0022250x.1971.9989794.

Taeuber, K.E., Taeuber, A.F.. Negroes in cities: residential segregation and neighborhood change. Chicago: Aldine Publishing, 1965.

Theil, H., Finizza, A.J.. A note on the measurement of racial integration of schools by means of informational concepts. The Journal of Mathematical Sociology 1971;1(2):187–193. doi:10.1080/0022250X.1971.9989795.

Vaz, D.V., Hoffmann, R.. Segregação ocupacional por sexo no setor público brasileiro no período 1995 e 2008. Revista ABET 2011;10(1):120–141. URL: http://periodicos.ufpb.br/ojs2/index.php/abet/article/view/15546.

Walker, D., Stinchcombe, A.L., McDill, M.S.. School desegregation in Baltimore. Technical Report; The Johns Hopkins University; Baltimore, [Report, The Johns Hopkins University, ED013168], 52 p.; 1967. URL: http://files.eric.ed.gov/fulltext/ED013168.pdf.

# Supplementary material

Table 7: Gender segregation evolution considering 7 groups of activities. Brazil, 1992-2015

| Year | Theil's $T$ index | $\varepsilon = 0.5$ | Theil's $L$ index | Dissimilarity | Gini |
|---|---|---|---|---|---|
| 1992 | 0.8094 | 0.7069 | 0.7478 | 0.4310 | 0.5828 |
| 1993 | 0.7934 | 0.6947 | 0.7204 | 0.4287 | 0.5830 |
| 1995 | 0.7989 | 0.7194 | 0.7922 | 0.4344 | 0.5871 |
| 1996 | 0.7874 | 0.7210 | 0.8110 | 0.4322 | 0.5873 |
| 1997 | 0.7831 | 0.7212 | 0.8175 | 0.4312 | 0.5885 |
| 1998 | 0.7533 | 0.6866 | 0.7414 | 0.4282 | 0.5778 |
| 1999 | 0.7485 | 0.6795 | 0.7283 | 0.4283 | 0.5770 |
| 2001 | 0.7366 | 0.6909 | 0.7892 | 0.4219 | 0.5731 |
| 2002 | 0.7383 | 0.6936 | 0.7913 | 0.4242 | 0.5744 |
| 2003 | 0.7297 | 0.6908 | 0.8027 | 0.4170 | 0.5717 |
| 2004 | 0.6986 | 0.6602 | 0.7563 | 0.4107 | 0.5614 |
| 2005 | 0.7025 | 0.6664 | 0.7687 | 0.4056 | 0.5608 |
| 2006 | 0.6782 | 0.6443 | 0.7362 | 0.4040 | 0.5531 |
| 2007 | 0.6804 | 0.6473 | 0.7495 | 0.4022 | 0.5526 |
| 2008 | 0.6640 | 0.6350 | 0.7285 | 0.3958 | 0.5504 |
| 2009 | 0.6546 | 0.6408 | 0.7636 | 0.3934 | 0.5509 |
| 2011 | 0.6427 | 0.6322 | 0.7700 | 0.3759 | 0.5456 |
| 2012 | 0.6467 | 0.6360 | 0.7659 | 0.3779 | 0.5472 |
| 2013 | 0.6468 | 0.6391 | 0.7680 | 0.3813 | 0.5503 |
| 2014 | 0.6189 | 0.6194 | 0.7531 | 0.3760 | 0.5413 |
| 2015 | 0.6190 | 0.6150 | 0.7348 | 0.3818 | 0.5410 |

Note: Data used to make Figure 1.

Table 8: Classification of the groups of activities (decomposition). Brazil, 2002-2015

| Group | Code V9907 | | | Activity |
|---|---|---|---|---|
| | 1101 | to | 1500 | Crop and Animal Production |
| Agriculture | 2001 | and | 2002 | Silviculture, Forestry and Logging |
| | 5001 | and | 5002 | Fishing and Aquaculture |
| | 15010 | to | 15055 | Food and Beverage Manufacturing |
| | | 16000 | | Tobacco Product Manufacturing |
| | 17001 | and | 17002 | Textile Products Manufacturing |
| | 18001 | and | 18002 | Apparel Manufacturing |
| | 19011 | to | 19020 | Leather Products Manufacturing |
| | | 20000 | | Wood Products Manufacturing |
| | 21001 | and | 21002 | Pulp, Paper and Paper Products Manufacturing |
| | | 22000 | | Publishing, Printing and Reproduction of Recordings |
| | 23010 | to | 23400 | Coal, Petroleum, and Alcohol Products Manufacturing |
| | 24010 | to | 24090 | Chemical Products Manufacturing |
| | 25010 | and | 25020 | Rubber and Plastics Product Manufacturing |
| Industry | 26010 | to | 26092 | Nonmetallic Mineral Product Manufacturing |
| | 27001 | to | 27003 | Metals Production and Basic Processing |
| | 28001 | and | 28002 | Metal Product Manufacturing (excluding machinery) |
| | 29001 | and | 29002 | Machinery and Equipment Manufacturing |
| | | 30000 | | Office and Data Processing Equipment Manufacturing |
| | 31001 | and | 31002 | Electrical Machinery and Supplies Manufacturing |
| | | 32000 | | Electronic and Communication Equipment Manufacturing |
| | 33001 | to | 33005 | General Instrumentation |
| + | 34001 | to | 34003 | Vehicle Assembly, Engine, and Manufacturing |
| | 35010 | to | 35090 | Other Transportation Equipment Manufacturing |
| | 36010 | and | 36090 | Furniture and Miscellaneous Manufacturing |
| | | 37000 | | Recycling |
| | 10000 | and | 11000 | Coal, Petroleum and Gas Extraction |
| | 12000 | to | 13002 | Metallic Mineral Mining |
| | 14001 | to | 14004 | Nonmetallic Mineral Mining |
| | 40010 | and | 40020 | Electricity, Gas and Hot Water |
| | | 41000 | | Water Collection, Treatment and Distribution |
| Construction | 45005 | and | 45999 | Construction |
| | | 55010 | | Accommodation |
| | 55020 | and | 55030 | Foodservice |
| Accommodation | 60010 | to | 60092 | Land transportation |
| and | | 61000 | | Water Transportation |
| Transportation | | 62000 | | Air Transport |
| | 63010 | to | 63030 | Storage and Transportation Auxiliary Activities |
| | 64010 | and | 64020 | Mail Services and Telecommunications |
| | 75011 | and | 75015 | Federal Public Administration |
| | 75012 | and | 75016 | State Public Administration |
| Public Administration | 75013 | and | 75017 | County Public Administrator |
| | | 75014 | | Defense |
| | | 75020 | | Social Security |
| Trade and | 50010 | to | 50050 | Vehicles and Fuels |
| Repair | 53010 | to | 53113 | Wholesale and Retail |
| | 80011 | to | 80090 | Education |
| | 85011 | to | 85030 | Health and Social Services |
| | | 90000 | | Urban Cleaning and Sewage |
| Other Services | 91010 | to | 91092 | Associative Activities |
| | 92011 | to | 92040 | Recreational Activities |
| | 93010 | to | 93092 | Personal Services |
| | | 95000 | | Domestic Services |

Source: Made by the authors using the PNAD variable V9907.

Table 9: Gender segregation evolution considering the overall and between-groups segregation. Brazil, 2002-2015

| | $\varepsilon$ | | | | |
|---|---|---|---|---|---|
| Year | 0 | 0.1 | 0.5 | 0.9 | 1 |
| Overall (53 activities) | | | | | |
| 2002 | 1.1168 | 1.0676 | 0.9809 | 1.0506 | 1.0951 |
| 2003 | 1.1033 | 1.0545 | 0.9713 | 1.0498 | 1.0984 |
| 2004 | 1.0800 | 1.0306 | 0.9418 | 1.0047 | 1.0465 |
| 2005 | 1.0744 | 1.0272 | 0.9448 | 1.0124 | 1.0553 |
| 2006 | 1.0490 | 1.0033 | 0.9223 | 0.9832 | 1.0225 |
| 2007 | 1.0528 | 1.0068 | 0.9271 | 0.9958 | 1.0388 |
| 2008 | 1.0283 | 0.9857 | 0.9122 | 0.9750 | 1.0137 |
| 2009 | 1.0045 | 0.9660 | 0.9062 | 0.9871 | 1.0331 |
| 2011 | 0.9684 | 0.9370 | 0.9008 | 1.0093 | 1.0649 |
| 2012 | 0.9557 | 0.9278 | 0.8993 | 1.0049 | 1.0573 |
| 2013 | 0.9712 | 0.9422 | 0.9104 | 1.0125 | 1.0633 |
| 2014 | 0.9200 | 0.8957 | 0.8744 | 0.9778 | 1.0281 |
| 2015 | 0.9150 | 0.8913 | 0.8694 | 0.9665 | 1.0136 |
| Between-groups (6 groups of activities) | | | | | |
| 2002 | 0.6726 | 0.6458 | 0.5710 | 0.5383 | 0.5358 |
| 2003 | 0.6643 | 0.6384 | 0.5671 | 0.5384 | 0.5371 |
| 2004 | 0.6371 | 0.6127 | 0.5448 | 0.5155 | 0.5135 |
| 2005 | 0.6389 | 0.6145 | 0.5479 | 0.5220 | 0.5213 |
| 2006 | 0.6177 | 0.5949 | 0.5321 | 0.5067 | 0.5055 |
| 2007 | 0.6187 | 0.5953 | 0.5298 | 0.5017 | 0.4998 |
| 2008 | 0.5991 | 0.5771 | 0.5155 | 0.4890 | 0.4872 |
| 2009 | 0.5857 | 0.5657 | 0.5105 | 0.4891 | 0.4886 |
| 2011 | 0.5716 | 0.5508 | 0.4920 | 0.4645 | 0.4617 |
| 2012 | 0.5746 | 0.5541 | 0.4970 | 0.4724 | 0.4706 |
| 2013 | 0.5761 | 0.5567 | 0.5028 | 0.4809 | 0.4799 |
| 2014 | 0.5460 | 0.5285 | 0.4794 | 0.4591 | 0.4580 |
| 2015 | 0.5510 | 0.5336 | 0.4854 | 0.4662 | 0.4654 |

Note: Data used to make Figure 3.

Table 10: Between-groups share in the overall segregation (in %). Brazil, 2002-2015

| | $\varepsilon$ | | | | |
|---|---|---|---|---|---|
| Year | 0 | 0.1 | 0.5 | 0.9 | 1 |
| 2002 | 66.11 | 67.41 | 70.71 | 72.03 | 72.25 |
| 2003 | 66.14 | 67.52 | 71.12 | 72.76 | 73.08 |
| 2004 | 64.68 | 66.17 | 70.10 | 71.93 | 72.27 |
| 2005 | 65.38 | 66.79 | 70.54 | 72.45 | 72.84 |
| 2006 | 64.65 | 66.07 | 69.86 | 71.66 | 72.00 |
| 2007 | 64.62 | 66.04 | 69.81 | 71.74 | 72.15 |
| 2008 | 64.57 | 65.94 | 69.61 | 71.49 | 71.87 |
| 2009 | 65.17 | 66.60 | 70.72 | 73.31 | 73.91 |
| 2011 | 66.37 | 67.42 | 70.18 | 71.86 | 72.31 |
| 2012 | 67.67 | 68.53 | 70.72 | 72.06 | 72.44 |
| 2013 | 66.60 | 67.58 | 70.20 | 71.81 | 72.22 |
| 2014 | 67.27 | 68.20 | 70.84 | 72.74 | 73.25 |
| 2015 | 67.64 | 68.49 | 70.74 | 72.13 | 72.49 |

Note: Data used to make Figure 4.