



WIDER Working Paper 2018/156

## **The gender gap, education, and the life cycle profile in the Brazilian formal labour market**

Cecilia Machado,<sup>1</sup> Marcelo Neri,<sup>2</sup> and Valdemar Pinho Neto<sup>1</sup>

December 2018

**Abstract:** We study the trajectory of the gender gap over time and over the life cycle, using a matched employer-employee data from the formal labour market in Brazil. We document the evolution of participation and earnings for both males and females during the period 1994–2015 and the gender earnings gap throughout the life cycle for different birth cohorts. We focus on the cohort of workers (male and female) born between 1967 and 1974, who were working in 1994, in order to understand the roles that occupation/industry and establishment play on the gender gap pattern throughout the life cycle and for different education levels. We find that the gender earnings gap increases with the educational level. For instance, at 40 years of age, women without high school degrees earned on average 28.8 per cent less than men with the same level of education and for the group with high school and college degrees, this difference was 32.6 and 47.4 per cent, respectively. After controlling for the occupation/industry and firm’s characteristics, we observe a remaining gender earnings gap lower than 20 per cent and greater than 10 per cent over the entire career and independent of the educational level.

**Keywords:** earnings inequality, linked employer-employee data, education, Brazilian inequality, entropy indexes, firms’ effects

**JEL classification:** I24, J31

**Acknowledgements:** This paper was fully supported by FGV Social and *Rede de Pesquisa e Conhecimento Aplicado* (RPCAP) from Getulio Vargas Foundation (FGV). We thank comments given at conferences held in Helsinki and Natal. The paper will be presented at the National Meeting of Brazilian Economists (ANPEC) in Rio de Janeiro, 2018.

---

<sup>1</sup> Getulio Vargas Foundation (FGV) – *Escola de Pós-Graduação em Economia* (FGV EPGE), Rio de Janeiro, Brazil; <sup>2</sup> FGV Social and FGV EPGE, Rio de Janeiro, Brazil, corresponding author: [Marcelo.Neri@fgv.br](mailto:Marcelo.Neri@fgv.br).

This study has been prepared within the UNU-WIDER project on ‘Inequality in the Giants’.

Copyright © The Authors 2018

Information and requests: [publications@wider.unu.edu](mailto:publications@wider.unu.edu)

ISSN 1798-7237 ISBN 978-92-9256-598-5 <https://doi.org/10.35188/UNU-WIDER/2018/598-5>

Typescript prepared by Ans Vehmaanperä.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

## 1 Introduction

Female labour force participation has experienced a remarkable progress over the past decades, resulting in a gender convergence regarding human capital and labour market outcomes. Therefore, the long-term trend has been a substantial reduction in the gender pay gap in most economically advanced countries (Blau and Khan 2008). Nevertheless, a gender pay differential remains in all countries and, although it has been intensively investigated for many decades, this topic is still an area of active and innovative research. The gender gap dynamics depend on many factors, such as age, time since leaving school, marital status, having small children, educational level, worked hours, work flexibility, etc. (see, for example, Bertrand et al. 2010). However, a well-known finding in the literature is that the gender pay gap persists even when controlled by a broad set of firms' and individuals' observable characteristics.

There is a vast international literature on the patterns of gender inequality over time. For instance, using information from the Michigan Panel Study of Income Dynamics (PSID) and the March Current Population Survey (CPS) microdata over 1980–2010, Blau and Khan (2016) provide new empirical evidence on the extent of gender wage gap and its trend, which declined considerably over this period. On the other hand, using American data from the Longitudinal Employer-Household Dynamics (LEHD) database linked to the 2000 Census, Goldin et al. (2017) studied the patterns of wage gender gap and found that the earnings gap between men and women has been expanding over the life cycle. The authors document that movements within and across establishments contribute to the expanding gender earnings gap until 14 years after school-leaving, for narrowly defined cohort and education groups.

In this paper, we study the gender gap over time and the life cycle in the formal labour market in Brazil, a large developing country that has experienced a strong process of formalization in the labour force in recent years, with massive entry of women into the workforce. Despite recent advances, gender inequality remains very high in Brazil<sup>1</sup>. We start with a comprehensive picture of the formal labour market in Brazil since 1994 until 2015, highlighting the evolution of gender differences in employment, earnings and worked hours per week. We also analyse the patterns of gender earnings gap throughout the life cycle, which may be influenced by the combination of two different processes. First, the career path varies within an employer, due to wage raises and promotions over time and, second, there is a sorting into high- versus low-paying employers. Additionally, we investigate the role played by the substantial employment differences in both industrial and occupational structure by gender. In sum, we explore how much of the change in the gender gap over the life cycle occurs within or between establishments and how much of it is driven by the firm's sector of activity or occupational choices.

We find that the gender gap expands over the life cycle until around the age of 40, when it starts to reduce until the end of the career. However, the gender earnings gap has been reducing over generations when we compare workers from older and younger birth cohorts. For a specific generation, born between 1967 and 1974, the contribution of the gender variable to the earnings inequality increases with the educational level: gender contributes with 1.14 per cent for workers with less than high school completed, 4.17 per cent for workers with high school degrees and 5.46 per cent for those with more than high school. However, in all of the three groups of schooling covered, the marginal contribution of gender drops to near zero after controlling by

---

<sup>1</sup> According to the Global Gender Gap Report, provided by the World Economic Forum, in 2016 Brazil occupied the 79<sup>th</sup> place among the 144 countries covered in the research.

occupation/activity (using 2-digits classification) and firm characteristics (size, average earning and fixed effects).

We split the cohort born in 1967–74 according to these three groups of schooling mentioned above in order to explore the gender gap on earning over the life cycle by level of education. We observe, for all of the three groups and independent of the specific age, that the gender gap on earning reduces when occupation/industry controls are added to the model. The inclusion of firm controls also accounts for a great part of the gender gap in earnings. In contrast with what we find for workers without high school degrees, which expanded the gender gap throughout the life cycle, the gender gap is relatively more stable over the life cycle for workers who were high school graduates. Furthermore, for this group, we see a reduction in the gender gap overall. Although college graduates show a higher gender gap over the entire life cycle, for this group the sector (occupation/activity) play a fundamental role in explaining the gender gap throughout the entire life cycle of workers. Finally, even after controlling for all the observable characteristics, we observe a gender gap greater than 10 per cent and lower than 20 per cent over the entire career, independent of the educational level.

This paper is divided as follows. The next section details the data used in this paper and the filters applied to the sample before the empirical analyses. Section 3 brings a summarized background of the formal labour market in Brazil and its evolution over the last two decades for key variables and for both men and women. Section 4 describes the empirical strategy to be used to achieve the objective proposed. Section 5 presents the findings and Section 6 concludes this paper.

## 2 Data

This research uses RAIS (Relação Anual de Informações Sociais), the matched employer-employee data for Brazil, provided by the Brazilian Ministry of Labour. RAIS is a longitudinal data set covering the universe of the formal labour market in Brazil through restricted-access administrative records with an average of 33 million observations per year from 1994 to 2015.

In Brazil, the firms are required to report all the workers formally employed at some point in the previous calendar year and each worker is identified by a unique number (PIS, *Programa de Integração Social*), which allows us to follow the employees over time and across firms. The firms also have a unique identifier (CNPJ, *Cadastro Nacional de Pessoa Jurídica*). Thus, our dataset allows us to track workers and firms over time. RAIS contains a set of variables on both firms' and employees' characteristics as well as about the features of the employment contract. To be precise, the information in the dataset includes firm-related information (sector of activity, size, state, etc.), worker-related characteristics (gender, age, schooling, etc.) and job-related features (earnings, occupation, weekly hours of work, etc.).

In this study, we restrict the analyses to those employment contracts that were active on December 31. In case of more than one employment, we select the job with a higher salary (in minimum wages). We calculate the real earnings in December 2015 by multiplying the variable '*wage in December (in minimum wages)*' by the value of the minimum wage in each year and using the INPC (*Índice Nacional de Preços ao Consumidor*) as the deflator. We construct an annual panel and study the time evolution of employment, earnings and the probability of working full-time (at least 40 hours weekly) by gender.

In order to study the earning gender gap between men and women over their life cycle, we restrict the panel dataset to a specific cohort and divide the sample into three education groups, similar to the study of Goldin et al. (2017). For each group, we pick specific ages at the start of the data set in order to track those individuals who would probably have completed their highest grade by then,

yet be young enough to be in their earliest jobs. We consider the following three groups: 1) individuals who were not High School graduates at ages 20–22 in 1994; 2) those who were High School graduates, but did not go to the college, at ages 23–25 in 1994 and 3) those who were College graduates at ages 25–27 in 1994. We track these groups of individuals, born between 1967 and 1974, in the formal Brazilian labour market and follow them over two decades after school-leaving, from 1994 to 2015.

We chose the start year of our data, i.e. 1994, because it is the earliest year when we have information about all the variables to be used in the regression analysis (later detailed). On the other hand, 2015 is the most recent year when we have access to the RAIS data set.

### 3 Background

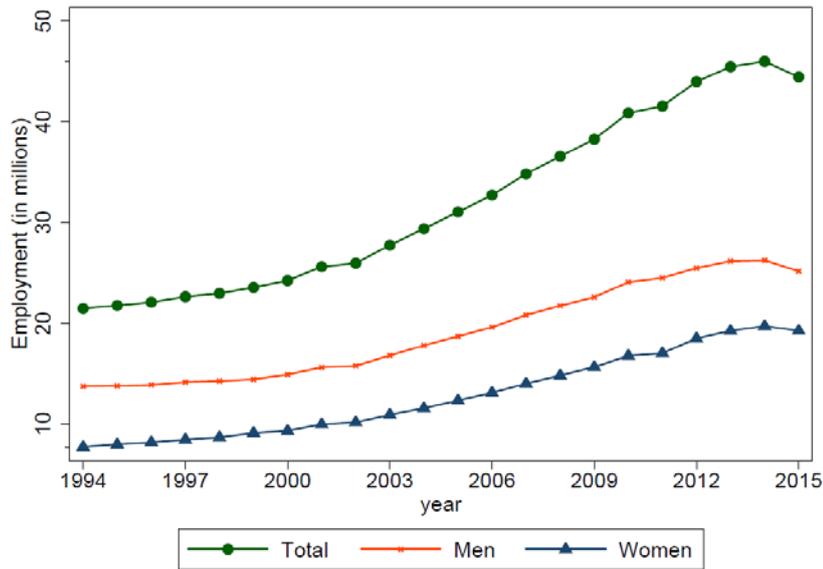
In the past two decades, Brazil has been very successful in reducing gender inequality. Our calculations from the *Pesquisa Nacional de Amostras a Domicílio* (PNAD, the main Brazilian National Household Survey, from the *Instituto Brasileiro de Geografia e Estatística* (IBGE)) show that overall gender labour differentials fell from 185.6 per cent to 86.1 per cent, reflecting relative gains in all labour market components such as participation rates, unemployment rates, hours worked and hourly wages. However, gender inequality remains high in many dimensions and, according to the Global Gender Gap Report, provided by the World Economic Forum, in 2016 Brazil occupied the 79<sup>th</sup> place among the 144 countries covered in the research.

A key challenge to achieve equality is to overcome gender gaps in access to formal employment and market income. For instance, the fraction of women occupied with formal jobs increased from 39.24 per cent in 1995 to 63.11 per cent in 2015, which is now higher than that of employed men, which, on the other hand, stood at 44.46 per cent in 1995 and 60.34 per cent in 2015 (see also Agénor et al 2013). 2011 was the year when the formalization of occupied females surpassed that of males. This comparison, controlled by basic socio-demographics characteristics, in particular education, is required. Years of schooling is perhaps the main single determinant of the increasing formalization trends found in the Brazilian literature (Barbosa e Moura 2015). In 1997 females' years of schooling overtook those of males in the population above 25 years of age. The chance of getting a formal job for individuals with the same observable characteristics, which was 67.41 per cent in 1995, fell to 18.79 per cent in 2015, still favourable to males<sup>2</sup>. In the appendix we present the evolution of various gender gaps, including formalization, for the whole Brazilian labour market using PNAD from 1995 to 2015. We develop in this section an analysis on similar data based on the formal employment from RAIS. Figure 1 shows the evolution of employment in the Brazilian formal labour market and disaggregates this information by gender. The total number of formal employees in Brazil has grown more than double over the last two decades, especially after 2000. The number of employed persons increased from around 24.2 million in 2000 to almost 44.4 million in 2015.

---

<sup>2</sup> The regression is controlled by race, migration status, Unit of Federation, quadratic polynomials for age and education. If we use similar controls, the difference in difference estimated of social security contribution between these two years is 1.34 in favour of females.

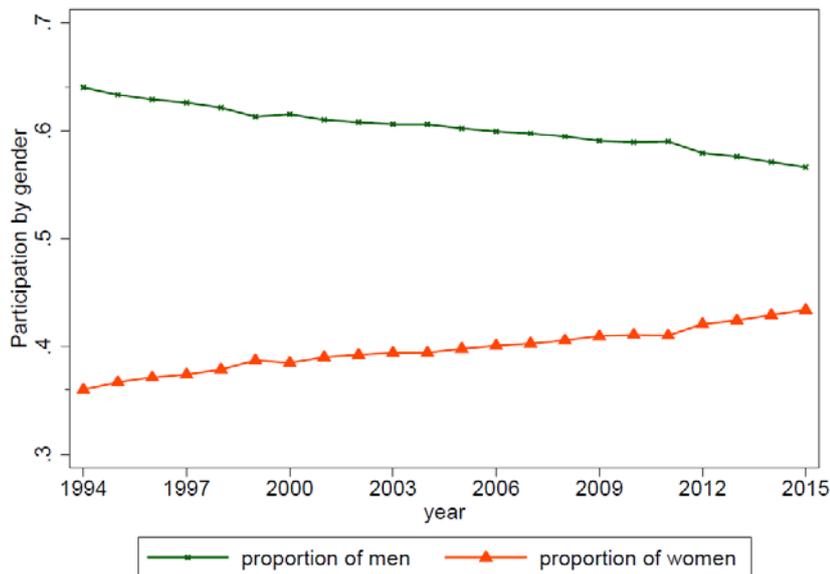
Figure 1: Evolution of employment in the Brazilian formal labour market for the period 1994–2015 by gender



Source: Authors' calculation based on RAIS microdata.

An important aspect of this period was the shift in the gender composition of the formal labour market, with the female participation increasing by more than seven percentage points (from 35.97 per cent to 43.37 per cent) over the period 1994–2015 (Figure 2).

Figure 2: Gender composition of employment in the Brazilian formal labour market for the period 1994–2015



Source: Authors' calculation based on RAIS microdata.

Despite considerable formalization of the Brazilian labour market (mainly among women), the period 1994–2015 was also marked by an increase in the real average earnings. Figure 3 shows the trajectory of the average earnings received by both men and women. For both genders, the time evolution of the earnings presented similar patterns, with a period of stability from 1994–2003 and a substantial increase from 2003 onwards.

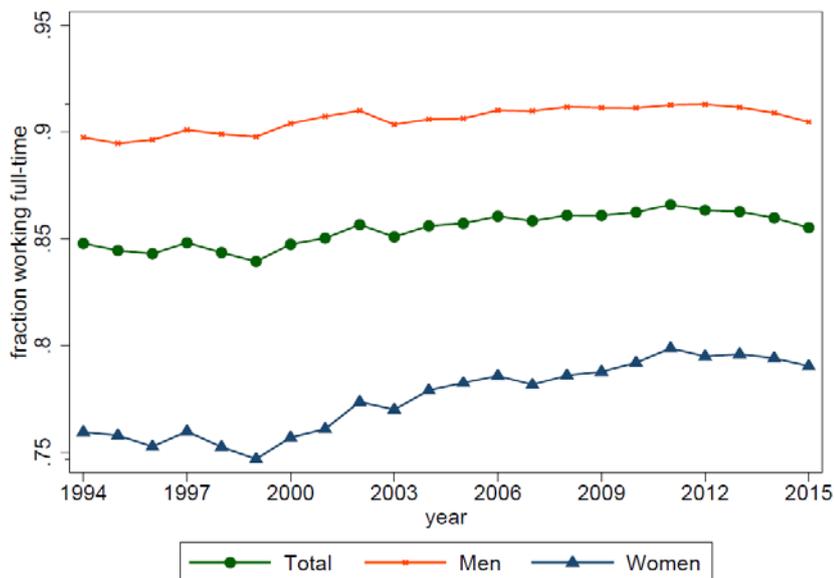
Figure 3: Evolution of real average earnings in the Brazilian formal labour market for the period 1994–2015 by gender



Source: Authors' calculation based on RAIS microdata.

When we compare the proportion of men and women working full-time, we observe that, relative to men, women are less likely to work full-time (Figure 4), i.e., at least 40 hours weekly. However, the hours spent by female workers in the formal labour market have increased over the period of analysis.

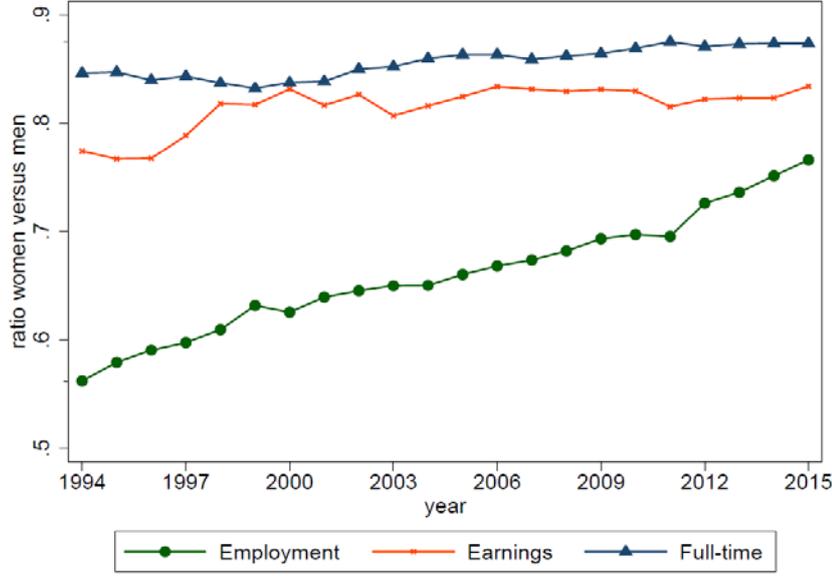
Figure 4: Evolution of share of workers working full-time in the Brazilian formal labour market for the period 1994–2015 by gender



Source: Authors' calculation based on RAIS microdata.

Finally, Figure 5 summarizes the information plotted on each previous graph, by considering the temporal evolution of the gender differences (measured as the ratio women/men) in employment, earnings, and in the probability of working full-time. As we can see, we find that the substantial change in the Brazilian formal labour market was the remarkable increase in the female participation in comparison to men.

Figure 5: Evolution of the gender differences (ratio women to men) in employment, real average earnings and probability of working full-time for the period 1994–2015



Source: Authors' calculation based on RAIS microdata.

The previous analysis considers the trends in gender gap over time. Although there has been a remarkable growth in female labour market participation over the last two decades, this change may have different patterns throughout the life cycle of individuals. In addition, these different patterns may depend on a variety of factors such as human capital, occupational choices, establishment characteristics and so on. Understanding these patterns is precisely the purpose of this study and the way that we investigate this issue is described below.

#### 4 Empirical strategy

We analyse a large panel of individuals  $i$  for each year  $t$  from 1994 to 2015. We estimate the following baseline econometric model, close related to Barth et al. (2017) and Goldin et al. (2017), for three education groups  $j$ :

$$Y_{ijt} = \alpha_{tj} + X_{ijt}'\beta + \sum_l \theta_{lj}I(a_{ijt} = l) + \sum_l \varphi_{lj}I(a_{ijt} = l) \times F_{ij} + \epsilon_{ijt} \quad (1)$$

where  $Y_{ijt}$  is the real earnings (in log levels) for individual  $i$  of education level  $j$  in year  $t$ . The term  $\alpha_{tj}$  is a year-fixed effect for each educational group. The vector  $X_{ijt}$  contains a set of individual characteristics (exact education and state) and  $F_{ij}$  is an indicator variable for women.  $I(a_{ijt} = l)$  is an indicator of age  $l$  for each individual  $i$  of education level  $j$  in year  $t$ . We normalize  $\theta_{\min\{l\}j} = 0$  and the effect of age can be interpreted relative to the youngest age in the data set (20-year-old for the first group, 23 for the second group and 25 for the third group). We are interested in studying the pattern of the coefficient  $\varphi_{lj}$ , which gives us the pattern of gender gap over the life cycle by the level of education. The idiosyncratic error is  $\epsilon_{ijt}$ , in which we suppose the classical hypothesis.

In addition, we extend the Equation 1 in order to evaluate the effect of the occupational/industry and firms characteristics (hereafter we omit the educational index  $j$ ). We estimate the following equation:

$$Y_{it} = \alpha_t + X'_{it}\beta + \sum_l \theta_l I(a_{it} = l) + \sum_l \varphi_l I(a_{it} = l) \times F_i + S'_{it}\mu + \epsilon_{it} \quad (2)$$

This equation includes variables indicating the occupation sector (for the employment observed in 1994) and the firm's sector of activity; in both of them the two-digit classification was considered. The comparison between  $\varphi_{lj}$ , estimated from Equation 1 and 2, allows us to understand in which extent the evolution of the gender gap over the life cycle can be explained by sector characteristics.

We also expand the model to include controls for job-specific effect as follows:

$$Y_{it} = \alpha_t + X'_{it}\beta + \sum_l \theta_l I(a_{it} = l) + \sum_l \varphi_l I(a_{it} = l) \times F_i + S'_{it}\mu + \Omega_{e(it)t} + \epsilon_{it} \quad (3)$$

The term  $\Omega_{e(it)t}$  is a job-specific effect, where  $e(it)t$  identifies the establishment  $e$  where individual  $i$  is employed at time  $t$ . A job is defined as a match between an individual and an establishment in a specific year  $t$ . Further, we decompose  $\Omega_{e(it)t}$  into an establishment-fixed effect,  $\Gamma_e$ , and a time-variant component containing the size,  $S_{e(it)t}$ , in terms of the total number of employees and the average earning of the establishment,  $MP_{e(it)t}$ , (both of them in log scale). We set  $\Omega_{e(it)t} = \Gamma_e + \delta S_{e(it)t} + MP_{e(it)t}$ . Notice that both the size and the average earning of the establishment may vary, either within the same establishment over time, or as the individual changes establishment. Similar to the previous analysis, by comparing the  $\varphi_{lj}$  in Equation 3 and Equation 2, we can observe how much the gender gap over the life cycle is driven by a sorting of employees into firms.

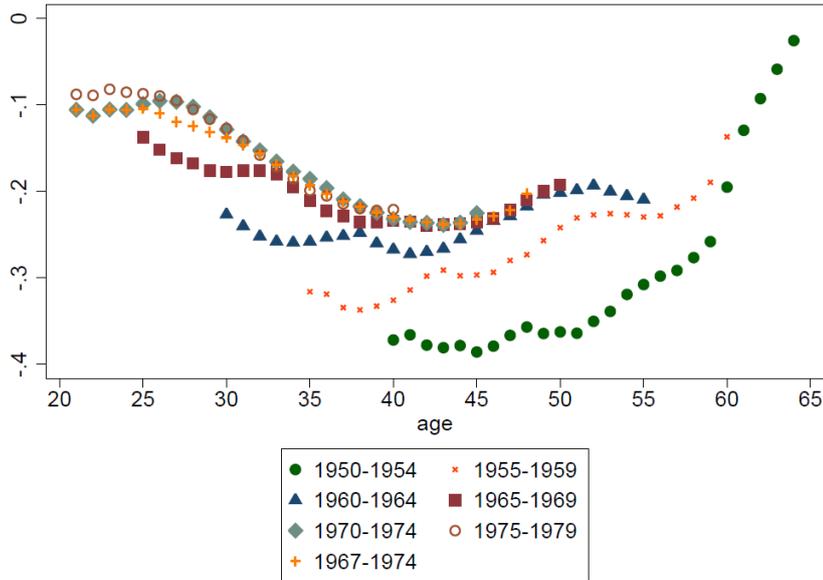
In sum, our approach consists in analysing the pattern of the coefficient  $\varphi_{lj}$ , which represents gender earnings gap, over the age profile for each of the three education groups. Additionally, we observe how the pattern of the coefficient changes as we incorporate more variables (representing the firm and job characteristics) in our econometric model, which will give us the relative contribution of each factor to explain the variability of gender gaps. This analysis uses an unbalanced panel, since individuals who were employed in 1994 may leave the labour market over time and hence we can no longer observe them nor their characteristics.

## 5 Results

Before presenting our main findings, we describe the temporal trajectory of the Brazilian formal labour market background for the period 1994–2015. Based on that longitudinal data, we divide the workers who were in the formal job market at some point during the period 1994–2015 into groups according to their birth cohorts. In Figure 6 we present the evolution of the earnings gender gap (in log differences) over the life cycle for the following cohorts: those born in 1950–54, 1955–59, 1960–64, 1965–69, 1970–74, 1975–79 and 1967–74.

Basically, we find an expanding gender gap over the life cycle until around the age of 40, when the gender gap starts to reduce towards the end of the career. In addition, we see that the gender gap is higher for the older cohorts, which shows that part of the gender inequality has reduced over generations. When looking at the cohort born between 1967–74, which will be the basis for our later analysis, we notice that at the beginning of the career 20-year old women earn 6.79 and men 6.85 (in log points) approximately, which means that women earn around 6 per cent less than men on average. This gap is increasing over the age profile and reaches its maximum value, almost 24 per cent (8.09–7.85), at around 43 years of age, when it begins to fall.

Figure 6: Evolution of the earnings gender gap (in log differences) throughout the life cycle by birth cohorts

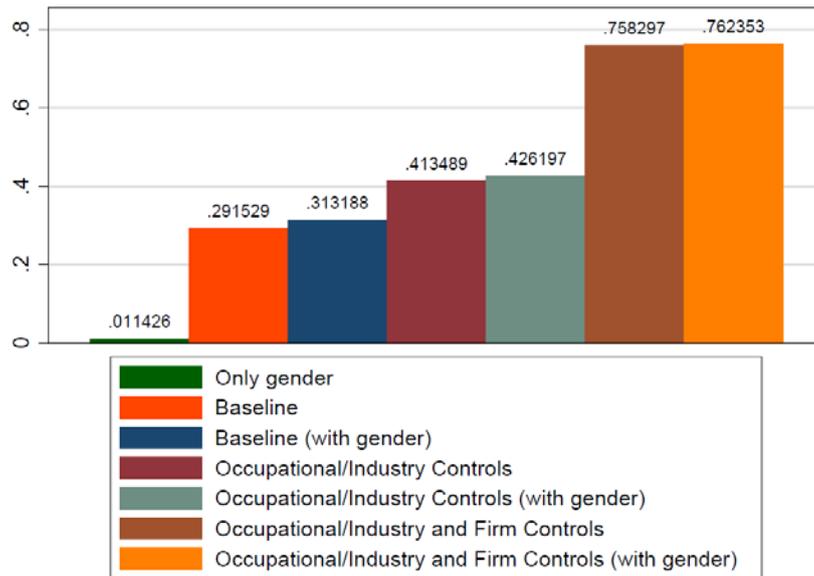


Source: Authors' calculation based on RAIS microdata.

According to what was described previously in the data section, we restrict our analysis of the cohort born in 1967–74 by dividing it into three age-education groups: 1) NOT High School Graduates at ages 20–22 in 1994; 2) High School graduates, but NOT College, at ages 23–25 in 1994 and 3) College Graduates at ages 25–27 in 1994. Notice that we restrict the analyses for those who were employed in the labour formal market in 1994. For each of these groups, the Figures 7, 8 and 9 show how much the R-squared of the regressions increase as we include occupation/activity and firm controls. Moreover, in each econometric model, we measure the marginal contribution of the gender variable to explain the variance of log earnings.

The first column of the Figures 7, 8 and 9 show the R-squared in a simple regression with only the gender variable included in the model and the real earnings (in log levels) as the dependent variable. For the first educational group (workers with less than high school degree) the gender variable alone accounts for 1.14 per cent of the inequality, measured by the variance of the log earnings. For the most educated workers, this percentage is higher, in order of 4.17 per cent for workers with high school degrees and 5.46 per cent for those with more than high school.

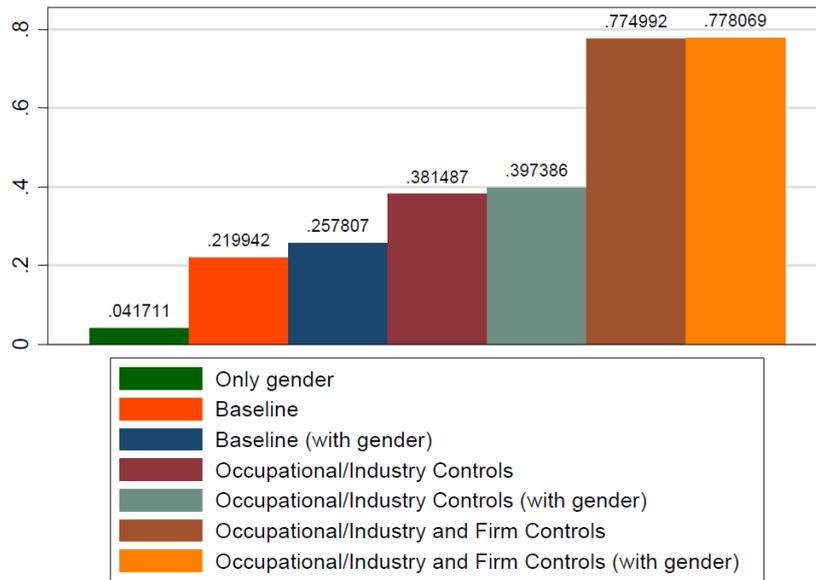
Figure 7: Relative importance of the controls to explain the variance of the log of earnings for NOT High School graduates at ages 20–22 in 1994



Source: Authors' calculation based on RAIS microdata.

The baseline econometric model includes other workers' characteristics such as age, state, exact educational level and year-fixed effects. As can be seen from Figure 7, these variables explain a considerable part of the earnings inequality, especially for the lower educational group, in which the R-squared increased from 1.14 per cent to 31.31 per cent. Despite the relatively smaller contribution, the baseline variables also explain a significant part of the earnings inequality for the most educated groups. On the other hand, we find that the marginal contribution of the occupation/industry controls is relatively higher for the more educated groups, explaining approximately 15.11 per cent (36.55%–21.44%) of the inequality when we look at the college graduates (Figure 9).

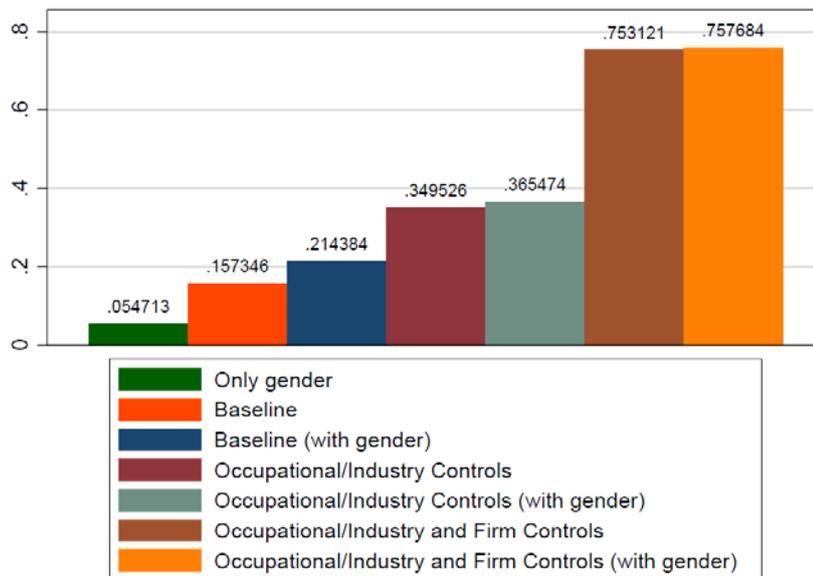
Figure 8: Relative importance of the controls to explain the variance of the log of earnings for High School graduates, but NOT College, at ages 23–25 in 1994



Source: Authors' calculation based on RAIS microdata.

Finally, when we include firm controls (size, average earnings and fixed effects) we observe that the R-squared rises substantially, reaching more than 75 per cent for all educational groups. This finding reveals that a large part of the inequality can be explained by a sorting between high-paying and low-paying companies rather than inequality within firms. Moreover, the marginal contribution of the gender variable becomes negligible when we consider the full model with all the controls.

Figure 9: Relative importance of the controls to explain the variance of the log of earnings for College Graduates at ages 25–27 in 1994



Source: Authors' calculation based on RAIS microdata.

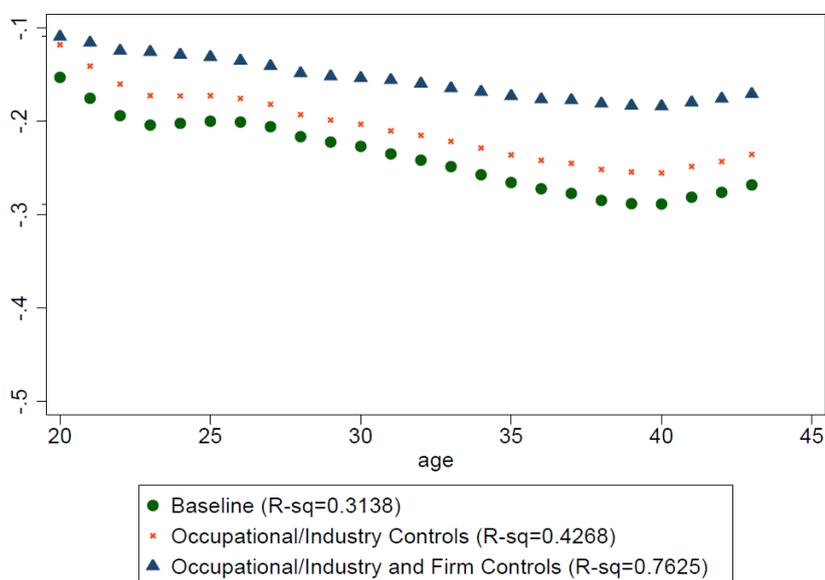
In the following figures, we present the main results based on our empirical strategy laid out in Section 4. Each of the graphs in Figures 10, 11 and 12 corresponds to a specific educational group as previously defined. The baseline model is as described in Equation 1 and we sequentially include occupation/industry and firms' controls, according to the Equations 2 and 3. It is worth

mentioning that, unlike the previous 3 graphs, in this part, we basically interact the gender variable with the indicators of age in order to measure the earnings gender gap throughout the life cycle.

The first important result is that the gender gap increases with the level of education. For instance, comparing the baseline model scenario (without Occupation/Industry and Firm controls) in Figures 10, 11 and 12, we observe that the gender gap at age of 25 is approximately 20 per cent for the group with less than high school education, 35 per cent for the group with exactly high school and this gap reaches almost 40 per cent for the college graduates' group.

Figure 10 shows that the gender gap is expanding throughout the life cycle, reaching its maximum at around 40 years of age. We also observe that the earning gender gap, independent of the specific age, reduces when occupation/industry controls are added to the model (considering 2-digits classification). However, in relative terms, the inclusion of firm controls (size, average earning and fixed effects) account for most of the gender gap in the earnings. When all the controls are added, a gender gap of less than 20 per cent remains throughout the entire worker life cycle.

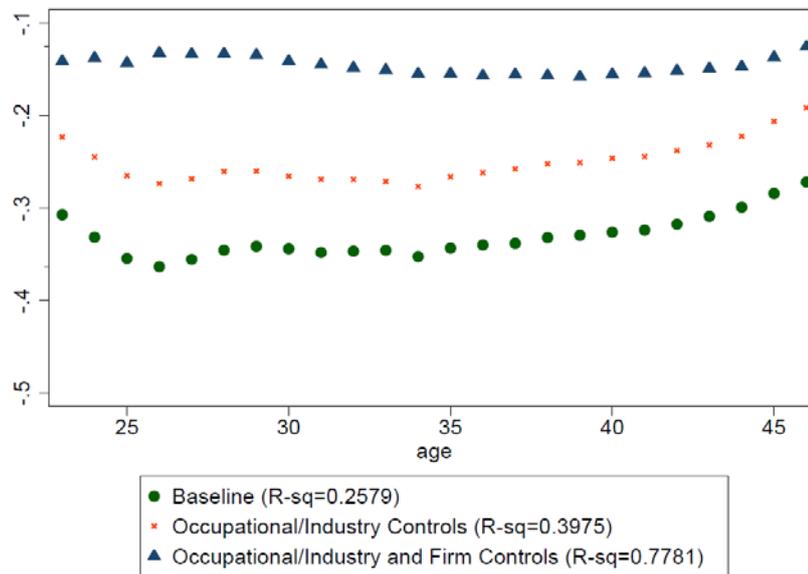
Figure 10: Importance of the controls in the earnings gender gap (in log differences) throughout the life cycle for NOT High School Graduates at ages 20–22 in 1994



Source: Authors' calculation based on RAIS microdata.

Differently from what we have found for workers without High School degrees, we observe that the gender gap is relatively more stable throughout the life cycle of workers who were High School graduates. Furthermore, we see a reduction in the gender gap overall, especially after 40 years of age. Regarding the sectorial effect (occupation/industry controls), we find that its relative importance in explaining the gender gap grows throughout the workers' life cycle. Especially early in the career, the characteristics of the firm represent most of the gender gap. For the High School graduates, when using all controls, the gender gap does not exceed 16 per cent at any point in the life cycle.

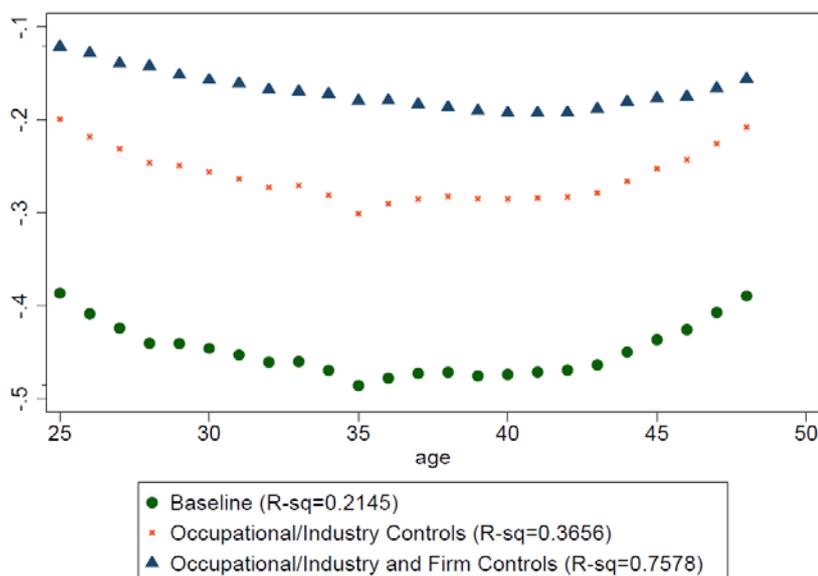
Figure 11: Importance of the controls in the earnings gender gap (in log differences) throughout the life cycle for High School graduates, but NOT College, at ages 23–25 in 1994



Source: Authors' calculation based on RAIS microdata.

Finally, for the college graduates, the sector (occupation/activity) plays a fundamental role in explaining the gender gap throughout the entire life cycle of workers. When controlling for occupational/industry categories, the gender gap reduces by almost 20 percentage points for any age level. On the other hand, the gender gap reduces by another 10 percentage points when we add firm controls in the econometric model. Thus, for the group with the highest level of human capital, even after controlling for all observable characteristics, there remains a gender gap greater than 10 per cent at the beginning of the career, which reaches its maximum value at age 40 and then decreases after that age.

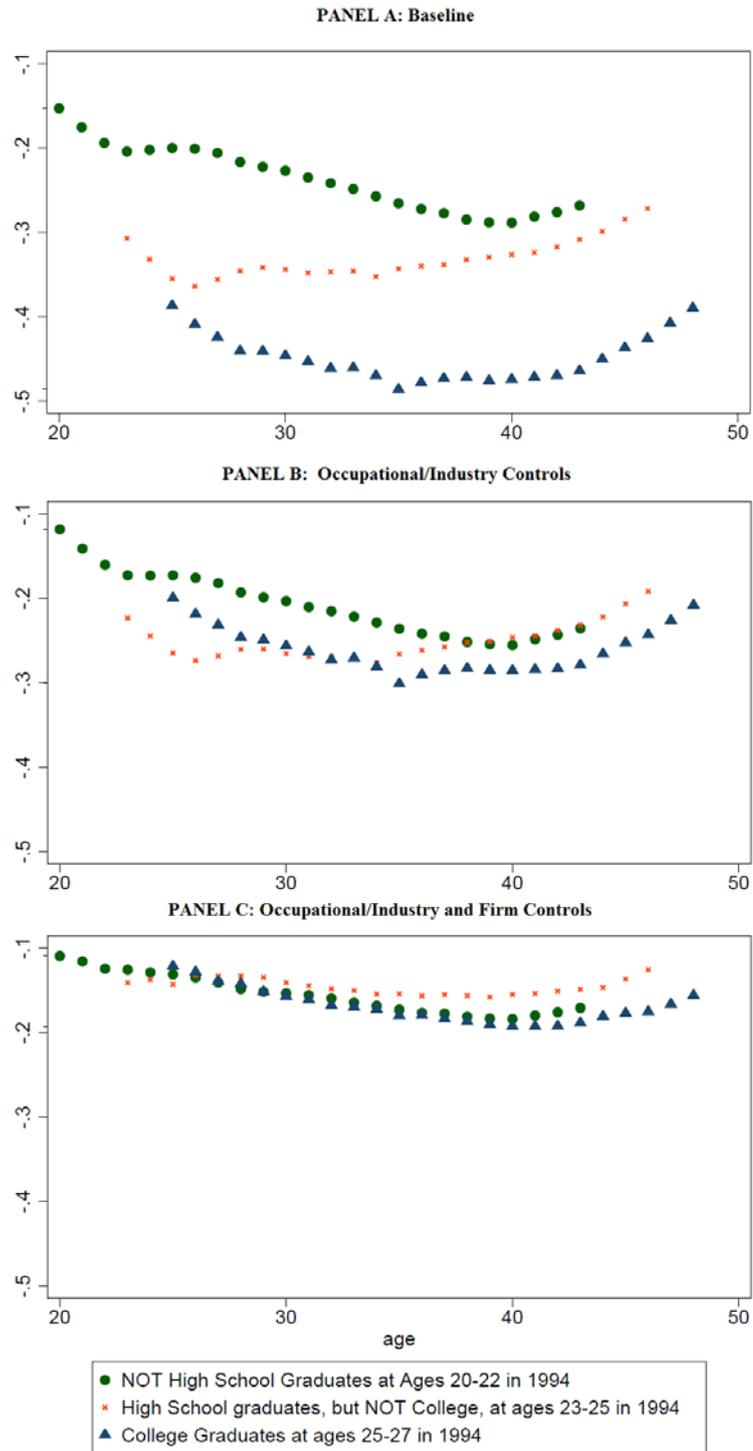
Figure 12: Importance of the controls in the earnings gender gap (in log differences) throughout the life cycle for College Graduates at ages 25–27 in 1994



Source: Authors' calculation based on RAIS microdata.

Figure 13 summarizes the results of the three previous graphs by considering a different perspective. It focuses on the differences in the gender gap by educational groups. The gender gap increases with the educational level in our baseline model, which excludes occupation/activity information and firm characteristics (Panel A). The gender gap for college graduates is around double the gender gap observed for workers without high school. The inclusion of occupation/industry controls significantly reduces the gap between male and female in terms of earnings (Panel B). Furthermore, controlling also for firm characteristics, we see that the gender earnings gap is around 13 per cent at 25 years of age and presents a similar pattern among educational groups. In addition, throughout the entire life cycle, the percentage difference in earnings between men and women does not exceed 20 per cent (Panel C).

Figure 13: Relative importance of the controls in the earnings gender gap (in log differences) throughout the life cycle by educational level



Source: Authors' calculation based on RAIS microdata.

## 6 Conclusion

In this paper we use a large linked employer-employee dataset to study the gender gap in the Brazilian formal labour market. We started with a comprehensive picture of the labour market in Brazil from 1994 to 2015, highlighting the fall of gender differences in employment, earnings and worked hours per week in one of the most unequal countries of the world.

We analyse the patterns of gender gap over the life cycle, disentangling the influences of three separate processes. First, the returns related to innate characteristics such as cohort and geographical location, and early acquired attributes, such as education. Second, we consider that part of the gender gap may reflect differences in gender composition among the occupation/activity sectors and the fact that employees in different sectors can be paid differently. Finally, the career path can vary, within an employer, due to wage raises and promotions over time or according to the sorting of employees into high- versus low-paying employers.

We find that the gender earnings gap expands over the life cycle until around the age of 40, when the gender gap starts to reduce until the end of the career. However, the gender earnings gap has reduced over generations. For the generation born between 1967 and 1974, the contribution of gender to the earning inequality increases with education ladder from 1.14 per cent for workers with less than high school completed to 4.17 per cent for workers with high school degrees and 5.46 per cent for those with more than high school. However, the marginal contribution of the gender variable becomes negligible (lower than 0.4 per cent in all education groups) when we consider the full model, which includes both occupation/industry (considering 2-digits classification) and firm controls (size, average earning and fixed effects).

The cohort born in 1967–74 was analysed by dividing it into three groups of schooling: 1) not High School graduates at ages 20–22 in 1994; 2) High School graduates, but not College, at ages 23–25 in 1994; and 3) College Graduates at ages 25–27 in 1994. We observe that the earning gender gap, independent of the specific age or level of education, reduces when occupation/industry controls are added to the model. The inclusion of firm controls also accounts for reducing a large part of the gender gap in the earnings over life cycle.

Differently from what we have found for workers without high school degrees, which shows an expanding gender earnings gap throughout the life cycle, we observe a more stable pattern for workers who were high school graduates. For this group, we find a reduction in the gender gap overall and an increase in the relative importance of occupation/industry controls to explain the gender gap throughout the workers' life cycle. Finally, for the College graduates, the sector (occupation/activity) plays a fundamental role in explaining the gender gap throughout the entire life cycle of workers. When controlling for occupation/industry categories, the gender gap reduces by almost 20 percentage points for any age level and it reduces by another 10 percentage points when we add firm controls. For all the levels of education, even after controlling for all observable characteristics, the remaining gender gap is greater than 10 per cent but lower than 20 per cent throughout the entire worker life cycle.

## References

- Agénor, P.-R., O. Canuto, and others (2013). “Gender equality and economic growth in Brazil”. *World Bank-Economic Premise*, 109: 1–5, The World Bank.
- Barbosa Filho F. H., R.L. Moura (2015). “Evolução recente da informalidade do emprego no Brasil: uma análise segundo as características da oferta de trabalho e o setor”. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), 45: 101–23.
- Barth, E., S. Pekkala Kerr, and C. Olivetti (2017). “The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data”. NBER Working Paper No. 23381. Cambridge, MA: National Bureau of Economic Research.
- Bertrand, M., C. Goldin, and L.F. Katz (2010). “Dynamics of the gender gap for young professionals in the financial and corporate sectors”. *American Economic Journal: Applied Economics*, 2(3): 228–55, American Economic Association.
- Blau, F.D., and L.M. Khan (2008). “Women’s work and wages”. *The new Palgrave dictionary of economics*, 8: 762–72.
- Blau, F.D., and L.M. Kahn (2016). “The gender wage gap: Extent, trends, and explanations”. NBER Working Paper No. 21913. Cambridge, MA: National Bureau of Economic Research.
- Goldin, C., and J. Mitchell (2017). “The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops”. *The Journal of Economic Perspectives*, 31(1): 161–82, American Economic Association.
- Goldin, C. (1989). “Life-cycle labor-force participation of married women: Historical evidence and implications”. *Journal of Labor Economics*, 7(1): 20–47, University of Chicago Press.
- Goldin, C., S. Pekkala Kerr, C. Olivetti, E. Barth, and others (2017). “The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census”. *American Economic Review*, 107(5): 110–14, American Economic Association.
- Heckman, J.J., and R.J. Willis (1977). “A beta-logistic model for the analysis of sequential labor force participation by married women”. *Journal of Political Economy*, 85(1): 27–58, The University of Chicago Press.

## Appendix

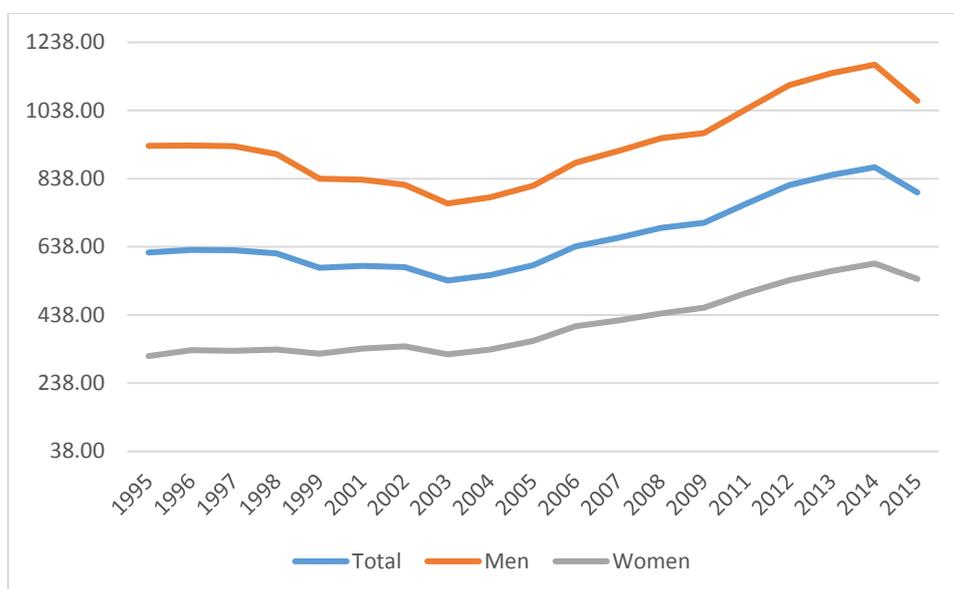
### The evolution of gender gaps in overall Brazilian labour market according to PNAD/IBGE harmonized series from 1995 to 2015

Figure A1: Evolution of share of formal jobs (social security contribution among occupied) in the Brazilian labour market (%) for the period 1995–2015 by gender



Source: Authors' calculation based on PNAD microdata.

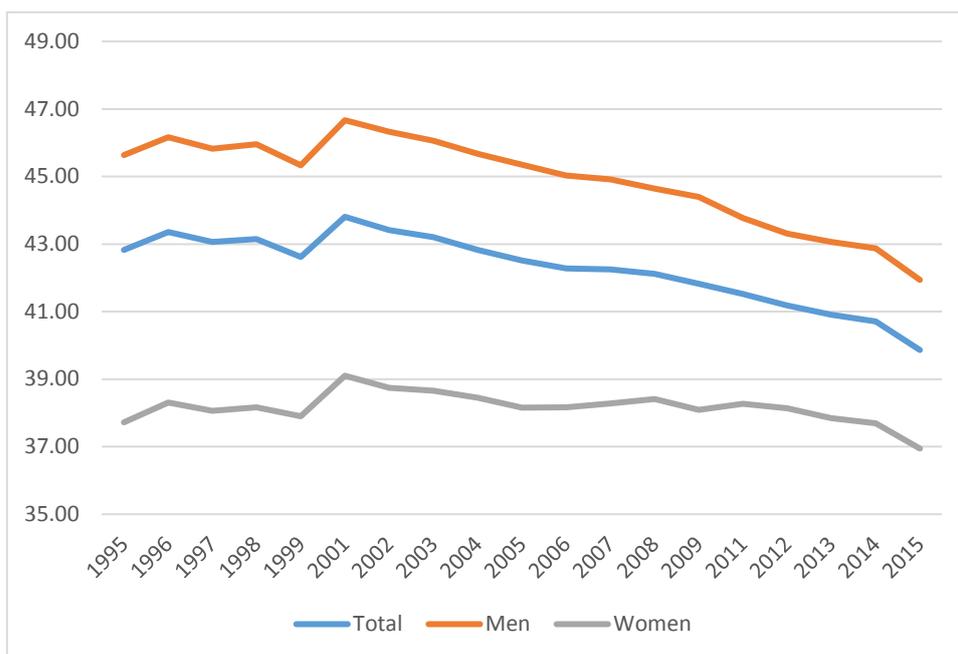
Figure A2: Evolution of real average earnings in the Brazilian labour market for the period 1995–2015 by gender<sup>3</sup>



Source: Authors' calculation based on PNAD microdata.

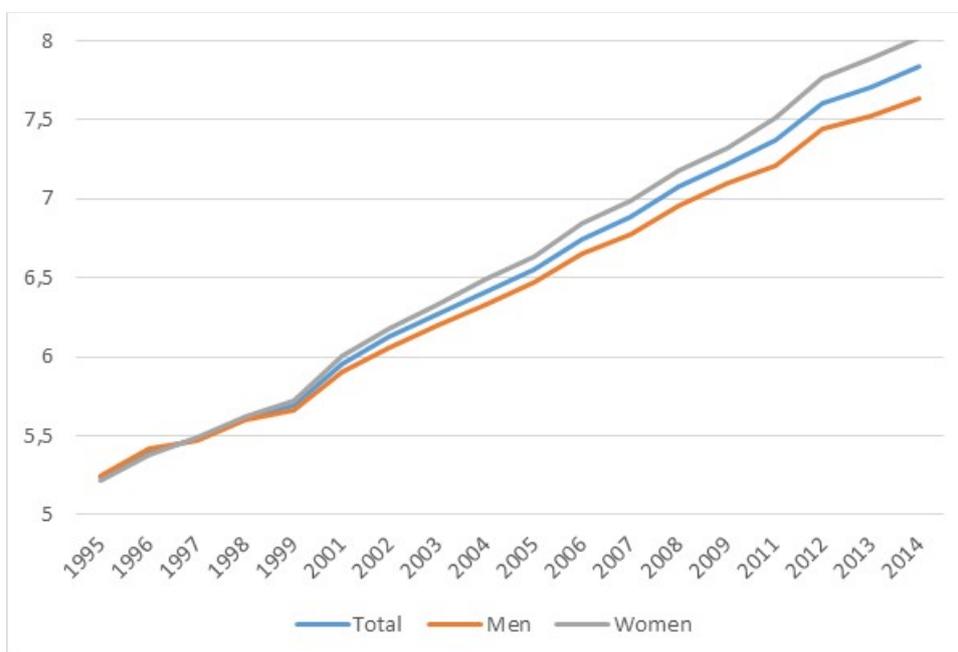
<sup>3</sup> In the overall labour market, the gross contribution of gender to inequality falls from 3.95 per cent in 1995 to 2.21 per cent in 2015, while the pure gender gap falls from 41.75 to 25.58 per cent. If we restrict the sample in PNAD to formal jobs the wage gap falls from 30.52 per cent in 1995 to 17.23 per cent in 2015.

Figure A3: Evolution of in the Brazilian labour market – weekly working hours - for the period 1995–2015 by gender



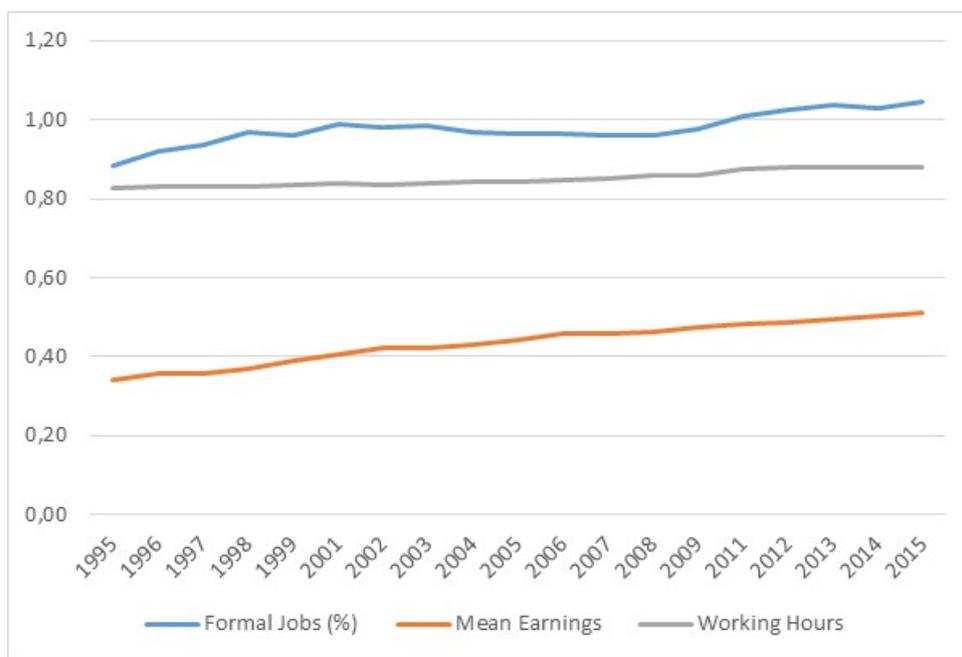
Source: Authors' calculation based on PNAD microdata.

Figure A4: Evolution of in the Brazilian labour market – years of schooling 25 years or more - for the period 1995–2014 by gender



Source: Authors' calculation based on PNAD microdata.

Figure A5: Evolution of the gender differences (ratio women to men) in formal employment (%), real average earnings and working hours for the period 1995–2015



Source: Authors' calculation based on PNAD microdata.