A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the NOVA School of Business and Economics.

THE ROLE OF FIRMS IN THE PORTUGUESE GENDER WAGE GAP

JOANA INÊS ALVES LEITÃO

Work project carried out under the supervision of:

Professor Pedro Portugal

16-12-2022

THE ROLE OF FIRMS IN THE PORTUGUESE GENDER WAGE GAP *

Joana Leitão †

December 16, 2022

Abstract

This paper follows the methodology proposed by Card et al. (2016) in order to investigate the role of firms in the Portuguese gender wage gap. It was found that firm components account, on average, for 9.8% of the gender pay gap and that its relevance increases for older and less educated workers. The sorting channel is the main driver of these components and the bargaining effect was proven to be sensitive to the normalization strategy. Additionally, it was found that, on average, women are 1.6 percentage points less likely to move to higher-paying firms than their male counterparts.

JEL Codes: J16, J31, J17.

Keywords: Gender wage gap, labour market discrimination, firm fixed effects, AKM model

---

*I would like to express my deepest gratitude to my advisor, professor Pedro Portugal, for all the patience and guidance throughout this process. Additionally, I would also like to thank my friends and family for the support.

†Nova School of Business and Economics, Universidade Nova de Lisboa. Email: 39134@novasbe.pt

1 Introduction

Gender discrimination in the labour market has been a widely discussed topic for decades and encompasses gender differences in earnings, employment and career progression, as well as sectoral and occupational segregation. Particularly, the gender wage gap and its potential sources have been the center of public and academic debates on the matter - as despite the rapid convergence of women in terms of productive characteristics (such as education and experience) differences in wages have persisted over time.

The Portuguese ”raw” gender wage gap has decreased in the last couple of decades, falling from 32 % in 1991[1] to 23.6% in 2020. Meanwhile, the ”adjusted” wage gap, which takes into consideration workers' characteristics, did not suffer significant changes, implying that the unexplained component of the wage gap - usually interpreted as discrimination - has remained constant (Cardoso et al., 2016).

The approximation between male and female workers in terms of observable characteristics has brought to light the potential role of firms in creating and maintaining the gender wage gap through their wage policies. One faction of the literature, in particular, focuses on the role of firm-specific wage premiums. Card et al. (2016) were the pioneers of this type of analysis in Portugal, introducing a simple methodology that allowed not only for the estimation of the contribution of firm components to the gap, but also for its decomposition into sorting and bargaining effects - combining 2 subtopics of wage premium literature.

This paper builds on the work of Card et al. (2016), as it uses the AKM model of Abowd et al. (1999) to estimate the firm fixed effects and then applies an Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1973) decomposition to the estimated effects. The main contribution of the

[1](Cardoso et al., 2016)

present study to the literature was the use of a different normalization strategy, which increased substantially the number of person-year observations used in the estimations and could, therefore, potentially improve their accuracy. However, the bargaining effects are, by construction, sensitive to the chosen normalization. Despite this sensitivity, Card et al. (2016) adopted an arbitrary normalization strategy, using some robustness checks to prove the strength of their conclusions. In this paper, one of these robustness checks was followed - namely the normalization according to a "low-surplus" industry. However, a different sector was chosen in order to potentially improve the results.[2] This small amendment to the original paper provided unconvincing results for the bargaining effects, highlighting the need for a less sensitive estimation method.

Overall, it was found that firm components represent 9.8% of the Portuguese gender wage gap, with the sorting effect being its main driver. The sorting effect, on average, accounted for 31% of the pay gap and its robustness was tested using a Gelbach decomposition (Gelbach, 2016).[3] The bargaining effects, which were inconsistent with the results of Card et al. (2016) were, on average, negative. Taking into consideration the importance of the sorting channel - and following the ideas of Casarico and Lattanzio (2019) - a probit regression was estimated in order to understand if women were less likely than men to switch to more generous firms when changing jobs. On average, women are 1.6 percentage points less likely than men to move to firms with a higher average wage premium than their previous employer.

The present paper has the following structure. Section 2 presents a review of the vast literature on the gender wage gap, giving particular attention to the replications of Card et al. (2016) in other countries. Section 3 discusses the methodology used, while section 4 describes, in great

[2]In their robustness test, Card et al. (2016) used the restaurant and accommodation industry, as it had the lowest surplus. However, as discussed in section 3.2, this sector had a considerable gap between male and female fixed effects, meaning it is not necessarily desirable to simply set them to 0. Therefore, the textile industry was used in this case.

[3]When considering the value of the sorting effect, please consider that the bargaining effect was negative and that the condition set out in section 3.3 holds at the individual level.

detail, our data set. Finally, while section 5 presents the main results of the analysis, section 6 concludes the paper.

2 Literature Review

There is extensive literature regarding the gender wage gap and its potential sources. Traditionally, greater attention has been provided to supply-side factors, such as differences between men and women regarding human capital (Mincer and Polachek, 1974,Goldin et al., 2006) and labour market participation (Goldin, 2006). However, as both genders converged in terms of education, training, and employment (Blau and Kahn, 2017), it became apparent that firms - through their hiring and wage-setting policies - could have a significant role in the gender pay gap.

Abowd et al. (1999) demonstrated the importance of firms' pay policies as a determinant of wages through a framework latter known as the AKM model , which included both observable and unobservable firm and individual components. Card et al. (2016) built on their methodology in order to measure the impact of firm-specific premiums on the gender wage gap in Portugal. They used matched worker-firm data from 2002 to 2009 to estimate (AKM) wage regression models for both genders, splitting the firm component into sorting and bargaining effects via a Blinder-Oaxaca decomposition (Oaxaca, 1973; Blinder, 1973). The sorting effect can arise from the potential differences in the share of men and women employed at different firms (Petersen and Morgan, 1995), while the bargaining effect might result from a certain degree of wage-setting power of firms. This could imply that women receive and negotiate lower wages, receiving a smaller share of the surplus generated by their work. Card et al. (2016) found that, together, the 2 effects explained one-fifth of the Portuguese gender wage gap, with the sorting effect being particularly important for low-skilled workers.

The work of Card et al. (2016) has been replicated in other countries. Coudin et al. (2018) found that, in France, the sorting channel explained 11% of the total gender wage gap while the bargaining channel was not relevant (if anything, it would be negative, as women were paid as well as their male colleagues when controlling for observable and unobservable individual characteristics). Coudin et al. (2018) discussed how differences in minimum wages between the 2 countries could be the source of the contrasting bargaining effects. Similarly, Bruns (2019) estimated that the sorting effect explained 15% of the German wage gap and the bargaining effect was not relevant. According to the author, de-unionization has an important role in the growing relevance of firm effects in the gender wage gap. In Chile (Cruz and Rau, 2017), the firm effects explain almost 49% of the wage gap, with the sorting component accounting for 70-80% and the bargaining effect having limited relevance. In contrast, Li et al. (2022) found that, in Canada, firm-specific premiums explain 27% of the gender wage gap, with the sorting and bargaining effects each accounting for half of it. Casarico and Lattanzio (2019) estimated that firm pay policies could be responsible for around 30% of the Italian wage gap. The sorting effect explained 20-22% of the total pay gap with the bargaining effect being more relevant at the top of the income distribution. In Estonia, Masso et al. (2022) found that the firm effects explained 40% of the gender wage gap and the bargaining channel was responsible for one-third to a half of the firm effects. These studies seem to agree on the importance of firm effects in the gender wage gap, as well as on the relevance of the sorting channel. However, there were divergent results regarding the bargaining channel, which could account for half of the firm effects (Li et al., 2022) or play an irrelevant role (Bruns, 2019; Coudin et al., 2018).

Card et al. (2016) did not research further the potential sources of the bargaining effect. This ”negotiation gap” could result from women not negotiating their salary as much as their male colleagues (Small et al., 2007) or not being as successful as them in the process (Gerhart and Rynes, 1991) - which could have a compounding effect throughout their career. Besides the social factors that might come into play at the negotiation table, there is one additional element that could constrain the bargaining power of women - job search costs. If women are discriminated in the labour market and have fewer job opportunities than men, it will be more costly for them to search and switch jobs, meaning they would have fewer incentives to refuse an offer or negotiate their salary. Firms would have, therefore, a certain degree of monopsony power which could be used to extract a higher surplus from women.

Despite being introduced by Robinson (1969) as a possible explanation for the gender wage gap, the monopsony model was given rare attention by researchers in the last decades. Recently, it gained more relevance through the ”new monopsony models” (Burdett and Mortensen, 1998; Manning, 2003) which were grounded on more realistic assumptions of labour market frictions and suggested a way of estimating the labour supply elasticity faced by an individual firm. However, there is still limited research relating monopsony to the gender wage gap. Studies by Webber (2016), Hirsch et al. (2010) and Sulis (2011) found that, indeed, female labour supply facing each firm is less elastic than male labour supply but none found evidence of strong monopsony power.

3 Empirical Framework

3.1 The AKM Model

In order to estimate the firm effects and evaluate their potential impact on the gender wage gap, the empirical strategy and decomposition proposed by Card et al. (2016) were followed. Firstly, the two-way fixed effects model introduced by Abowd et al. (1999) was estimated, using the following specification:

$$\omega_{it} = \alpha_i + \psi^{G(i)}_{J(i,t)} + X'_{it}\beta^{G(i)} + r_{it} \tag{1}$$

where ω_{it} represents the logarithm of real hourly wage of worker i ($i \in \{1, ..., N\}$) at time t ($t \in \{1, ..., T\}$); α_i denotes the worker fixed-effect and $\psi^{G(i)}_{J(i,t)}$ stands for the gender-specific firm fixed-effect (*G(i)* ∈ (*F*,*M*)) in firm J ($J \in \{1, ..., J\}$). X'_{it} represents a set of time-varying observable characteristics of each individual (such as age, tenure and education years), meaning that $\beta^{G(i)}$ captures the gender-specific returns to these covariates. Lastly, r_{it} is the residual error.

Equation (1) is a fundamental piece of this study, as it allows the estimation of gender-specific firm fixed effects. However, as shown by Abowd et al. (2002), it can only be accurately estimated (by conventional methods) within connected groups of individuals and firms - which are referred to as *connected sets*. These connected sets contain all the workers that have ever worked for any of the firms within the group and all the firms that have ever employed one of the workers of the group.[4]

Following the steps of Card et al. (2016), Equation (1) was estimated separately for male and female workers, using the largest connected set of each gender. In both cases, the largest connected set contained more than 98% of all person-year observations.[5]

3.2 Normalization of Firm Fixed Effects

The estimated firm fixed effects were, however, not comparable, since they were estimated separately for each gender.[6] Therefore, there was the need to normalize them with respect to a

[4]As highlighted by Abowd et al. (2002), the connected sets can be seen as "realized mobility networks" within an economy.

[5]For more detailed information on the number of observations within the male and female connected sets, please resort to Tables A1 and A2 in the Appendix

[6]Simply put, firm effects represent the average wage premium of a firm compared to a base category. Therefore, as male and female firm effects were estimated within 2 different connected sets, they were not presented in a comparable scale.

common criterion. In order to do that, only observations within the largest dual-connected set were considered.[7]

Card et al. (2016) normalized the firm effects by setting the average wage premium of a group of "low-surplus firms" to 0, using mean log value added per worker as a surplus measure.[8] This choice led them to merge the information from *Quadros de Pessoal* with data from *SABI*, greatly restricting the number of observations used in their estimations.

To overcome this issue, another normalization strategy proposed by Card et al. (2016) was followed - namely the usage of a "low-surplus" industry instead of a group of "low-surplus" firms. The reasoning, however, was similar, as it implied setting the average fixed effects of this industry to 0.[9] The goal was to choose an industry in which there was structurally little rent to share and where, consequently, the wage premiums and wage differential between genders was expected to be virtually 0. As shown in Table A4 in the Appendix, even though the accommodation and restaurant industry (used by Card et al. (2016)) presented the lowest firm effects, it had a considerable gap between the average male and female firm effects. Therefore, assuming that this gap would be 0 could potentially bias the conclusions of this paper. Hence, when considering both the levels of firm effects and the differential between genders, the textile industry was considered the most viable option.

The normalized fixed effects could, therefore, be defined as follows:

$$\psi_{J(i,t)}^{G(i)} = \widehat{\psi_{J(i,t)}^{G(i)}} - \mathbb{E}[\widehat{\psi_{J(i,t)}^{G(i)}}|Textile\ industry] \tag{2}$$

[7]The dual-connected set is composed only by firms which were within the female and male connected sets. Therefore, it eliminates firms that did not have, at least, one mobile female and male worker. For more information, check Table A3 in the Appendix

[8]Card et al. (2016) followed a 2-step approach to normalize the firm effects. Firstly, they arbitrarily normalized them by setting the fixed effects of the largest firm to 0. Then, they grouped firms into percentiles of log value-added per worker and plotted the mean firm effects in each percentile. The attained "hockey stick" figure allowed them to identify a threshold of mean log valued-added which defined firms as being "low-surplus".

[9]They used this approach as a robustness check of their main normalization strategy.

in which $\psi^{G(i)}_{J(i,t)}$ represents the normalized fixed effects, $\widehat{\psi^{G(i)}_{J(i,t)}}$ denotes the estimated effects of Equation (1) and $\mathbb{E}[\widehat{\psi^{G(i)}_{J(i,t)}}|Textile\ industry]$ stands for the average estimated firm effects in the textile industry.

3.3 Decomposition

In order to get a better grasp of the mechanisms behind the role of firms in the gender wage gap, the difference in normalized firm effects was decomposed into bargaining and sorting channels, following the style of the classic Blinder-Oaxaca decomposition (Oaxaca, 1973; Blinder, 1973):

$$[\mathbb{E}_{\psi^M_{J(i,t)}|male}] - [\mathbb{E}_{\psi^F_{J(i,t)}|female}] = \mathbb{E}[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|male] + \mathbb{E}[\psi^F_{J(i,t)}|male] - \mathbb{E}[\psi^F_{J(i,t)}|female] \tag{3}$$

$$[\mathbb{E}_{\psi^M_{J(i,t)}|male}] - [\mathbb{E}_{\psi^F_{J(i,t)}|female}] = \mathbb{E}[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|female] + \mathbb{E}[\psi^M_{J(i,t)}|male] - \mathbb{E}[\psi^M_{J(i,t)}|female] \tag{4}$$

in which $[\mathbb{E}_{\psi^M_{J(i,t)}|male}]$ is the average pay premium received by men and $[\mathbb{E}_{\psi^F_{J(i,t)}|female}]$ is the average pay premium received by women. The first term on the right-hand side of Equation (3) - $\mathbb{E}[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|male]$- represents the average bargaining effect, which captures the gender difference in the ability to extract rents from firms. It can, therefore, be computed by comparing $\psi^M_{J(i,t)}$ and $\psi^F_{J(i,t)}$ across the distribution of jobs held by men. On the other hand, the second term - $\mathbb{E}[\psi^F_{J(i,t)}|male]$ - $\mathbb{E}[\psi^F_{J(i,t)}|female]$ - denotes the sorting effect, which results from women being commonly sorted into lower-paying firms. This effect can be calculated by comparing the average of $\psi^F_{J(i,t)}$ in jobs held by men versus jobs held by women.

Alternatively, the difference in firm effects could be decomposed according to Equation (4). In this case, the bargaining effect is the difference between the gendered firm effects across the distribution of jobs held by women, while the sorting effect results from the comparison of male firm effects in jobs held by men versus jobs held by women. In order to check the robustness of the results, both decompositions will be presented in section 5.2.

Additionally, it is important to note that in this type of decomposition the choice of a reference group for the normalization is not insignificant for the results (Oaxaca and Ransom, 1999). While the estimated sorting effects are, in principle, invariant of the chosen reference group, the bargaining effects are not - as taking different constants from the fixed effects will undoubtedly result in different values for the first term of Equations (3) and (4). This issue will be discussed throughout the paper.

4 Data and Descriptive Analysis

This study resorted to *Quadros de Pessoal ("Personnel Records")* - a longitudinal matched employer-employee data set that results from an annual compulsory employment survey collected by the Portuguese Ministry of Employment. The survey includes all firms with at least one paid employee, covering almost all Portuguese wage earners - with the exception of civil servants and independent contractors.

The data set provides detailed information on firms (such as location, industry and sales) and workers (including gender, age, education, occupation and earnings). To be more precise, data on earnings encompasses not only base wages, but also irregular and regular benefits, overtime pay and details on collective agreements. The compulsory nature of the survey, as well as the fact that earnings are reported by employers, attenuate common problems such as panel attrition and measurement error (Cardoso et al. 2013).

As discussed in section 3.2, the main reason behind our choice of normalization strategy was the possible gains in terms of data. As a result of not matching data from *Quadros de Pessoal* with another data set, we were able not only to use all observations within the period studied by

Card et al. (2016), but also to expand the number of years covered by the study.[10] This allowed for a greater representativeness of the sample used in the estimations, which could provide more accurate results. This is also important due to the formulation of the connected sets - since having less observations in the overall sample will impact the percentage of person-year observations included in the largest connected sets.

The main sample used in this study covers the period between 1994[11] and 2020, including all individuals aged from 18 to 64 years old, which worked full schedule and received more than 80% of the minimum wage. The wage level was measured through real hourly wages, which were adjusted for 1985 prices. This resulted in an overall sample of 46,343,946 person-year observations, 5,325,654 individuals and 1,162,880 firms. The male connected set counted with 25,677,303 observations, 2,878,243 males and 518,886 firms, while the female connected set had 19,702,449 observations, 2,263,353 females and 479,432 firms. Lastly, the dual-connected counted with 40,106,381 observations, 4,780,992 individuals and 543,786 firms. Table A6 in the Appendix presents a descriptive summary of the 3 groups.

5 Results

5.1 AKM Estimation

The results of the estimation of Equation (1) within the largest connected set of each gender can be found in Table 1 . The male largest connected set included 98.23% of all person-year observations for male workers, while the largest connected set of females covered 97.52% of female person-year observations. Most importantly, as shown in Table A6, the observations

[10]Card et al. (2016) focused on the period between 2002 and 2009, but financial data from SABI was only reliable after 2006. Additionally, they were only able to match 66% of the person-year observations for the period between 2006-2009.

[11]The year 1994 was chosen in order to maximize the number of years used in the estimation but avoid some initial measurement errors and structural changes in the code of the data set.

within the gendered connected sets have similar characteristics (such as average age and years of education] to the ones included in the main sample of analysis.

The top panel of Table 1 presents a summary of the estimates of the two-way fixed effect model for each gender, while the bottom panel displays their adjusted R-squareds. The covariates used in the estimation included firm, worker and year fixed effects, as well as years of schooling, tenure and the squared terms of both age and tenure.[12]

It is shown that the standard deviation of person (worker) effects is greater than the standard deviation of firm effects and of the covariates of both genders. This might imply that, in both cases, wage inequality between individuals is mainly driven by worker characteristics and those seem to have a higher relative importance for males - in which the standard deviation of worker effects more than doubles the standard deviation of firm effects. Firm effects are still considerably important in both cases.

Table 1: Summary of the Estimated Two-Way Fixed Effects Models

	Males Connected Set	Female Connected Set
Standard deviation of log wages	0.587	0.523
Number of person-year observations	25,675,086	19,700,207
Summary of parameter estimates:		
Number of person effects	2,877,926	2,262,981
Number of firm effects	518,803	479,347
Std. dev.of person effects	0.607	0.389
Std. dev.of firm effects	0.277	0.224
Std. dev.of Xb	0.547	0.493
Correlation of person/firm effects	0.191	0.224
Correlation of male/female firm effects	0.733	
Model fit:		
Adjusted R-squared	0.851	0.888
RMSE	0.227	0.175

Note: Models were estimated using years of schooling, tenure, the quadratic terms of age and tenure, as well as worker, firm and year dummies.

[12]The linear term of age was excluded from the equation in order to avoid collinearity issues with the year-fixed effects.

Another important element to consider is the positive correlation between worker and firm effects, which was 0.19 for males and 0.22 for females. This evidence of a possible positive assortative matching could be an indicator that more productive workers tend to be employed at more productive firms. However, one should be careful when making such a statement. While it seems reasonable to assume that high worker effects might be a synonym of more productive workers, it is not necessarily true that highly productive firms are also high-paying firms. In an imperfectly competitive market, highly productive firms - which usually have better outside options than their competitors - might use their degree of monopsony power to lower the wages of their employees, which would be translated into lower firm effects (Portugal et al., 2020).

Additionally, it is also interesting to notice the high correlation between the male and female firm effects (0.73), especially taking into consideration that they were estimated separately. This correlation could be an indicator that firms that tend to pay higher premiums to males tend to pay higher premiums to females as well. Finally, both models present considerably high adjusted R-squareds.[13]

5.2 Decomposition

The estimated firm effects were normalized according to the strategy described in section 3.2. Then, they were decomposed into sorting and bargaining channels, following the framework set out by Equations (3) and (4). For robustness, both equations were estimated. Table 2 presents the final results, focusing, as previously explained, only on observations within the dual-connected set. The analysis covered the overall connected set, as well as different age and educational groups.

The first column presents the "raw" wage gap between men and women in Portugal, which was

[13]The coefficients of the actual estimation can be found in Tables A7 and A8 in the Appendix, as they were not essential for the analysis.

obtained by taking the difference in the mean log hourly wages of male and female workers. On average, in our sample, women received 23.6% less than their male counterparts, without accounting for observable or unobservable characteristics. The second and third columns present the average male firm premiums among men and the average female premiums among women. In the fourth column, one can see the contribution of firm premiums to the gender wage gap, which is the difference between the 2 previous columns. Men receive, on average, higher wage premiums than women (0.023) which represents 10% of the total gender wage gap.

In the fifth and sixth columns, one can find the estimates of the sorting effect according to the male and female distributions. They were obtained by taking the average difference of either the male or female firm effects across the male and female distributions. Considering both results, the sorting effect seems to be between 0.075 and 0.065, meaning it explains around 31% / 27% of the gender wage gap.

The bargaining effect can be found in the last 2 columns of Table 2. This effect is the difference between male and female firm effects within firms, averaged across one on the gender's distribution. As mentioned previously, while the sorting effect is insensitive to the used normalization strategy, the bargaining effect is not - and the chosen normalization might alter significantly the end results. This seems to be the case of the present paper. Choosing a control "low-surplus" industry to normalize the firm effects granted us a considerably higher number of observations, especially regarding the number of firms. While this might have improved the accuracy of the estimates of the overall firm and sorting effects, it also provided unconvincing bargaining effects when compared to the results of Card et al. (2016). Overall, they obtained very small values for the bargaining effects, which were, on average, 0.015 using the female distribution and 0.003 using the male distribution. In 2 of the subgroups - namely for individuals with less than 30 years and individuals with without a high school diploma - they obtained negative bargaining

Table 2: Contribution of Firm Wage Premiums to the Gender Wage Gap

		Mean of Firm Premiums			Decomposition			
					Sorting		Barganning	
	Gender Wage Gap	Male Premium	Female Premium	Firm components	Male	Female	Male	Female
Overall Sample	0.236	0.151	0.128	0.023	0.075	0.065	-0.042	-0.052
By age group:								
Less than 35years	0.151	0.130	0.126	0.004	0.047	0.051	-0.042	-0.047
Ages 36-50	0.277	0.168	0.135	0.033	0.074	0.081	-0.039	-0.048
Over 50 years	0.317	0.163	0.116	0.047	0.087	0.111	-0.043	-0.054
By education group:								
Less than High school	0.306	0.108	0.069	0.040	0.092	0.090	-0.047	-0.054
High school	0.269	0.202	0.170	0.033	0.068	0.078	-0.037	-0.047
University	0.330	0.287	0.254	0.033	0.050	0.075	-0.026	-0.038

Note: This analysis is only focused on individuals within the dual-connected set. The first column shows the raw gender wage gap, while the second and third columns present the average male and female premiums, respectively. Column 4 is the difference between the 2 previous columns. Columns 5 and 6 show the sorting effects using the male and female distribution, while columns 7 and 8 show the bargaining effects. For more details, please resort to the text.

effects. On the other hand, this study obtained negative bargaining effects for all subgroups. The negative results imply that, on average, women are able to capture a higher share of the rents than men - implying that the great driver of the gender wage gap caused by firms is the sorting effect. The bargaining effects are very similar across all subgroups and between the male and female distribution, ranging from -0.054 to -0.0026.

The second panel of Table 2 displays the analysis by age group. The raw gender wage gap visibly widens with age, ranging from 15% for individuals with less than 35 years to 31% for workers over 50. The role of firms in this gap also increases as individuals get older, accounting only for 2% of the pay gap for younger workers and 15% for the eldest group. This seems to be mostly due to an increase in the sorting effects across age groups, which account for 31%/35% of the wage gap for younger individuals and 27%/35% for those over 50. There are many possible explanations for this - for example, older women might be more influenced by social norms and, therefore, might tend to work in female-dominated industries with lower firm effects. The bargaining effects remain negative for all groups and, counter-intuitively, decrease with age.

The last panel of Table2 shows how the effects change for different levels of education. There doesn't seem to be a clear relationship between education and the gender wage gap, as it is wider for individuals with higher education (33%) and for those who didn't finish high school (30.6%) than it is for workers with just a high school diploma (26.9%). There is, however, a trend in terms of the relative role that firms play in these gaps - with firm components representing 13% of the pay gap between individuals without a high school diploma, 12% of the gap between individuals who finished high school and 10% of the gender wage gap between university-educated workers. Once again, the fact that wage premiums for both men and women increase with education level is not a sufficient condition to say that there is necessarily positive assortative matching between highly productive workers and firms. The greatest contributor to the role of firms in the gender wage gap is, once more, the sorting effect - which also decreases with the level of education. The bargaining effects, which remained negative, were still higher for individuals with higher levels of education. If the results were positive, this could be a sign of the so-called ”glass-ceiling”, which affects women’s chances of rising to senior-level positions and capture higher firm effects (or at least the same as their male counterparts).

5.3 Sorting : Robustness check

This paper illustrated that the results of the bargaining effects following the methodology of Card et al. (2016) are, indeed, very sensitive to the normalization procedure. The estimates of that variable do seem counter-intuitive, as they are, on average, negative across the sample. Even though some replications of this method in other countries, such as France (Coudin et al., 2018), also had negligible and negative bargaining effects, it does not seem to be a good sign for the robustness of the overall conclusions of this study.

The sorting effects, contrarily to bargaining, are not sensitive to different normalization strategies. Hence, they are a good way of testing whether our overall results (such as the role of firm components on the gender wage gap) are reliable. Comparing our results to the ones obtained by Card et al. (2016) is not the best strategy - as despite using the same information source, our sample included a considerably higher number of observations, which led to different average firm effects. Therefore, we opted for using a different method to compute the sorting effects - namely by following a Gelbach (Gelbach, 2016) decomposition of the firm effects.

Firstly, the firm effects were estimated following the methodology set out in section 3.1. This time, however, Equation (1) was estimated using observations within the overall dual-connected set instead of being estimated separately for each gender. Consequently, the estimated firm effects were able to capture the impact of wage premiums for our overall sample. This procedure was repeated for each age and educational group present in Table 2.

Then, in order to find the sorting effect, a simple OLS regression was used:

$$\widehat{\psi_{J(i,t)}} = \alpha + X'_{it}\beta + \gamma Male_i + r_{it} \tag{5}$$

in which $\widehat{\psi_{J(i,t)}}$ represents the estimated firm effects in the dual-connected set; $X'_{it}\beta$ is a set of covariates and *Male* is a gender dummy. It is important to mention that $X'_{it}\beta$ contains all the covariates used in the AKM estimation of the fixed effects (squared term of age, tenure, squared term of tenure, and education years), as well as year dummies and the linear term of age.

The coefficient of the gender dummy - γ - captures the sorting effect since it measures, on average, ceteris paribus, how much the male firm effects are higher than the female firm effects given the set of observable characteristics. Table 3 compares the results of the Oaxaca decomposition proposed by Card et al. (2016) with the ones from the Gelbach decomposition.

Table 3: Comparison between the Oaxaca-Blinder and Gelbach sorting effects

	Sorting		
	Card et al.(2016) method		Gelbach
	Male distribution	Female distribution	
Overall Sample	0.075	0.065	0.078
By age group:			
Less than 35years	0.047	0.051	0.057
Ages 36-50	0.074	0.081	0.073
Over 50 years	0.087	0.111	0.075
By education group:			
Less than High school	0.092	0.090	0.102
High school	0.068	0.078	0.077
University	0.050	0.075	0.030

Note: Columns (1) and (2) show the sorting effect taken from Table 2 while the last column shows the coefficient of the gender dummy of Equation (5)

The sorting effects seem to be similar regardless of the estimation methods or the gender distribution taken into consideration, which corroborates the idea that the results of this paper - with the notable exception of the bargaining effects - are robust.

5.4 Gender Mobility Gap

As discussed in previous sections, the sorting of women into lower-paying firms is one of the main drivers of the gender wage gap in Portugal. Cardoso et al. (2016) delved deeper into the possible mechanisms behind this phenomenon, arguing that women might be less efficient in their job search than their male counterparts. This can be a consequence of women underestimating the true distribution of wage offers - either because they expect to be discriminated against or because they underestimate their own market value, having, therefore, lower reservation wages.

Casarico and Lattanzio (2019) presented another possible contributing factor to the sorting channel - the existence of a gender mobility gap. They focused their analysis on mobile workers, comparing the average wage premiums of previous and current firms, in other to understand if men were more likely then women to move to firms with more generous premiums than their previous employeer.

This section aims at making a similar investigation for Portugal, using a simple probit model :

$$P[Y_i = 1|X_i] = \Phi(\alpha + \gamma Female_i + X'_{it}\beta + \lambda_s + \delta_t) \tag{6}$$

in which

$$Y_i = \begin{cases} 1 & \text{if } \psi^{G(i)}_{J1(i,t)} > \psi^{G(i)}_{J0(i,t)} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The dependent variable Y_i represents the improvement, in terms of wage premium, that an individual had from switching firms. Therefore, as explicit in equation (7), Y_i is equal to 1 when the average wage premium of the current firm ($\psi^{G(i)}_{J1(i,t)}$) is higher than the average premium of the previous firm ($\psi^{G(i)}_{J0(i,t)}$) and 0 otherwise. $X'_{it}\beta$ is composed by a set of covariates (such as education, age and tenure) while λ_s and δ_t represent, respectively, sector and year fixed effects. Finally, $Female_i$ is a gender dummy which will be the center piece of this analysis.

Table 4: Average Marginal Effects from Probit Estimation

	(1) 0
Female	-0.0160***
	(-21.12)
Age	0.00224***
	(8.71)
Age^2	-0.0000195***
	(-5.84)
Tenure	0.0115***
	(63.82)
$Tenure^2$	-0.000308***
	(-45.85)
Education	-0.00227***
	(-20.94)
N	1,972,919

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As it is possible to see in Table 4, the average marginal effect of the gender dummy is equal to -0.016 and it is significant at a 1% significance level. This means that, on average, women are 1.6 percentage points less likely to move to firms with higher average wage premiums than males. This not only provides evidence for one possible mechanism behind the sorting effect, but might also contribute to the understanding of the bargaining channel - at least in theory, since our bargaining estimates were not robust. Simply put, if women are less likely to move to a firm with a more generous pay policy, their current firms will be able to extract more rents from them then from their male colleagues - which have better exist options (Casarico and Lattanzio, 2019).

A logit regression following a similar specification to the one of Equation (6) was run in order to confirm the robustness of the results.[14] As expected, the coefficient of the gender dummy was similar to the one obtained with the probit regression (-0.0159) and it was also significant at a 1% significance level. The remaining results can be found on Table A9 in the Appendix.

Finally, it is important to understand that this is still far from a perfect measure of the gender mobility gap. Firstly, it does not explain *why* women are less likely than males to move to higher-paying firms when they switch jobs. Additionally, it also fails to consider the different types of mobility - since workers can voluntary switch firms or be force to do so, either due to a dismissal or firm closure. Therefore, there is still space for improvement for future research on the matter.

6 Conclusion

There is an increasing interest in understanding the role of firms in the gender wage gap, especially in what concerns the importance of firm-specific wage premiums. This paper aimed at contributing to the growing literature on the matter by building on the work of Card et al. (2016) and their analysis of the impact of firm components on the Portuguese pay gap. The main contribution was the choice of a different normalization strategy, which allowed for a greater number of individuals and years to be included in the estimations. However, as it was discussed throughout the paper, the methodology proposed by Card et al. (2016) was not as robust as expected and this small change led to unconvincing bargaining effects.

[14]Due to the distribution of our dependent variable, this step was not necessary, as it was not expected for the 2 types of regressions to give different results.

Overall, firm components accounted for 9.8% of the gender wage gap in Portugal, being increasingly relevant for older individuals and for those with fewer years of education. The sorting effect was the main driver of firm components accounting, on its own, for 31.7% of the pay gap.[15] Sorting is also more meaningful for less educated individuals. The bargaining effects, which are by definition sensitive to the applied normalization strategy, were, on average, negative for all age and educational groups. In their original paper, Card et al., 2016 argued that, despite the sensitivity of the bargaining effects and the arbitrary normalization used, the bargaining estimates were still robust. The current study could be taken as evidence that, indeed, arbitrary normalizations lead to different results and a better method is necessary in order to correctly identify the bargaining channel. A robustness check was done for the sorting effects - using a Gelbach decomposition - in order to understand if the differences from the original paper came from the normalization procedure or from an incorrect estimation. The check corroborated the results of the sorting effects. Additionally, following the ideas and methodology of Casarico and Lattanzio (2019) it was found that, on average, women were 1.6 percentage points less likely than men to move to firms with higher wage premiums than their previous firm. This could be a potential driver of both sorting and bargaining effects.

To conclude, the present paper provides robust estimates of the total contribution of firm components to the gender wage gap, as well as for the importance of the sorting channel in the matter. Additionally, it also emphasizes the need for other measures of bargaining effects, as the linear additive structure of the AKM model might not be the best to truly grasp the way firms extract rents from male and female workers. A better measure could be to directly compare the arrival rates of job offers of men and women and understand how that would impact the average wage premium received. Indirectly, it could also be useful to look at the new monopsony

[15]Please take into consideration that i) the bargaining effects were negative and ii) the decomposition of firm components into sorting and bargaining holds at the *individual* level. In Card et al. (2016) it is also possible to observe certain sorting effects which were greater than the overall firm components.

literature, namely at the estimations of the labour supply curves facing individual firm and compare the elasticilities of both genders. Ideally, one could combine more than one method in order to obtain more robust conclusions - which could then be used to tackle the issue of gender inequality in the labour market.

References

Abowd, J. M., R. H. Creecy, F. Kramarz, et al. (2002). Computing person and firm effects using linked longitudinal employer-employee data. Technical report, Center for Economic Studies, US Census Bureau.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature 55*(3), 789–865.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436–455.

Bruns, B. (2019). Changes in workplace heterogeneity and how they widen the gender wage gap. *American Economic Journal: Applied Economics 11*(2), 74–113.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly journal of economics 131*(2), 633–686.

Cardoso, A. R., P. Guimarães, and P. Portugal (2013). Everything you always wanted to know about sex discrimination. *Available at SSRN 2199792*.

Cardoso, A. R., P. Guimarães, and P. Portugal (2016, 01). What drives the gender wage gap? A look at the role of firm and job-title heterogeneity. *Oxford Economic Papers 68*(2), 506–524.

Casarico, A. and S. Lattanzio (2019). What firms do: Gender inequality in linked employer-employee data.

Coudin, E., S. Maillard, and M. Tô (2018). Family, firms and the gender wage gap in france. Technical report, IFS Working Papers.

Cruz, G. and T. Rau (2017). The effects of firms' pay policies and equal pay laws on the gender wage gap in chile. *Unpublished manuscript*.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Gerhart, B. and S. Rynes (1991). Determinants and consequences of salary negotiations by male and female mba graduates. *Journal of Applied Psychology 76*(2), 256.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American economic review 96*(2), 1–21.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic perspectives 20*(4), 133–156.

Hirsch, B., T. Schank, and C. Schnabel (2010). Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from germany. *Journal of Labor Economics 28*(2), 291–330.

Li, J., B. Dostie, and G. Simard-Duplain (2022). Firm pay policies and the gender earnings gap: The mediating role of marital and family status. *ILR Review*, 00197939221093562.

Manning, A. (2003). The real thin theory: monopsony in modern labour markets. *Labour economics 10*(2), 105–131.

Masso, J., J. Meriküll, and P. Vahter (2022). The role of firms in the gender wage gap. *Journal of Comparative Economics 50*(2), 454–473.

Mincer, J. and S. Polachek (1974). Family investments in human capital: Earnings of women. *Journal of political Economy 82*(2, Part 2), S76–S108.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

Oaxaca, R. L. and M. R. Ransom (1999). Identification in detailed wage decompositions. *Review of Economics and Statistics 81*(1), 154–157.

Petersen, T. and L. A. Morgan (1995). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology 101*(2), 329–365.

Portugal, P. et al. (2020). The sources of wage variability in portugal: a binge reading survey. *Economic Bulletin and Financial Stability Report Articles and Banco de Portugal Economic Studies*.

Robinson, J. (1969). *The economics of imperfect competition*. Springer.

Small, D. A., M. Gelfand, L. Babcock, and H. Gettman (2007). Who goes to the bargaining table? the influence of gender and framing on the initiation of negotiation. *Journal of personality and social psychology 93*(4), 600.

Sulis, G. (2011). What can monopsony explain of the gender wage differential in italy? *International Journal of Manpower*.

Webber, D. A. (2016). Firm-level monopsony and the gender pay gap. *Industrial Relations: A Journal of Economy and Society 55*(2), 323–345.

Appendix

Table A1: Largest Connected Set - Males

	Freq.	Percent
0	463,624	1.77
1	25,677,303	98.23
Total	26,140,927	100.00

Table A2: Largest Connected Set - Females

	Freq.	Percent
0	500,570	2.48
1	19,702,449	97.52
Total	20,203,019	100.00

Table A3: Dual Largest Connected Set

	Freq.	Percent
0	3,481	0.01
1	40,106,381	99.99
Total	40,109,862	100.00

Table A4: Average Premia per sector

	F_Males	F_Females
Agriculture,Animal Production, Hunting and Forestry	-.1222934	-.0667355
Fishing	.1436812	.1437068
Extraction of Energy Products	.3767246	.5686677
Extractive Industries Other than Extraction of Energy Products	.1281945	.1632107
Food, Beverage and Tobacco Industries	-.0196939	-.0312894
Textile Industry	-.1229209	-.1109729
Leather and Leather Products Industry	-.1641574	-.0948525
Wood and Cork Industries	-.0503749	.0126688
Paper and Cardboard Pulp, Publishing and Printing Industries	.1007967	.0879026
Manufacture of Coke, Refined Petroleum Products and Nuclear Fuel	.7604408	.662201
Manufacture of Chemicals and Synthetic or Artificial Fibers	.1985386	.1894739
Manufacture of Rubber and Plastic Articles	.0605574	.0344145
Manufacture of Other Non-Metallic Mineral Products	.0776095	.0466779
Base Metallurgical Industries and Metallic Products	-.008462	.0172761
Machinery and N.E Equipment Manufacture	.02184	.0422139
Manufacture of Electrical and Optical Equipment	.1147243	.1821519
Manufacturing of Transport Material	.1044071	.0882335
N.E Manufacturing Industries	-.1565679	-.0711455
Production and Distribution of Electricity, Gas and Water	.3169713	.3544882
Construction	-.0004502	.0367354
Wholesale and Retail Trade, Repair of Motor Vehicles and Motorcycles	-.0296257	.0195723
Accommodation and Catering (Restaurants and Similar)	-.1696629	-.0671103
Transport, Storage and Communications	.2103402	.2963272
Financial Activities	.4060593	.4085441
Real Estate Activities, Leases and Services Provided to Companies	.0501952	.0650599
Public Administration, Defense and Mandatory Social Security	-.0661179	.0822117
Education	-.0361958	-.0062058
Health and Social Action	-.1174644	-.0536701
Other Collective, Social and Personal Service Activities	.0579822	.024107
International Organizations and Other Extra-Territorial Institutions	.2098261	.4402862

Table A5: Number of Observations Within the Different Samples

	Overal Sample		Connected Sets		Dual-Connected Set	
	Male	Female	Male	Female	Male	Female
N of person-year	26,140,927	20,203,019	25,677,303	19,702,449	22,567,309	17,539,072
N persons	2,970,319	2,355,335	2,878,24	2,263,353	2,652,809	2,128,183
N firms	598,644	564,236	518,886	479,432	271,893	271,893

Table A6: Descriptive Statistics of the Different Samples Used (QP), 1994-2020

	Overall Sample		Connected Sets		Dual-Connected Set	
	Males	Females	Males	Females	Males	Females
Mean age	39.06	37.97	39.04	37.88	39.10	37.93
Proportion with less than 30 years	0.26	0.28	0.26	0.28	0.25	0.28
Proportion with more than 50 years	0.20	0.16	0.20	0.16	0.20	0.16
Mean years of schooling	8.62	9.50	8.62	9.52	8.85	9.60
Proportion with high school diploma	0.33	0.44	0.33	0.45	0.36	0.46
Proportion with university diploma	0.09	0.12	0.09	0.12	0.10	0.13
Mean log real hourly wage	0.57	0.35	0.58	0.36	0.63	0.39
	(0.59)	(0.52)	(0.59)	(0.52)	(0.59)	(0.53)
Mean monthly hours	170.64	169.27	170.61	169.21	170.41	169.00
Mean firm size (no. employees)	866.43	993.95	882.01	1,019.10	1,002.31	1,143.78
Average share of female workers in companies	0.27	0.66	0.27	0.66	0.30	0.62
	(0.22)	(0.26)	(0.22)	(0.26)	(0.22)	(0.25)
Observations	26,140,927	20,203,019	25,677,303	19,702,449	22,567,309	17,5390,72

Table A7: Coefficients of AKM Estimation Using the Male Largest Connected Set

	Log of real hourly wage
Age^2	-0.000387***
	(5.79e-07)
Tenure	0.00924***
	(2.53e-05)
$Tenure^2$	-0.000155***
	(7.52e-07)
Education	0.00614***
	(3.87e-05)
Constant	1.102***
	(0.00106)
Observations	25,675,086
R-squared	0.870

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A8: Coefficients of AKM Estimation Using the Female Largest Connected Set

	Log of real hourly wage
Age^2	-0.000171***
	(5.70e-07)
Tenure	0.0115***
	(2.47e-05)
$Tenure^2$	-0.000238***
	(7.49e-07)
Education	0.00692***
	(3.75e-05)
Constant	0.495***
	(0.000997)
Observations	19,700,207
R-squared	0.892

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A9: Marginal Average Effects of Logit Estimation

	(Average Marginal Effects)
Female	-0.0159***
	(-21.05)
Age	0.00221***
	(8.60)
Age^2	-0.0000192***
	(-5.77)
Tenure	0.0116***
	(64.26)
$Tenure^2$	-0.000311***
	(-45.86)
Education	-0.00228***
	(-21.04)
N	1,972,919

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$